

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 586 210**

51 Int. Cl.:

C12N 15/62 (2006.01)

C07K 14/00 (2006.01)

C12N 15/82 (2006.01)

C12N 15/09 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.12.2007 E 11184934 (5)**

97 Fecha y número de publicación de la concesión europea: **11.05.2016 EP 2415872**

54 Título: **Proteínas de dedo de zinc no canónicas optimizadas**

30 Prioridad:

14.12.2006 US 874911 P

30.05.2007 US 932497 P

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

13.10.2016

73 Titular/es:

SANGAMO BIOSCIENCES, INC. (50.0%)
Point Richmond Tech. Center II 501 Canal
Boulevard, Suite A100
Richmond, CA 94804, US y
DOW AGROSCIENCES, LLC (50.0%)

72 Inventor/es:

CAI, QIHUA C.;
MILLER, JEFFREY;
URNOV, FYODOR;
SHUKLA, VIPULA K.;
PETOLINO, JOSEPH F.;
BAKER, LISA W.;
GARRISON, ROBBI J.;
BLUE, RYAN C.;
MITCHELL, JON C.;
ARNOLD, NICOLE L. y
WORDEN, SARAH E.

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 586 210 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

Proteínas de dedo de zinc no canónicas optimizadas

Campo técnico

5 La presente descripción se encuentra en los campos de ingeniería genómica, direccionamiento génico, integración cromosómica dirigida, expresión de proteínas y edición epigenómica.

Antecedentes

10 La unión de proteínas a secuencias específicas de ADN, ARN, proteínas y otras moléculas participa en diversos procesos celulares, tales como, por ejemplo, la transcripción, la replicación, la estructura de la cromatina, la recombinación, la reparación del ADN, el procesamiento del ARN y la traducción. La especificidad de unión de las proteínas de unión celulares que participan en las interacciones entre proteína-ADN, proteína-ARN y proteína-proteína contribuye al desarrollo, la diferenciación y la homeostasis.

15 Las proteínas con dedos de zinc (ZFP, acrónimo inglés de *Zinc Finger Proteins*) son proteínas que pueden unirse al ADN de una manera específica de secuencia. Los dedos de zinc se identificaron por primera vez en el factor de transcripción TFIIIA de los ovocitos del sapo africano con uñas, *Xenopus laevis*. Un solo dominio de dedo de zinc de esta clase de ZFP tiene una longitud de aproximadamente 30 aminoácidos, y diversos estudios estructurales han demostrado que contiene un giro beta (que contiene dos restos de cisteína conservados) y una hélice alfa (que contiene dos restos de histidina conservados), que están sujetos en una conformación particular a través de la coordinación de un átomo de zinc con las dos cisteínas y las dos histidinas. Esta clase de ZFP también se conoce como ZFP C2H2. También se han propuesto otras clases de ZFP. Véase, por ejemplo, Jiang *et al.* (1996) J. Biol. Chem. 271: 10723-10730 para un análisis de ZFP de tipo Cys-Cys-His-Cys (C3H). Hasta ahora, se han identificado más de 10.000 secuencias de dedos de zinc en diversos miles de factores de transcripción conocidos o hipotéticos. Los dominios de los dedos de zinc están implicados no solo en el reconocimiento del ADN, sino también la unión al ARN y en la unión de proteína-proteína. Las estimaciones actuales son que esta clase de moléculas constituirá aproximadamente el 2 % de todos los genes humanos.

25 La mayoría de las proteínas con dedos de zinc han conservado restos de cisteína e histidina que coordinan tetraédricamente con el único átomo de zinc de cada dominio del dedo. En particular, la mayoría de las ZFP se caracterizan por componentes de los dedos de la secuencia general: -Cys(X)₂₋₄-Cys(X)₁₂-His(X)₃₋₅-His- (SEQ ID NO: 1), donde X representa cualquier aminoácido (las ZFP C₂H₂). Las secuencias de coordinación de zinc de esta clase más ampliamente representada contienen dos cisteínas y dos histidinas con espacios particulares. La estructura plegada de cada dedo contiene un giro β antiparalelo, una región punta de dedo y una hélice α anfipática corta. Los ligandos de coordinación del metal se unen al ion zinc y, en el caso de los dedos de zinc de tipo zif268 la hélice α anfipática corta se une en el surco mayor del ADN. Además, los aminoácidos hidrófobos conservados (por ejemplo, el resto que precede directamente a la primera Cys conservada y el resto en la posición +4 del segmento helicoidal del dedo) y la coordinación del zinc por los restos de cisteína e histidina conservados, estabilizan la estructura del dedo de zinc.

35 Se han descrito proteínas con dedos de zinc (C2H2) canónicas que tienen alteraciones en posiciones que establecen contactos directos con la base, restos de 'apoyo' o 'yuxtaposición' inmediatamente adyacentes a las posiciones de contacto con la base, y posiciones capaces de establecer contacto con el esqueleto de fosfato del ADN. Véanse, por ejemplo, las patentes de Estados Unidos n.º 6.007.988; 6.013.453 6.140.081 6.866.997 6.746.838 6.140.081 6.610.512 7.101.972 6.453.242 6.785.613 7.013.219; PCT WO 98/53059; Choo *et al.* (2000) Curr. Opin. Struct. Biol. 10: 411-416; Segal *et al.* (2000) Curr. Opin. Chem. Biol. 4:34-39.

45 Además, también se han descrito proteínas con dedos de zinc que contienen dedos de zinc con restos de coordinación de zinc modificados (véanse, por ejemplo, las solicitudes de patente de Estados Unidos n.º 20030108880, 20060246567 y 20060246588; cuyas descripciones se incorporan por referencia). Sin embargo, aunque las proteínas con dedos de zinc que contienen estos dedos de zinc no canónicos conservan la función reguladora de la transcripción génica, su capacidad para actuar como nucleasas con dedos de zinc (ZFN, acrónimo inglés de *Zinc Finger Nucleases*) está, en algunos casos, disminuida con respecto a las proteínas con dedos de zinc que constan exclusivamente de dedos de zinc C2H2, canónicos.

50 El documento WO 02/057293 A2 desvela proteínas de unión de dedo de zinc con sitios de coordinación de metal modificados, mediante lo cual las proteínas de unión contienen tres dedos de coordinación de zinc y uno o más de estos dedos son componentes de dedo no canónicos, modificados.

El documento WO 2005/014791 A2 se refiere a polipéptidos de fusión que comprenden un dominio de unión de dedo de zinc modificado por ingeniería genética y un dominio de escisión, y a polipéptidos de fusión que comprenden un dominio de unión de dedo de zinc modificado por ingeniería genética y el semidominio de escisión.

55 Caravec *et al.* (Molecular and Cellular Biology, Vol. 17, n.º 1, 1997, págs. 482-494) desvelan un dedo de zinc CCHC que tiene la secuencia 945-CRHC GK KYR WK ST LR RH EN VE CG G-969.

Yang *et al.* (Chinese Journal of Biochemistry and Molecular Biology, vol. 20, 2004, págs. 725-731) se refieren a un gen de ratón denominado BSG2. La proteína predicha contiene 9 dedos de zinc C2H2 típicos en su región C terminal. No hay información acerca de si los dedos de zinc participan en la unión al ADN.

- 5 Por tanto, sigue habiendo una necesidad, particularmente en la construcción de nucleasas con dedos de zinc, de proteínas de unión de dedo de zinc modificadas por ingeniería genética adicionales que contengan dedos de zinc que tengan regiones de coordinación de zinc no canónicas, optimizadas.

Compendio

La presente invención está definida por el conjunto de reivindicaciones adjuntas.

- 10 La presente descripción proporciona dominios de unión de ADN de dedo de zinc con alteraciones en al menos un resto de coordinación de zinc. En particular, en la presente memoria se describen dedos de zinc CCHC. También se describen polipéptidos y proteínas de fusión con dedos de zinc que comprenden uno o más de estos dedos de zinc CCHC, polinucleótidos que codifican estos dedos de zinc y proteínas de fusión y métodos de uso de estos polipéptidos y/o proteínas de fusión con dedos de zinc.

Breve descripción de los dibujos

- 15 La **figura 1** es un gráfico que representa tasas de corrección génicas, medidas por el porcentaje de células que expresan GFP, en un sistema de ensayo indicador de células con GFP como se describe en la patente de Estados Unidos n.º 2005/0064474 y más adelante. Las variantes ZFN se designan con "X-Y," donde "X" se refiere al número de tabla e "Y" se refiere al número dado al dedo de zinc en la tabla particularmente seleccionada. Por ejemplo, "2-21" se refiere a una ZFN que tiene un dedo que comprende la secuencia mostrada en la tabla 2 en la fila numerada 20 21, concretamente HAQRCGLRGSQVL (SEQ ID NO:53).

- La **figura 2** es un gráfico que representa el porcentaje de señal Cel-1 resultante de la escisión usando diversos pares de variantes ZFN. Se muestran los resultados de dos experimentos para cada par de ZFN por referencia al número de muestra. Los pares de variantes usados para cada muestra se muestran en el recuadro en el ángulo derecho superior, donde "ts 5-8" y "ts 5-9" se refiere a pares de ZFN canónicos desvelados en el ejemplo 14 (tabla 25 17) de la solicitud de patente de Estados Unidos n.º 2005/0064474. En las muestras 3-12, la región C terminal de las hélices de reconocimiento del dedo 2 o dedo 4 de las 5-8 o 5-9 ZFN canónicas, está reemplazada con secuencias no canónicas. La secuencia parcial de las variantes ZFN no canónicas designadas 20, 21, 43, 45, 47 y 48 en las muestras 3-12 y la posición del dedo de estas variantes dentro de las ZFN de cuatro dedos se muestra en el ángulo izquierdo superior por encima del gráfico. El asterisco encima de la barra que representa resultados del experimento 30 2 para las muestras 8 y 9 indica el fondo en el carril, que producen una subestimación de la eficacia de ZFN.

- La **figura 3** es un gráfico que representa tasas de corrección génica en el sistema de ensayo indicador de células con GFP descrito en la patente de Estados Unidos n.º 2005/0064474 y en la presente memoria. Los pares ZFN ensayados en cada muestra se muestran debajo de cada barra, donde los números de dedos de zinc 20, 21, 43, 45, 35 47 y 48 son los descritos en el ejemplo 3 y los dedos de zinc CCHC 1a a 10a comprenden la secuencia mostrada en las tablas 3 y 4. Los dedos de zinc 20, 21, 7a, 8a, 9a y 10a se usaron en el dedo 4; los dedos de zinc 43, 45, 47, 48, 1a, 2a, 3a, 4a, 5a y 6a se usaron en el dedo 2.

La **figura 4** es una representación esquemática lineal del plásmido pDAB1585, un vector diana para el tabaco.

La **figura 5** es una representación esquemática del plásmido pDAB1585, un vector diana para el tabaco.

- 40 Las **figuras 6A** y **6B** representan nucleasas con dedos de zinc (ZFN). La FIG. 6A es un esquema que representa la unión de ZFN. La FIG. 6B muestra la secuencia de la secuencia diana.

La **figura 7** es una representación esquemática del plásmido pDAB1400.

La **figura 8** es una representación esquemática del plásmido pDAB782.

La **figura 9** es una representación esquemática del plásmido pDAB1582.

La **figura 10** es una representación esquemática del plásmido pDAB354.

- 45 La **figura 11** es una representación esquemática del plásmido pDAB1583.

La **figura 12** es una representación esquemática del plásmido pDAB2407.

La **figura 13** es una representación esquemática del plásmido pDAB1584.

La **figura 14** es una representación esquemática del plásmido pDAB2418.

La **figura 15** es una representación esquemática del plásmido pDAB4045.

- La **figura 16** es una representación esquemática del plásmido pDAB1575.
- La **figura 17** es una representación esquemática del plásmido pDAB1577.
- La **figura 18** es una representación esquemática del plásmido pDAB1579.
- La **figura 19** es una representación esquemática del plásmido pDAB1580.
- 5 La **figura 20** es una representación esquemática del plásmido pDAB3401.
- La **figura 21** es una representación esquemática del plásmido pDAB1570.
- La **figura 22** es una representación esquemática del plásmido pDAB1572.
- La **figura 23** es una representación esquemática del plásmido pDAB4003.
- La **figura 24** es una representación esquemática del plásmido pDAB1571.
- 10 La **figura 25** es una representación esquemática del plásmido pDAB7204.
- La **figura 26** es una representación esquemática del plásmido pDAB1573.
- La **figura 27** es una representación esquemática del plásmido pDAB1574.
- La **figura 28** es una representación esquemática del plásmido pDAB1581.
- La **figura 29** es una representación esquemática del plásmido pDAB1576.
- 15 La **figura 30** es una representación esquemática del plásmido pDAB1600.
- La **figura 31** es una representación esquemática del plásmido pDAB3731.
- La **figura 32** es una representación esquemática del plásmido pDAB4322.
- La **figura 33** es una representación esquemática del plásmido pDAB4331.
- La **figura 34** es una representación esquemática del plásmido pDAB4332.
- 20 La **figura 35** es una representación esquemática del plásmido pDAB4333.
- La **figura 36** es una representación esquemática del plásmido pDAB4334.
- La **figura 37** es una representación esquemática del plásmido pDAB4336.
- La **figura 38** es una representación esquemática del plásmido pDAB4339.
- La **figura 39** es una representación esquemática del plásmido pDAB4321.
- 25 La **figura 40** es una representación esquemática del plásmido pDAB4323.
- La **figura 41** es una representación esquemática del plásmido pDAB4341.
- La **figura 42** es una representación esquemática del plásmido pDAB4342.
- La **figura 43** es una representación esquemática del plásmido pDAB4343.
- La **figura 44** es una representación esquemática del plásmido pDAB4344.
- 30 La **figura 45** es una representación esquemática del plásmido pDAB4346.
- La **figura 46** es una representación esquemática del plásmido pDAB4330.
- La **figura 47** es una representación esquemática del plásmido pDAB4351.
- La **figura 48** es una representación esquemática del plásmido pDAB4356.
- La **figura 49** es una representación esquemática del plásmido pDAB4359.
- 35 La **figura 50** es una representación esquemática del plásmido pDAB7002.
- La **figura 51** es una representación esquemática del plásmido pDAB7025.
- La **figura 52** es una representación esquemática del plásmido pDAB1591.

- La **figura 53** es una representación esquemática del plásmido pcDNA3.1-SCD27a-L0-FokI, el molde de ADN usado para la amplificación por PCR de Scd27 ZFN.
- La **figura 54** es una representación esquemática del plásmido pDAB1594.
- La **figura 55** es una representación esquemática del plásmido pDAB1598:
- 5 La **figura 56** es una representación esquemática del plásmido pDAB1577.
- La **figura 57** es una representación esquemática del plásmido pDAB1578.
- La **figura 58** es una representación esquemática del plásmido pDAB1601, el vector de control del gen PAT.
- La **figura 59** es una representación esquemática de recombinación homóloga intracromosómica predicha estimulada por la proteína de fusión IL-1-FokI.
- 10 La **figura 60** es una representación esquemática del plásmido pDAB1590, un control positivo que expresa GFP.
- La **figura 61** es una representación esquemática de recombinación homóloga intercromosómica predicha estimulada por la proteína de fusión IL-1 dedo zinc-FokI.
- La **figura 62** es una representación esquemática de recombinación homóloga intercromosómica predicha estimulada por la proteína de fusión Scd27 dedo zinc-FokI.
- 15 La **figura 63** es un gel que representa el análisis por PCR de los recombinantes. Los primeros 4 carriles de la izquierda se marcan por encima del gel. Los carriles 1-5 marcados muestran acontecimientos de RH de la transformación de BY2-380 con el gen de la proteína de fusión C3H IL-1-FokI y los carriles 6-7 marcados muestran acontecimientos de RH de la transformación de BY2-380 con el gen de la proteína de fusión C3H SCD27-FokI.
- 20 La **figura 64** muestra una secuencia del gen IPP2K del maíz (SEQ ID NO: 6), derivado del cultivo de células Hill y que sirvió como un molde diseñado para la modificación por ingeniería genética de las ZFN dirigidas al IPP2K.
- La **figura 65, paneles A a E**, representa un esquema de clonación del vector de expresión de ZFN. Se usó una estrategia de clonación gradual para generar construcciones de expresión ZFN. Los genes que codifican ZFN individuales se clonaron en los vectores pVAX-N2A-NLSop2-EGFP-FokMono (A) y pVAX-C2A-NLSop2-EGFP-FokMono (B) para crear un casete (C) de proteína doble. Este casete se ligó en pDAB3872 (D) para generar un plásmido final (E) para expresión del heterodímero ZFN.
- 25 La **figura 66** representa la unión de ZFN en un gen IPP2K de maíz. Se necesitan dos proteínas ZFN para llevar a cabo la escisión de ADN bicatenario. Se muestra la secuencia (SEQ ID NO: 7) que rodea el sitio de escisión (indicado con una flecha hacia abajo). Una proteína (8705) se unió a la secuencia CTGTGGGGCCAT (cadena superior) (SEQ ID NO: 8), donde otra proteína (8684, 8685 o 8686) se unió a la secuencia aguas abajo (CTTGACCAACTCAGCCAG, cadena inferior) (SEQ ID NO: 9).
- 30 La **figura 67** representa secuencias de tipo silvestre (secuencia superior, SEQ ID NO: 10) y el clon 127 de ZFN (secuencia inferior, SEQ ID NO: 11). La diana de escisión para esta ZFN se destaca en un recuadro gris.
- La **figura 68** muestra un alineamiento de deleciones múltiples resultante de la unión de extremos no homólogos (NHEJ, acrónimo inglés de *Non-Homologous End Joining*) de una rotura de ADNbc mediada por ZFN en el gen IPP2K del maíz, detectado por secuenciación 454. La diana de escisión para esta ZFN se destaca en un recuadro gris.
- 35 La **figura 69** es un gráfico que representa tasas de corrección génica en el sistema de ensayo indicador de células con GFP descrito en la patente de Estados Unidos n.º 2005/0064474 y en la presente memoria. Los pares ZFN ensayados en cada muestra se muestran debajo de cada barra.
- 40 La **figura 70** representa el plásmido pDAB7471, construido como se describe en el Ejemplo 18B.
- La **figura 71** representa el plásmido pDAB7451, construido como se describe en el Ejemplo 18C.
- La **figura 72** es una representación esquemática de un casete de expresión génica de tolerancia a herbicida autónomo ejemplar. Esta construcción comprende una unidad transcripcional promotora (UTP) completa que contiene un promotor, un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) como se describe en el Ejemplo 18D.
- 45 La **figura 73** representa el plásmido pDAB7422, construido como se describe en el Ejemplo 18E. El plásmido incluye una unidad transcripcional-promotora (UTP) completa que contiene un promotor, un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) insertada en un esqueleto plasmídico en posición 1.

- La **figura 74** representa el plásmido pDAB7452, construido como se describe en el Ejemplo 18E. El plásmido incluye una unidad transcripcional promotora (UTP) completa que contiene un promotor, un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) insertada en un esqueleto plasmídico en posición 2.
- 5 La **figura 75** es una representación esquemática de un casete de expresión génico de tolerancia a herbicida no autónomo. Esta construcción comprende una unidad transcripcional promotora (UTP) incompleta que contiene un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) como se describe en el Ejemplo 18F.
- La **figura 76** representa el plásmido pDAB7423, construido como se describe en el Ejemplo 18G. Este plásmido incluye una unidad transcripcional promotora (UTP) incompleta que contiene un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) insertada en un esqueleto plasmídico en posición 1.
- 10 La **figura 77** representa el plásmido pDAB7454, construido como se describe en el Ejemplo 18G. El plásmido incluye una unidad transcripcional promotora (UTP) incompleta que contiene un gen de tolerancia a herbicidas y una secuencia de terminación de poliadenilación (poliA) insertada en un esqueleto plasmídico en posición 2 como se describe en el Ejemplo 18G.
- 15 La **figura 78** representa el plásmido pDAB 7424 (un donante autónomo en posición 1 adaptado a Gateway® ejemplar), construido como se describe en el Ejemplo 18H.
- La **figura 79** representa el plásmido pDAB 7425 (un donante autónomo en posición 1 adaptado a Gateway® ejemplar), construido como se describe en el Ejemplo 18H.
- 20 La **figura 80** representa el plásmido pDAB 7426, construido como se describe en el Ejemplo 18H. pDAB 7426 es un plásmido de combinación que contiene el donante autónomo en posición 1 con un casete de expresión ZFN.
- La **figura 81** representa el plásmido pDAB 7427, construido como se describe en el Ejemplo 18H. pDAB 7427 es un plásmido de combinación que contiene el donante autónomo en posición 1 con un casete de expresión ZFN.
- La **figura 82** representa la amplificación de secuencias específicas de ADN donante del ADN genómico. La presencia de un producto de 317 pb es diagnóstica para la presencia de ADN donante que contiene el gen PAT insertado en el genoma de líneas de callos de maíz n.º 61 - 72 como se describe en el Ejemplo 20C. Hill indica un control negativo de tipo silvestre.
- 25 La **figura 83** representa la amplificación de límite 5' entre el ADN donante y las secuencias genómicas de maíz específicas para IPP2K. Se diagnosticaron productos de PCR secundarios procedentes de la integración dirigida del donante en el gen IPP2K por la presencia de fragmentos de ADN de 1,65 Kpb como se describe en el Ejemplo 21A. Hill indica un control negativo de tipo silvestre.
- 30 La **figura 84** representa la amplificación de límite 3' entre el ADN donante y las secuencias genómicas de maíz específicas para IPP2K. Se diagnosticaron productos de PCR secundarios procedentes de la integración dirigida del donante en el gen IPP2K por la presencia de fragmentos de ADN de 1,99 Kpb como se describe en el Ejemplo 21A. Hill indica un control negativo de tipo silvestre.
- 35 La **figura 85** representa la amplificación del límite (5') aguas arriba entre el genoma y el donante. Se diagnosticaron productos de PCR procedentes de la integración dirigida del donante en el gen IPP2K (límite 5') por la presencia de fragmentos de ADN de un tamaño de 1,35 Kpb como se describe en el Ejemplo 21B. Hill indica un control negativo de tipo silvestre.
- 40 La **figura 86** representa la amplificación del límite (3'-) aguas abajo entre el donante y el genoma. Se diagnosticaron productos de PCR procedentes de la integración dirigida del donante en el gen IPP2K (límite 3') por la presencia de fragmentos de ADN de un tamaño de 1,66 Kpb como se describe en el Ejemplo 21B. Hill indica un control negativo de tipo silvestre.
- La **figura 87** representa la secuencia del flanco de homología 5' en la posición 1 (SEQ ID NO: 171).
- La **figura 88** representa la secuencia del flanco de homología 3' en la posición 1 (SEQ ID NO: 172).
- 45 La **figura 89** representa la secuencia del flanco de homología 5' en la posición 2 (SEQ ID NO: 139).
- La **figura 90** representa la secuencia del flanco de homología 3' en la posición 2 (SEQ ID NO: 140).
- La **figura 91** representa la secuencia de una secuencia genómica IPP2K (5'-) aguas arriba de las regiones ZFN dirigidas (SEQ ID NO: 141).
- 50 La **figura 92** representa la secuencia de una secuencia genómica de IPP2K (3'-) aguas abajo de las regiones ZFN dirigidas (SEQ ID NO: 142).

Descripción detallada

En la presente memoria se desvelan composiciones que comprenden polipéptidos de unión a dedos de zinc (ZFP) que contienen dedos de zinc no canónicos de formato Cys-Cys-His-Cys como se reivindica. Puesto que la coordinación de zinc proporciona la energía de plegamiento principal para los dedos de zinc, el ajuste de restos de coordinación de zinc proporciona un medio fácil para modificar la estabilidad y la estructura del dedo, que influye sobre una diversidad de características funcionales importantes de las proteínas con dedos de zinc, incluyendo, por ejemplo, la semivida celular, interacciones con otros factores celulares, la especificidad y afinidad de unión del ADN y la orientación relativa de dominios funcionales.

Se ha mostrado que las proteínas con dedos de zinc, que comprenden dedos de zinc no canónicos, tales como los desvelados en las solicitudes de patente de Estados Unidos nos. 20030108880, 20060246567 y 20060246588, se unen al ADN y alteran la transcripción. Sin embargo, cuando se incorporan en nucleasas con dedos de zinc (ZFN, véase, por ejemplo, la solicitud de patente de Estados Unidos publicación N.º 2005/0064474), estas proteínas con dedos de zinc no canónicas previamente descritas pueden presentar, algunas veces, actividad subóptima en la escisión del ADN diana.

En la presente memoria se describen proteínas con dedos de zinc que comprenden uno o más dedos de zinc CCHC como se reivindica, en los que se han alterado secuencias específicas que rodean el par C terminal de los restos que coordinan el zinc. En la presente memoria también se describen proteínas de fusión, por ejemplo, nucleasas con dedos de zinc (ZFN), que comprenden estos de dedos de zinc no canónicos optimizados, en los que las ZFN escinden el ADN diana a tasas comparables a las de la escisión obtenida usando las ZFN que comprenden dedos de zinc (CCHH) canónicos

Como se desvela en la presente memoria, los polipéptidos de fusión pueden potenciar o suprimir la transcripción de un gen y/o escindir una secuencia diana. También se proporcionan polinucleótidos que codifican dedos de zinc no canónicos optimizados y polinucleótidos que codifican proteínas de fusión que comprenden uno o más dedos de zinc no canónicos optimizados. Adicionalmente se proporcionan composiciones farmacéuticas que comprenden una cantidad terapéuticamente eficaz de cualquiera de los polipéptidos de unión a nucleótidos de dedos de zinc descritos en la presente memoria o sus fragmentos funcionales; o una cantidad terapéuticamente eficaz de una secuencia de nucleótidos que codifica cualquiera de los polipéptidos de unión a nucleótidos de dedos de zinc o sus fragmentos funcionales, en combinación con un transportador farmacéuticamente aceptable. Adicionalmente se proporcionan composiciones agrícolas que comprenden una cantidad eficaz desde el punto de vista agronómico de cualquiera de los polipéptidos de unión a nucleótidos de dedos de zinc descritos en la presente memoria o sus fragmentos funcionales; o una cantidad eficaz desde el punto de vista agronómico de una secuencia de nucleótidos que codifica cualquiera de los polipéptidos de unión a nucleótidos de dedos de zinc modificados o sus fragmentos funcionales, en combinación con un transportador aceptable desde el punto de vista agrícola. También se proporcionan métodos de exploración para obtener un polipéptido de unión a nucleótidos de dedos de zinc modificados que se unen a una secuencia genómica.

Las secuencias genómicas incluyen las que están presentes en cromosomas, episomas, genomas organulares (por ejemplo, mitocondria, cloroplasto), cromosomas artificiales y cualquier otro tipo de ácido nucleico presente en una célula tal como, por ejemplo, secuencias amplificadas, cromosomas pequeños dobles y los genomas de bacterias y virus endógenos o infecciosos. Las secuencias genómicas pueden ser normales (*es decir*, de tipo silvestre) o mutantes; las secuencias mutantes pueden comprender, por ejemplo, inserciones, deleciones, sustituciones, translocaciones, reordenamientos y/o mutaciones puntuales. Una secuencia genómica también puede comprender uno de los diversos alelos diferentes.

Generalidades

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

Definiciones

Las expresiones "ácido nucleico" "polinucleótido" y "oligonucleótido" se usan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y tanto en forma monocatenaria como bicatenaria. Para los fines de la presente descripción, estas expresiones no deben

considerarse como limitantes con respecto a la longitud de un polímero. Las expresiones pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en la base, el azúcar y/o fracciones fosfato (por ejemplo, esqueletos de fósforoato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo A se emparejará con la base T.

- 5 Los términos "polipéptido," "péptido" y "proteína" se usan indistintamente para referirse a un polímero de restos de aminoácido. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de aminoácidos correspondientes de origen natural.

10 "Unión" se refiere a una interacción no covalente específica de secuencia entre macromoléculas (por ejemplo, entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión necesitan ser específicos de la secuencia (por ejemplo, contactos con restos de fosfato en un esqueleto de ADN), siempre y cuando la interacción en conjunto sea específica de la secuencia. Dichas interacciones se caracterizan generalmente por una constante de disociación (K_d) de 10^{-6} M⁻¹ o inferior. "Afinidad" se refiere a la fuerza de unión: estando una mayor afinidad de unión correlacionada con una K_d inferior.

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

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

25 Los dominios de unión de dedo de zinc pueden "modificarse genéticamente" para unirse a una secuencia de nucleótidos predeterminada. Los ejemplos no limitantes de métodos para modificar genéticamente proteínas con dedos de zinc son el diseño y la selección. Una proteína con dedo de zinc diseñada es una proteína que no se produce en la naturaleza cuyo diseño/composición es resultado principalmente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y de algoritmos computerizados para procesar información en una bases de datos que almacena información de diseños de ZFP y de los datos de unión existentes Véase, por ejemplo, las patentes de Estados Unidos 6.140.081; 6.453.242 6.534.261 y 6.785.613; véanse, también los documentos WO 98/53058, WO 98/53059, WO 98/53060, WO 02/016536 y WO 03/016496; y las patentes de Estados Unidos 6.746.838; 6.866.997 y 7.030.215.

35 Una proteína con dedo de zinc "seleccionada" es una proteína que no se encuentra en la naturaleza cuya producción es el resultado principalmente de un proceso empírico, tal como la presentación en fagos, la trampa de la interacción o la selección de híbridos. Véanse, por ejemplo, los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; US 6.733.970; US RE39.229; y WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197 y WO 02/099084.

40 Una proteína con dedo de zinc "no canónica" es una proteína que comprende un dedo de zinc no canónico (no C2H2). Por tanto un dedo de zinc no canónico comprende una sustitución, adición y/o deleción de al menos un aminoácido, en comparación con una proteína de dedo de zinc C2H2 de origen natural. Los ejemplos no limitantes de dedos de zinc no canónicos incluyen los que comprenden restos de coordinación de zinc (del extremo amino al carboxilo) de Cys-Cys-His-Cys (por ejemplo, C3H).

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

En la materia se conocen técnicas para determinar la identidad de secuencia de ácidos nucleicos y aminoácidos. Típicamente, dichas técnicas incluyen la determinación de la secuencia de nucleótidos del ARNm para un gen y/o la determinación de la secuencia de aminoácidos codificada por la misma, y comparar estas secuencias con una segunda secuencia de nucleótidos o de aminoácidos. También pueden determinarse secuencias genómicas y compararse de esta manera. En general, identidad se refiere a una correspondencia exacta de nucleótido con nucleótido o de aminoácido con aminoácido de dos secuencias de polinucleótidos o de polipéptidos, respectivamente. Dos o más secuencias (de polinucleótidos o de aminoácidos) pueden compararse determinando su porcentaje de identidad. El porcentaje de identidad de dos secuencias, bien secuencias de ácidos nucleicos o de aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas dividido entre la longitud de la secuencias más cortas y multiplicado por 100. Un alineamiento aproximado para las secuencias de ácido nucleico se proporciona en el algoritmo de homología local de Smith y Waterman, Advances in Applied Mathematics 2:482-489 (1981). Este algoritmo puede aplicarse a secuencias de aminoácidos usando la matriz de 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 normalizada por Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). Una implementación ejemplar de este algoritmo para determinar el porcentaje de identidad de una secuencia se proporciona en el grupo Genetics Computer (Madison, WI) en la aplicación de utilidad "BestFit". Los parámetros por defecto para este método se describen en el Wisconsin Sequence Analysis Package Program Manual, versión 8 (1995) (disponible en el Grupo Genetics Computer, Madison, WI). Un método ejemplar de establecimiento del porcentaje de identidad en el contexto de la presente descripción es el uso del paquete de programas MPSRCH registrado por la universidad de Edimburgo, desarrollado por F. Collins y Shane S. Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). Para este juego de paquetes puede emplearse el algoritmo de Smith-Waterman en donde se usan parámetros por defecto para la tabla de puntuación (por ejemplo, penalización de 12 por apertura de hueco, penalización de uno por extensión de hueco y un hueco de seis). A partir de los datos generados, el valor de "coincidencia" refleja la identidad de secuencia. Generalmente, en la técnica se conocen otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias, por ejemplo, otro programa de alineamiento es BLAST, usado con parámetros por defecto. Por ejemplo, pueden usarse BLASTN y BLASTP con los siguientes parámetros por defecto: código genético = estándar; filtro = ninguno; cadena = ambas; valor de corte = 60; esperado =10; Matriz = BLOSUM62; Descripciones = 50 secuencias; clasificado mediante= PUNTUACIÓN MÁS ALTA; Base de datos = no redundante, GenBank + EMBL + DDBJ + PDB + traducciones CDS de GenBank + Swiss protein + Spupdate + PIR. Los detalles de estos programas pueden encontrarse en internet. Con respecto a las secuencias descritas en la presente memoria, el intervalo de grados de identidad de secuencia deseados es de aproximadamente del 35 % al 100 % y cualquier valor entero entre ellos. Típicamente el porcentaje de identidad entre las secuencias es una identidad de secuencia de al menos 35%- 40 %; 40 % - 45 %; 45 %- 50 %; 50 %- 60 %; 60 %- 70 %; 70 - 75 %, preferentemente 80 - 82 %, más preferentemente 85 - 90 %, incluso más preferentemente 92 %, aún más preferentemente 95 % y lo más preferentemente de 98 %.

Como alternativa, el grado de similitud de secuencia entre polinucleótidos puede determinarse por hibridación de polinucleótidos en condiciones que permitan la formación de dúplex estables entre regiones homólogas, seguido por la digestión con una o más nucleasas específicas de cadena sencilla y la determinación del tamaño de los fragmentos digeridos. Dos ácidos nucleicos, o dos secuencias polipeptídicas son sustancialmente homólogas entre sí cuando las secuencias presentan una identidad de secuencia de al menos aproximadamente 70 % - 75 %, preferentemente 80 % - 82 %, más preferentemente 85 %- 90 %, incluso más preferentemente 92 %, aún más preferentemente 95 % y lo más preferentemente 98 %, sobre una longitud definida de las moléculas, determinado usando los métodos anteriores. Como se emplea en la presente memoria, sustancialmente homólogo también se refiere a secuencias que muestran una identidad completa con una secuencia de ADN o polipeptídica específica. Las secuencias de ADN que son sustancialmente homólogas pueden identificarse en un experimento de hibridación Southern, por ejemplo, en condiciones rigurosas, tal y como se definen para este sistema particular. La definición de las condiciones de hibridación apropiadas está dentro de los conocimientos del experto en técnica. Véase, por ejemplo, Sambrook *et al.*, citado anteriormente; 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 de la siguiente manera. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta a la eficiencia y a la fuerza de acontecimientos de hibridación entre dichas moléculas. Una secuencia de ácido nucleico parcialmente idéntica inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica a una molécula diana. La inhibición de la hibridación de la secuencia completamente idéntica puede evaluarse usando ensayos de hibridación que son muy conocidos en la técnica (por ejemplo, transferencia (de ADN) de Southern, transferencia (de ARN) de Northern, hibridación en solución, o similares véase Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, segunda edición, (1989) Cold Spring Harbor, N.Y.). Dichos ensayos pueden realizarse usando diversos grados de selectividad, por ejemplo, usando condiciones que varían de una baja a una alta rigurosidad. Si se emplean condiciones de baja rigurosidad, la ausencia de unión inespecífica puede evaluarse usando una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda que tenga una identidad de secuencia de menos de aproximadamente el 30 % con la molécula diana), de tal manera que, en ausencia de acontecimientos de unión inespecífica, la sonda secundaria no se hibridará con la diana.

5 Cuando se utiliza un sistema de detección basado en hibridación, se selecciona una sonda de ácido nucleico que es complementaria a una secuencia de ácido nucleico de referencia y después por selección de condiciones apropiadas, la sonda y la secuencia de referencia se hibridan, o se unen, selectivamente entre sí para formar una molécula bicatenaria. Una molécula de ácido nucleico que es capaz de hibridarse selectivamente con una secuencia de referencia en condiciones de hibridación moderadamente rigurosas 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 una identidad de secuencia de al menos aproximadamente 70% con la secuencia de la sonda de ácido nucleico seleccionado. Típicamente, las condiciones de hibridación rigurosas permiten la detección de secuencias de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tiene una identidad de secuencia de más de aproximadamente 90 - 95 % con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de la secuencias de la sonda/referencia, en donde la sonda y la secuencia de referencia tienen un grado específico de identidad de secuencia, pueden determinarse como se sabe en la técnica (véase, por ejemplo, *Nucleic Acid Hybridization: A Practical Approach*, editores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

15 Los expertos en la técnica conocen bien las condiciones de hibridación. La rigurosidad de la hibridación se refiere al grado al cual las condiciones de hibridación no favorecen la formación híbridos que contienen nucleótidos mal emparejados, correlacionándose una rigurosidad elevada con una menor tolerancia para los híbridos con emparejamiento erróneo. Los factores que afectan a la rigurosidad de la hibridación son muy conocidos por los expertos en la técnica e incluyen, pero no se limitan a, temperatura, pH, fuerza iónica y la concentración de disolventes orgánicos tales como, por ejemplo, formamida y dimetilsulfóxido. Como saben los expertos en la técnica, la rigurosidad de la hibridación aumenta a temperaturas más altas, a una fuerza iónica más baja y a concentraciones de disolvente más bajas.

25 Con respecto a las condiciones rigurosas para la hibridación, es bien conocido en la técnica que puedan emplearse numerosas condiciones equivalentes para establecer una rigurosidad particular modificando, por ejemplo, los siguientes factores: la longitud y la naturaleza de las secuencias, la composición de las bases de las diferentes secuencias, la concentración de las sales y de otros componentes de la solución de hibridación, la presencia o ausencia de agentes bloqueantes en las soluciones de hibridación (por ejemplo, sulfato de dextrano y polietilenglicol), la temperatura de la reacción de hibridación y los parámetros del tiempo, así como diferentes condiciones de lavado. La selección de un conjunto particular de condiciones de hibridación se selecciona siguiendo los métodos convencionales en la técnica (véase, por ejemplo, Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, segunda edición, (1989) Cold Spring Harbor, N.Y).

35 "Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los propósitos de la presente descripción "recombinación homóloga (RH)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas bicatenarias en las células. Este proceso requiere homología entre las secuencias de nucleótidos, utiliza una molécula "donante" para reparar el molde de una molécula "diana" (es decir, la que sufrió la rotura bicatenaria), y se conoce con nombres tales como "conversión génicas sin retrocruzamiento" o "conversión génica de un segmento corto", porque conduce a la transferencia de información genética desde el donante a la diana. Sin desear limitarse a ninguna teoría particular, dicha transferencia puede implicar la corrección de los emparejamientos erróneos del ADN heteroduplex que forman entre la diana rota y el donante y/o la "hibridación de cadena dependiente de la síntesis", en donde el donante se usa para volver a sintetizar la información genética que formará parte de la diana, y/o los procesos relacionados. Dicha RH especializada a menudo da lugar a una alteración de la secuencia de la molécula diana de tal manera que parte o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

45 "Escisión" se refiere a la rotura del esqueleto covalente de una molécula de ADN. La escisión puede iniciarse mediante una serie de métodos que incluyen, pero no se limitan a, hidrólisis enzimática o química de un enlace fosfodiéster. Es posible la escisión tanto monocatenaria como bicatenaria, y la escisión bicatenaria puede producirse como resultado de dos acontecimientos de escisión distintos monocatenarios. La escisión del ADN puede dar lugar a la producción de extremos romos o de extremos cohesivos. En determinadas realizaciones, se usan polipéptidos de fusión para la escisión del ADN bicatenario diana.

50 Un "dominio de escisión" comprende una o más secuencias polipeptídicas que poseen actividad catalítica para la escisión del ADN. Un dominio de escisión puede estar incluido en una sola cadena polipeptídica o la actividad de escisión puede resultar de la asociación de dos (o más) polipéptidos.

Un "semidominio de escisión" es una secuencia polipeptídica que, junto con un segundo polipéptido (idéntico o diferente) forma un complejo que tiene actividad de escisión (por ejemplo, actividad de escisión bicatenaria).

55 Las expresiones "dominio de escisión" y "semidominio de escisión" incluyen dominios de tipo silvestre y partes o mutantes de dominios de escisión o semidominios de escisión que conservan la capacidad de multimerizar (por ejemplo, dimerizar) para formar un dominio de escisión funcional.

"Cromatina" es la estructura nucleoproteica que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteínas, que incluyen las proteínas cromosómicas histónicas y no

- 5 histónicas. La mayor parte de la cromatina celular eucariota existe en forma nucleosomas, en donde un núcleo del 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 el ADN enlazador (de longitud variable dependiendo del organismo) se extiende entre los núcleos del nucleosoma. Una molécula de histona H1 se asocia generalmente al ADN enlazador. Para los fines de la presente descripción, el término "cromatina" significa que incluye todos los tipos de nucleoproteínas celulares, tanto procariotas como eucariotas. La cromatina celular incluye cromatina cromosómica y la episómica.
- 10 Un "cromosoma" es un complejo de cromatina que comprende toda o una parte del genoma de una célula. El genoma de una célula se caracteriza a menudo por su cariotipo, que es el conjunto 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.
- 15 Una "región accesible" es un sitio en la cromatina celular al cual puede unirse un sitio diana presente en el ácido nucleico mediante una molécula exógena que reconoce el sitio diana. Sin desear ligarse a ninguna teoría particular, se piensa que una región accesible es una que no está empaquetada en una estructura nucleosómica. La estructura distinta de una región accesible a menudo puede detectarse por su sensibilidad a sondas químicas y enzimáticas, por ejemplo, nucleasas.
- 20 Un "sitio diana" o "secuencia diana" es una secuencia de ácido nucleico que define una parte de un ácido nucleico al cual se unirá una molécula de unión, siempre que existan 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.
- 25 Una molécula "exógena" es una molécula que normalmente no está presente en una célula, pero que puede introducirse en una célula mediante uno o más métodos genéticos, bioquímicos u otros. "La presencia normal en la célula" se determina con respecto a la fase de desarrollo y a las condiciones ambientales particulares de la célula. Por tanto, por ejemplo, una molécula que solo está presente durante el desarrollo embrionario de músculo es una molécula exógena, con respecto a una célula de músculo adulto. De manera similar, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula que no ha experimentado choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión funcional de una molécula endógena no funcional o una versión no funcional de una molécula endógena normalmente funcional.
- 30 Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como generada mediante un proceso químico combinatorio, o una macromolécula tal como una proteína, un ácido nucleico, un hidrato de carbono, un lípido, una glucoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores o cualquier complejo que comprenda una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser mono o bicatenarios; lineales, ramificados o circulares y pueden tener cualquier longitud. Los ácidos nucleicos incluyen los que pueden formar dúplex, así como los ácidos nucleicos que forman triplex. Véanse, por ejemplo, las patentes de Estados Unidos n.º 5.176.996 y 5.422.251. Las proteínas incluyen, pero no se limitan a, proteínas de unión a ADN, factores de transcripción, factores de remodelación de la cromatina, proteínas de unión a ADN metilado, polimerasas, metilasas, desmetilasas, acetilasas, quinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.
- 35 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína exógena o un ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma de infección vírico, una cadena T de *Agrobacterium tumefaciens*, un plásmido o episoma introducido en una célula, o un cromosoma que no está normalmente presente en la célula. Sin embargo, los ácidos nucleicos o polinucleótidos exógenos, pueden contener secuencias que son homólogas o idénticas a secuencias endógenas. Con respecto a una región genómica endógena particular, una "secuencia exógena" se refiere a una secuencia de nucleótidos que no está presente en esa región. Dicha secuencia exógena puede estar presente en otra localización cromosómica endógena o puede no estar presente en todo el genoma. Por tanto, un polinucleótido exógeno puede contener secuencias tanto exógenas como endógenas: por ejemplo, un transgén flanqueado por secuencias homólogas a una región genómica. Dichos ácidos nucleicos exógenos se usan en métodos de integración y recombinación dirigida como se describe más adelante. Los expertos en la técnica conocen métodos de introducción de moléculas exógenas en células e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores de virus.
- 40 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína exógena o un ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma de infección vírico, una cadena T de *Agrobacterium tumefaciens*, un plásmido o episoma introducido en una célula, o un cromosoma que no está normalmente presente en la célula. Sin embargo, los ácidos nucleicos o polinucleótidos exógenos, pueden contener secuencias que son homólogas o idénticas a secuencias endógenas. Con respecto a una región genómica endógena particular, una "secuencia exógena" se refiere a una secuencia de nucleótidos que no está presente en esa región. Dicha secuencia exógena puede estar presente en otra localización cromosómica endógena o puede no estar presente en todo el genoma. Por tanto, un polinucleótido exógeno puede contener secuencias tanto exógenas como endógenas: por ejemplo, un transgén flanqueado por secuencias homólogas a una región genómica. Dichos ácidos nucleicos exógenos se usan en métodos de integración y recombinación dirigida como se describe más adelante. Los expertos en la técnica conocen métodos de introducción de moléculas exógenas en células e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores de virus.
- 45 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína exógena o un ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma de infección vírico, una cadena T de *Agrobacterium tumefaciens*, un plásmido o episoma introducido en una célula, o un cromosoma que no está normalmente presente en la célula. Sin embargo, los ácidos nucleicos o polinucleótidos exógenos, pueden contener secuencias que son homólogas o idénticas a secuencias endógenas. Con respecto a una región genómica endógena particular, una "secuencia exógena" se refiere a una secuencia de nucleótidos que no está presente en esa región. Dicha secuencia exógena puede estar presente en otra localización cromosómica endógena o puede no estar presente en todo el genoma. Por tanto, un polinucleótido exógeno puede contener secuencias tanto exógenas como endógenas: por ejemplo, un transgén flanqueado por secuencias homólogas a una región genómica. Dichos ácidos nucleicos exógenos se usan en métodos de integración y recombinación dirigida como se describe más adelante. Los expertos en la técnica conocen métodos de introducción de moléculas exógenas en células e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores de virus.
- 50 Una molécula exógena puede ser el mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína exógena o un ácido nucleico exógeno. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma de infección vírico, una cadena T de *Agrobacterium tumefaciens*, un plásmido o episoma introducido en una célula, o un cromosoma que no está normalmente presente en la célula. Sin embargo, los ácidos nucleicos o polinucleótidos exógenos, pueden contener secuencias que son homólogas o idénticas a secuencias endógenas. Con respecto a una región genómica endógena particular, una "secuencia exógena" se refiere a una secuencia de nucleótidos que no está presente en esa región. Dicha secuencia exógena puede estar presente en otra localización cromosómica endógena o puede no estar presente en todo el genoma. Por tanto, un polinucleótido exógeno puede contener secuencias tanto exógenas como endógenas: por ejemplo, un transgén flanqueado por secuencias homólogas a una región genómica. Dichos ácidos nucleicos exógenos se usan en métodos de integración y recombinación dirigida como se describe más adelante. Los expertos en la técnica conocen métodos de introducción de moléculas exógenas en células e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, incluyendo lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vectores de virus.
- 55 Por otro lado, una molécula "endógena" es una que está normalmente presente en una célula particular en una fase de desarrollo particular en condiciones ambientales particulares. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, cloroplasto u otro orgánulo, o un ácido nucleico episómico de origen natural. Las moléculas endógenas adicionales pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

- Una molécula de "fusión" es una molécula en donde dos o más moléculas subunitarias están unidas, por ejemplo, de forma covalente. Las moléculas subunitarias pueden ser moléculas del mismo tipo químico o pueden ser moléculas de tipos químicos diferentes. Los ejemplos del primer tipo de molécula de fusión incluyen, pero no se limitan a, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión al ADN de ZFP y un dominio de escisión) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita anteriormente). Como ejemplos del segundo tipo de molécula de fusión se incluyen, pero no se limitan a, una fusión entre un ácido nucleico de forma triplex y un polipéptido, y una fusión entre un ligando del surco menor y un ácido nucleico.
- La expresión de una proteína de fusión en una célula puede producirse por el suministro de la proteína de fusión a la célula o por el suministro de un polinucleótido que codifica la proteína de fusión a una célula, en donde el polinucleótido se transcribe, y el transcrito se traduce, para generar la proteína de fusión. El corte y empalme en *trans*, la escisión de polipéptidos y el ligamiento de polipéptidos también pueden estar implicados en la expresión de una proteína en la célula. En cualquier parte de esta descripción se presentan métodos para el suministro de polinucleótidos y polipéptidos a las células.
- Un "gen", para los fines de la presente descripción, incluye una región de ADN que codifica un producto génico (véase más adelante), así como todas las regiones de ADN que regulan la producción del producto génico, ya sean o no dichas secuencias reguladoras, adyacentes a las secuencias codificantes y/o transcritas. Por consiguiente, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias traduccionales reguladoras, tales como sitios de unión a ribosomas y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aislantes, elementos de contorno, orígenes de replicación, sitios de fijación a la matriz y regiones de control del locus.
- "Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto directo de la transcripción de un gen (por ejemplo, ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por la traducción de un ARNm. Los productos génicos también incluyen los ARN que se modifican, mediante procesos tales como encapuchamiento (*capping*), poliadenilación, metilación y edición y proteínas modificadas, por ejemplo, por metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristilación y glucosilación.
- "Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitación, tanto a la activación como a la represión génica.
- Las células de "plantas" incluyen, pero no se limitan a, células de plantas monocotiledóneas (monocots) o dicotiledóneas (dicots). Como ejemplos no limitantes de monocots se incluyen plantas de cereales tales como maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, cebolla, banana y coco. Los ejemplos no limitantes de dicots incluyen tabaco, tomate, girasol, algodón, remolacha, patata, lechuga, melón, semilla de soja, canola (colza) y alfalfa. Las células de las plantas pueden ser de cualquier parte de la planta y/o de cualquier fase del desarrollo de la planta.
- Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante, dentro o adyacente a un gen, en donde es deseable unir una molécula exógena. La unión puede ser con el fin de efectuar escisión de ADN y/o recombinación dirigidas. Una región de interés puede estar presente en un cromosoma, un episoma, un orgánulo celular (por ejemplo, mitocondria, cloroplasto), o en un genoma de infección vírica, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias líder, secuencias remolque o intrones, o dentro de regiones no transcritas, bien aguas arriba o aguas abajo de la región codificante. Una región de interés puede tener una longitud tan pequeña como de un solo par de nucleótidos o de hasta 25.000 pares de nucleótidos, o cualquier valor entero de pares de nucleótidos.
- Las expresiones "unión operativa" y "unido(a) operativamente", se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en donde los componentes se disponen de tal manera que ambos componentes funcionen normalmente y permitan la posibilidad de que al menos uno de los componentes pueda actuar como mediador en una función que se ejerce después de al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora transcripcional, tal como un promotor, está unido operativamente a una secuencia codificante si la secuencia reguladora transcripcional controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o a la ausencia de uno o más factores reguladores transcripcionales. Una secuencia reguladora transcripcional está, en general, unida operativamente en *cis* con una secuencia codificante, pero no tiene que estar necesariamente adyacente a esta. Por ejemplo, un potenciador es una secuencia reguladora transcripcional que está unida operativamente a una secuencia codificante, incluso aunque no sean contiguas.
- Con respecto a los polipéptidos de fusión, la expresión "unido operativamente" puede referirse al hecho de que cada uno de los componentes realiza la misma función unido al otro componente como si lo realizase si no estuviese unido. Por ejemplo, con respecto a un polipéptido de fusión en donde un dominio de unión al ADN de ZFP está fusionado a un dominio de escisión, el dominio de unión al ADN de ZFP y el dominio de escisión están en unión operativa si, en el polipéptido de fusión, la parte del dominio de unión al ADN de ZFP puede unir su sitio diana y/o su

sitio de unión, mientras que el dominio de escisión puede escindir el ADN cerca 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, sin embargo conserva 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 restos que la molécula nativa correspondiente, y/o puede contener una o más sustituciones de aminoácidos o de nucleótidos. En la técnica se conocen bien métodos para determinar la función de un ácido nucleico (por ejemplo, función de codificación, capacidad de hibridar con otro ácido nucleico). Del mismo modo, se conocen bien métodos para determinar la función de las proteínas. Por ejemplo, la función de unión al ADN de un polipéptido puede determinarse, por ejemplo, mediante unión a filtros, cambios de movilidad electroforética, o ensayos de inmunoprecipitación. La escisión del ADN puede ensayarse mediante electroforesis en gel. Véase Ausubel *et al.*, citado anteriormente. La capacidad de una proteína para interactuar con otra proteína puede determinarse, por ejemplo, por coinmunoprecipitación, ensayos de doble híbrido o complementación, tanto genéticos como bioquímicos. Véase, por ejemplo, Fields *et al.* (1989) Nature 340:245-246; patente de Estados Unidos n.º 5.585.245 y documento PCT WO 98/44350.

15 Dominios de unión de dedos de zinc

En la presente memoria se describen dominios de unión de dedos de zinc no canónicos y polinucleótidos que codifican estos dominios de unión de dedos de zinc. En determinadas realizaciones, los dominios de unión de dedos de zinc no canónicos descritos en la presente memoria son dedos de zinc C₃H, en los que uno de los restos de histidina de coordinación de zinc conservados se transforma en cisteína. En realizaciones adicionales, el resto de histidina más C terminal se transforma en un resto de cisteína, generando una "proteína CCHC".

Un dominio de unión de dedo de zinc puede comprender uno o más dedos de zinc (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de zinc), y puede modificarse por ingeniería genética para unirse a cualquier secuencia diana (por ejemplo, una secuencia genómica). Los dominios de unión de dedos de zinc pueden unirse a ADN, ARN y/o a una proteína. Típicamente, un solo dominio de dedo de zinc tiene una longitud de aproximadamente 30 aminoácidos. Los dedos de zinc incluyen dedos de zinc tanto canónicos C₂H₂ (es decir, aquellos en los que el ion zinc está coordinado por dos restos de cisteína y dos restos de histidina) y dedos de zinc no canónicos incluyendo, por ejemplo, dedos de zinc C₃H (aquellos en los que el ion zinc está coordinado por tres restos de cisteína y un resto de histidina). Véanse también las solicitudes de patente de Estados Unidos n.º 20030108880, 20060246567 y 20060246588.

Estudios estructurales han demostrado que un dominio (motivo) de dedo de zinc canónico contiene dos láminas beta (sujetadas en un giro beta que contiene los dos restos de cisteína invariantes) y una hélice alfa (que contiene los dos restos de cisteína invariantes), que se sujetan en una conformación particular a través de la coordinación de un átomo de zinc mediante las dos cisteínas y las dos histidinas. Los dedos de zinc no canónicos descritos en la presente memoria conservan esta estructura beta-beta-alfa.

Los dedos de zinc no canónicos descritos en la presente memoria pueden ser dominios de unión de dedos de zinc de origen natural. Sin embargo, más típicamente, los dedos de zinc no canónicos, como los descritos en la presente memoria, incluyen uno o más componentes de dedo de zinc donde al menos uno de los restos de cisteína o histidina que coordinan el zinc se ha reemplazado por uno o más aminoácidos. Por ejemplo, en determinadas realizaciones, el resto de His C terminal de una molécula de unión de dedo de zinc canónica se reemplaza por un resto de Cys.

Los dedos de zinc CCHC descritos en la presente memoria también pueden comprender una o más alteraciones (con respecto a la secuencia de un dedo de zinc C₂H₂ de origen natural) en la secuencia de restos de aminoácidos distinta de la de los restos de coordinación del zinc. Dichas alteraciones pueden comprender sustituciones, deleciones y/o inserciones. Los aminoácidos pueden alterarse en cualquier parte en el dedo de zinc. Como ejemplos no limitantes de alteraciones se incluyen: (1) sustituciones de restos sencillos que rodean el resto de coordinación de zinc alterado; (2) adición de restos extra antes o después del resto de coordinación de zinc alterado, (por ejemplo, en casos en los que el resto de His más C terminal se transforma en Cys, la adición de restos de aminoácidos extra puede facilitar la coordinación de zinc compensando la cadena lateral de cisteína más corta); y/o (3) sustitución de restos localizados entre los restos de His y Cys de un dedo de zinc CCHC de origen natural en la región correspondiente de un dedo de zinc CCHC no canónico.

Las proteínas de dedo de zinc descritas en la presente memoria incluyen al menos un dedo de zinc que comprende un dedo de zinc no canónico (no C₂H₂), en donde el dedo de zinc no canónico tiene una parte helicoidal implicada en la unión al ADN y en donde al menos un dedo de zinc comprende la secuencia: Cys-(X^A)₂₋₄-Cys-(X^B)₁₂-His-(X^C)₃₋₅-Cys-(X^D)₁₋₁₀ (SEQ ID NO:3), en donde X^A, X^B, X^C y X^D representan cualquier aminoácido e His-(X^C)₃₋₅-Cys(X^D)₁₋₁₀ comprende una de las secuencias como se expone en cualquiera de las SEQ ID NO: 14-80, 83-87 y 89-92 y adicionalmente en donde la proteína de dedo de zinc está modificada por ingeniería genética para unirse a una secuencia diana. En realizaciones en las que X^C comprende 3 restos (i) al menos uno de estos restos está alterado en comparación con un dedo de zinc CCHC canónico; y/o (ii) X^D comprende al menos una deleción, sustitución o inserción en comparación con un dedo de zinc CCHH canónico. En determinadas realizaciones, X^D comprende la secuencia QLV o QKP. En otras realizaciones X^D comprende uno o más (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10) restos de Gly (G).

Las secuencias parciales de aminoácidos (incluyendo y C terminal al 3^{er} resto de coordinación de zinc) de dedos de zinc no canónicos ejemplares se muestran en las tablas 1, 2, 3 y 4. La materia objeto de la invención se refiere solo a los dedos de zinc definidos en las reivindicaciones. En todas las tablas, los dos restos de coordinación de zinc (H y C) más C- terminales (es decir, el tercero y el cuarto) están subrayados. Las alteraciones (por ejemplo, sustituciones, inserciones, deleciones) en comparación con la secuencia del dedo no canónico de "tipo silvestre" (fila 2 de las tablas 1 y 3) se muestran con doble subrayado.

5

Tabla 1

| | |
|-------|--|
| C2H2 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>H</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (canónico tipo silvestre) (SEQ ID NO:12) |
| 1=C3H | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (tipo silvestre alterado a no canónico) (SEQ ID NO:13) |
| 2 | <u>H</u> <u>T</u> <u>K</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:14) |
| 3 | <u>H</u> <u>T</u> <u>K</u> <u>A</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:15) |
| 4 | <u>H</u> <u>T</u> <u>K</u> <u>V</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:16) |
| 5 | <u>H</u> <u>T</u> <u>K</u> <u>L</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:17) |
| 6 | <u>H</u> <u>T</u> <u>K</u> <u>S</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:18) |
| 7 | <u>H</u> <u>T</u> <u>K</u> <u>N</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:19) |
| 8 | <u>H</u> <u>T</u> <u>K</u> <u>K</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:20) |
| 9 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:21) |
| 10 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:22) |
| 11 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:23) |
| 12 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:24) |
| 13 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:25) |
| 14 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:26) |
| 15 | <u>H</u> <u>L</u> <u>K</u> <u>G</u> <u>N</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:27) |
| 16 | <u>H</u> <u>L</u> <u>K</u> <u>G</u> <u>N</u> <u>C</u> <u>P</u> <u>A</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:28) |
| 17 | <u>H</u> <u>S</u> <u>E</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:29) |
| 18 | <u>H</u> <u>S</u> <u>E</u> <u>G</u> <u>G</u> <u>C</u> <u>P</u> <u>G</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:30) |
| 19 | <u>H</u> <u>S</u> <u>S</u> <u>S</u> <u>N</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:31) |
| 20 | <u>H</u> <u>S</u> <u>S</u> <u>S</u> <u>N</u> <u>C</u> <u>T</u> <u>I</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:32) |

10

Tabla 2

| | |
|----|--|
| 1 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:33) |
| 2 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:34) |
| 3 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:35) |
| 4 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:36) |
| 5 | <u>H</u> <u>T</u> <u>K</u> <u>I</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:37) |
| 6 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:38) |
| 7 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:39) |
| 8 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:40) |
| 9 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:41) |
| 10 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:42) |
| 11 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:43) |
| 12 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:44) |
| 13 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:45) |
| 14 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>G</u> <u>C</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:46) |
| 15 | <u>H</u> <u>T</u> <u>K</u> <u>R</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:47) |
| 16 | <u>H</u> <u>L</u> <u>K</u> <u>G</u> <u>N</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:48) |
| 17 | <u>H</u> <u>L</u> <u>K</u> <u>G</u> <u>N</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:49) |
| 18 | <u>H</u> <u>L</u> <u>K</u> <u>G</u> <u>N</u> <u>C</u> <u>G</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:50) |
| 19 | <u>H</u> <u>K</u> <u>E</u> <u>R</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:51) |
| 20 | <u>H</u> <u>T</u> <u>R</u> <u>R</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:52) |
| 21 | <u>H</u> <u>A</u> <u>O</u> <u>R</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:53) |
| 22 | <u>H</u> <u>K</u> <u>K</u> <u>F</u> <u>Y</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:54) |
| 23 | <u>H</u> <u>K</u> <u>K</u> <u>H</u> <u>Y</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:55) |
| 24 | <u>H</u> <u>K</u> <u>K</u> <u>Y</u> <u>T</u> <u>C</u> <u>G</u> <u>L</u> <u>R</u> <u>G</u> <u>S</u> <u>Q</u> <u>L</u> <u>V</u> (SEQ ID NO:56) |

| | |
|----|--|
| 25 | <u>HKYYCGLRGSQ</u> LV (SEQ ID NO:57) |
| 26 | <u>HKOYYCGLRGSQ</u> LV (SEQ ID NO:58) |
| 27 | <u>HLLKKCGLRGSQ</u> LV (SEQ ID NO:59) |
| 28 | <u>HQKPCGLRGSQ</u> LV (SEQ ID NO:60) |
| 29 | <u>HQKCLCGLRGSQ</u> LV (SEQ ID NO:61) |
| 30 | <u>HOIRGCGLRGSQ</u> LV (SEQ ID NO:62) |
| 31 | <u>HIKROSCGLRGSQ</u> LV (SEQ ID NO:63) |
| 32 | <u>HIRRYTCGLRGSQ</u> LV (SEQ ID NO:64) |
| 33 | <u>HISSKKCGLRGSQ</u> LV (SEQ ID NO:65) |
| 34 | <u>HKIOKACGLRGSQ</u> LV (SEQ ID NO:66) |
| 35 | <u>HKRIYTCGLRGSQ</u> LV (SEQ ID NO:67) |
| 36 | <u>HLKGONCGLRGSQ</u> LV (SEQ ID NO:68) |
| 37 | <u>HLKKDGCGLRGSQ</u> LV (SEQ ID NO:69) |
| 38 | <u>HLKYTPCGLRGSQ</u> LV (SEQ ID NO:70) |
| 39 | <u>HTKRCGRGSQ</u> LV (SEQ ID NO:71) |
| 40 | <u>HTKIGCGGRGSQ</u> LV (SEQ ID NO:72) |
| 41 | <u>HLKGNCGRGSQ</u> LV (SEQ ID NO:73) |
| 42 | <u>HLKGNCGGGSQ</u> LV (SEQ ID NO:74) |
| 43 | <u>HIRTCTGSQKP</u> (SEQ ID NO:75) |
| 44 | <u>HIRTCTGSQKP</u> (SEQ ID NO:76) |
| 45 | <u>HIRTGCTGSQKP</u> (SEQ ID NO:77) |
| 46 | <u>HIRTGCGTGSQKP</u> (SEQ ID NO:78) |
| 47 | <u>HIRRCTGSQKP</u> (SEQ ID NO:79) |
| 48 | <u>HIRRGCTGSQKP</u> (SEQ ID NO:80) |

Tabla 3

| | |
|----|------------------------------------|
| ts | <u>HTKIHTGSQKP</u> (SEQ ID NO:81) |
| 1a | <u>HTKICTGSQKP</u> (SEQ ID NO:82) |
| 2a | <u>HTKRCTGSQKP</u> (SEQ ID NO:83) |
| 3a | <u>HAORCTGSQKP</u> (SEQ ID NO:84) |
| 4a | <u>HTKICGTGSQKP</u> (SEQ ID NO:85) |
| 5a | <u>HTKRCGTGSQKP</u> (SEQ ID NO:86) |
| 6a | <u>HAORCGTGSQKP</u> (SEQ ID NO:87) |

Tabla 4

| | |
|-----|-------------------------------------|
| ts | <u>HTKIHLRGSQ</u> LV (SEQ ID NO:88) |
| 7a | <u>HAORCGG</u> (SEQ ID NO:89) |
| 8a | <u>HAORCGGG</u> (SEQ ID NO:90) |
| 9a | <u>HTKICGGG</u> (SEQ ID NO:91) |
| 10a | <u>HTKRCGGG</u> (SEQ ID NO:92) |
| 11a | <u>HAORCG</u> (SEQ ID NO:93) |

Como se ha indicado anteriormente, una ZFP puede incluir cualquier número de dominios de unión de dedos de zinc, por ejemplo, al menos 3 dedos de zinc. Adicionalmente, uno, más de uno, o todos los dedos de zinc pueden ser de dedos de zinc no canónicos como se describe en la presente memoria.

5 En determinadas realizaciones, el dedo más C terminal de una proteína de dedo de zinc de múltiples dedos comprende un dedo de zinc canónico. En otras realizaciones, el dedo más C terminal de una proteína de dedo de zinc de múltiples dedos comprende un dedo CCHC como se describe en la presente memoria, por ejemplo, un dedo CCHC comprende una o más inserciones de aminoácidos C terminal en el resto de Cys de coordinación zinc más C terminal. Véanse los ejemplos 1-5 que describen proteínas de dedos de zinc con 4 dedos en las que el dedo 2 (D2) y/o el dedo 4 (D4) son dedos de zinc no canónicos como se describe en la presente memoria.

10 Los dominios de unión de dedos de zinc pueden modificarse por ingeniería genética para unirse a una secuencia de elección. Véase, por ejemplo, Beerli *et al.* (2002) *Nature Biotechnol.* 20:135-141; Pabo *et al.* (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan *et al.* (2001) *Nature Biotechnol.* 19:656-660; Segal *et al.* (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo *et al.* (2000) *Curr. Opin. Struct. Biol.* 10:411-416. Un dominio de unión de dedo de zinc modificado por ingeniería genética puede tener una nueva especificidad de unión, en comparación con una proteína de dedo de zinc de origen natural. Los métodos de modificación por ingeniería genética incluyen, pero no se limitan a, diseño racional y diversos tipos de selección (por ejemplo, métodos en los que explora una pluralidad de secuencias de dedos de zinc diferentes frente a una sola secuencia de nucleótidos diana). El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos triples (o cuádruples) y secuencias de aminoácidos de dedo de zinc individuales, en las que cada secuencia de nucleótidos triple o cuádruple está asociada con una o más secuencias de aminoácidos de dedos de zinc que se unen a la secuencia triple o cuádruple particular. Véanse, por ejemplo, las patentes de Estados Unidos 6.453.242 y 6.534.261 de propiedad conjunta. Se desvelan métodos de diseño adicionales, por ejemplo, en las patentes de Estados Unidos n.º 6.746.838; 6.785.613, 6.866.997 y 7.030.215. La potenciación de la especificidad de unión para los dominios de unión des dedo de zinc se ha descrito, por ejemplo, en la patente de Estados Unidos n.º 6.794.136 de propiedad conjunta.

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

30 Dado 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 en donde un dominio de dedo de zinc se modifica por ingeniería genética para unirse (por ejemplo, una secuencia diana) determinará el número de dedos de zinc en un dominio de unión de dedo de zinc modificado por ingeniería genética. Por ejemplo, para las ZFP en las que los motivos de dedos de zinc no se unen a subsitios solapantes, una secuencia diana de seis nucleótidos está unida por un dominio de unión de dos dedos de zinc; una secuencia diana de nueve nucleótidos está unida por un dominio de unión de tres dedos, *etc.* No es necesario que en un sitio diana, los sitios de unión para dedos de zinc individuales (es decir, subsitios) sean contiguos, sino que pueden estar separados por uno o varios nucleótidos, dependiendo de la longitud y de la naturaleza de las secuencias de aminoácidos entre los dedos de zinc (es decir, los enlazadores inter-dedo) en un dominio de unión de múltiples dedos. Véanse, por ejemplo, las patentes de Estados Unidos 6.479.626; 6.903.185 y 7.153.949 y solicitud de patente de Estados Unidos n.º 2003/0119023.

40 En un dominio de unión de dedo de zinc de múltiples dedos, los dedos de zinc adyacentes pueden estar separados por secuencias de aminoácidos enlazadoras de aproximadamente 5 aminoácidos (denominadas enlazadores inter-dedo "canónicos") o, como alternativa, por uno o más enlazadores no canónicos. Véanse, por ejemplo, las patentes de Estados Unidos 6.453.242 y 6.534.261, de propiedad conjunta. Para la modificación por ingeniería genética de los dominios de unión de dedos de zinc que comprenden más de tres dedos, la inserción de secuencias enlazadoras inter-dedo ("no canónicas") más largas entre determinados dedos de zinc puede aumentar la afinidad y/o especificidad de unión por el dominio de unión. Véase, por ejemplo, la patente de Estados Unidos n.º 6.479.626 y la publicación de solicitud de patente de Estados Unidos n.º 2003/0119023.

50 Por consiguiente, los dominios de unión de dedo de zinc de múltiples dedos también pueden caracterizarse con respecto a la presencia y a la localización de enlazadores inter-dedo no canónicos. El uso de enlazadores inter-dedo más largos también puede facilitar la unión de una proteína de dedo de zinc con sitios diana que comprenden nucleótidos no contiguos. Como resultado, uno o más subsitios, en un sitio diana para un dominio de unión de dedo de zinc, pueden separarse entre sí por 1, 2, 3, 4, 5 o más nucleótidos. Como ejemplo, un dominio de unión de cuatro dedos puede unirse a un sitio diana de 13 nucleótidos que comprende, en la secuencia, dos subsitios contiguos de 3 nucleótidos, un nucleótido interviniente y dos subsitios triplete contiguos.

55 Un subsitio diana es una secuencia de nucleótidos (generalmente de 3 o 4 nucleótidos) que está unida por un solo dedo de zinc. Sin embargo, no es necesario que un sitio diana sea un múltiplo de tres nucleótidos. Por ejemplo, en casos en los que se producen interacciones de cadena cruzada (véanse, por ejemplo, las patentes de Estados Unidos 6.453.242 y 6.794.136), uno o más de los dedos de zinc individuales de un dominio de unión de múltiples dedos pueden unirse a subsitios cuádruples solapantes. Véanse también las patentes de Estados Unidos 6.746.838 y 6.866.997. Como ejemplo, un dominio de unión de tres de dedos puede unirse a un sitio diana de 10 nucleótidos

que comprende tres subsitios de 4 nucleótidos solapantes.

5 La selección de una secuencia en la cromatina celular para la unión mediante un dominio de dedo de zinc (por ejemplo, un sitio diana) puede realizarse, por ejemplo, de acuerdo con los métodos desvelados en la patente de Estados Unidos n.º 6.453.242 (17 de septiembre, 2002) de mismo propietario que la presente, que también desvela métodos para el diseño de ZFP que se unen a una secuencia seleccionada. Será obvio para los expertos en la técnica, que para la selección de un sitio diana también pueda usarse inspección visual simple de una secuencia de nucleótidos. Por consiguiente, en los métodos descritos en la presente memoria puede usarse cualquier medio para la selección del sitio diana.

10 Las proteínas de dedo de zinc de múltiples dedos pueden construirse uniendo dedos de zinc individuales obtenidos, por ejemplo, por diseño o selección. Como alternativa, los módulos de unión que constan de dos dedos de zinc pueden unirse entre sí, usando o enlazadores inter-dedo canónicos o canónicos, más largos (véase anteriormente) para generar proteínas con cuatro y seis dedos de zinc. Dichos dos módulos de dos dedos pueden obtenerse, por ejemplo, seleccionando dos dedos adyacentes, que se unen a una secuencia diana particular de seis nucleótidos, en el contexto de una proteína de múltiples dedos (generalmente de tres dedos). Véase, por ejemplo, el documento WO 15 98/53057 y la publicación de solicitud de patente de Estados Unidos n.º 2003/0119023.

Como alternativa, pueden construirse módulos de dos dedos por ensamblaje de dedos de zinc individuales.

Por tanto, los dominios de dedos de zinc descritos en la presente memoria pueden usarse individualmente o en diversas combinaciones para construir proteínas de dedo de zinc de múltiples dedos que se unan a cualquier sitio diana.

20 La distancia entre secuencias (por ejemplo, sitios de diana) se refiere al número de nucleótidos o pares de nucleótidos que intervienen entre dos secuencias, medido a partir de los bordes de las secuencias más próximas entre sí.

25 En realizaciones que utilizan ZFN, por ejemplo, en las que la escisión depende de la unión de dos moléculas de fusión de dominio de dos dedos de zinc/semidominio de escisión para separar sitios diana, los dos sitios diana pueden estar en cadenas de ADN opuestas. En otras realizaciones, ambos sitios diana están en la misma cadena de ADN. Véase, por ejemplo, el documento WO 2005/084190.

30 Los polinucleótidos que codifican dedos de zinc o proteínas de dedos de zinc también están dentro del alcance de la presente descripción. Estos polinucleótidos pueden construirse usando técnicas convencionales e insertarse en un vector, y el vector puede introducirse en una célula (véase más adelante para una descripción adicional con respecto a vectores y a métodos para la introducción de polinucleótidos en células) de tal manera que la proteína codificada se expresa en la célula.

Proteínas de fusión

También se proporcionan proteínas de fusión que comprenden uno o más componentes de dedos de zinc no canónicos.

35 Las moléculas de fusión se construyen mediante métodos bioquímicos de clonación y conjugación, que son muy conocidos por los expertos en la técnica. Las moléculas de fusión comprenden una ZFP que contiene CCHC y, por ejemplo, un dominio de escisión, un semidominio de escisión, un dominio de activación transcripcional, un dominio de represión transcripcional, un componente de un complejo de remodelación de la cromatina, un dominio aislante, un fragmento funcional de cualquiera de estos dominios; y/o cualquiera de las combinaciones de dos o más 40 dominios funcionales o sus fragmentos.

En determinadas realizaciones, las moléculas de fusión comprenden una proteína de dedo de zinc de plantas modificada y al menos dos dominios funcionales (por ejemplo, un dominio aislante o un dominio de proteína de unión a metilo y, adicionalmente, un dominio de activación o de represión transcripcional).

45 Las moléculas de fusión también comprenden opcionalmente una señal de localización nuclear (tal como, por ejemplo, la del antígeno T del SV40 o la SLN (señal de localización nuclear) de maíz Opaco-2) y una etiqueta epitópica (tal como, por ejemplo, FLAG o hemaglutinina). Las proteínas de fusión (y los ácidos nucleicos que las codifican) se diseñan de tal manera que la fase de lectura traduccional está conservada entre los componentes de la fusión.

50 Los expertos en la técnica conocen métodos para diseñar y construir proteínas de fusión (y polinucleótidos que las codifican). Por ejemplo, en las patentes de Estados Unidos 6.453.242 y 6.534.261 de propiedad conjunta, se describen métodos para el diseño y la construcción de proteínas de fusión que comprenden proteínas de dedos de zinc (y los polinucleótidos que las codifican).

Los polipéptidos que codifican dichas proteínas de fusión también están dentro del alcance de la presente descripción. Estos polinucleótidos pueden construirse usando técnicas convencionales e insertarse en un vector y el

vector puede introducirse en una célula (véase más adelante para una descripción adicional con respecto a vectores y a métodos para introducir polinucleótidos en células).

Un ejemplo de dominio funcional para la fusión con un dominio de unión a ADN de una ZFP, que se utiliza para reprimir la expresión génica, es un dominio de represión KRAB de la proteína KOX-1 humana (véase, por ejemplo, Thiesen *et al.*, *New Biologist* 2, 363-374 (1990); Margolin *et al.*, *Proc. Natl. Acad. Sci. USA* 91, 4509-4513 (1994); Pengue *et al.*, *Nucl. Acids Res.* 22:2908-2914 (1994); Witzgall *et al.*, *Proc. Natl. Acad. Sci. USA* 91, 4514-4518 (1994). El dominio KOX es también adecuado para su uso como un dominio de represión. Otro dominio de represión adecuado es el dominio de unión a metilo de la proteína 2B (MBD-2B, acrónimo inglés de *methyl binding domain protein 2*) (véase, también Hendrich *et al.* (1999) *Mamm Genome* 10:906-912 para la descripción de proteínas MBD). Otro dominio de represión útil es el que está asociado con la proteína v-ErbA. Véase, por ejemplo, Damm, *et al.* (1989) *Nature* 339:593-597; Evans (1989) *Int. J. Cancer Suppl.* 4:26-28; Pain *et al.* (1990) *New Biol.* 2:284-294; Sap *et al.* (1989) *Nature* 340:242-244; Zenke *et al.* (1988) *Cell* 52:107-119; y Zenke *et al.* (1990) *Cell* 61:1035-1049. Otros ejemplos de dominios de represión incluyen, pero no se limitan a, el receptor de la hormona tiroidea (TR), SID, MBD1, MBD2, MBD3, MBD4, proteínas similares a MBD, miembros de la familia DNMT (por ejemplo, DNMT1, DNMT3A, DNMT3B), Rb, MeCP1 y MeCP2. Véase, por ejemplo, Zhang *et al.* (2000) *Ann Rev Physiol* 62:439-466; Bird *et al.* (1999) *Cell* 99:451-454; Tyler *et al.* (1999) *Cell* 99:443-446; Knoepfler *et al.* (1999) *Cell* 99:447-450; y Robertson *et al.* (2000) *Nature Genet.* 25:338-342. Otros ejemplos de dominios de represión incluyen, pero no se limitan a, ROM2 y AtHD2A. Véase, por ejemplo, Chern *et al.* (1996) *Plant Cell* 8:305-321; y Wu *et al.* (2000) *Plant J.* 22:19-27.

Los dominios adecuados para conseguir la activación incluyen el dominio de activación de VP16 de HSV (véase, por ejemplo, Hagmann *et al.*, *J. Virol.* 71, 5952-5962 (1997)) receptores nucleares de hormonas (véase, por ejemplo, Torchia *et al.*, *Curr. Opin. Cell. Biol.* 10:373-383 (1998)); la subunidad p65 del factor nuclear kappa B (Bitko & Barik, *J. Virol.* 72:5610-5618 (1998) y Doyle & Hunt, *Neuroreport* 8:2937-2942 (1997)); Liu *et al.*, *Cancer Gene Ther.* 5:3-28 (1998)), o dominios funcionales quiméricos artificiales tales como VP64 (Seifpal *et al.*, *EMBO J.* 11, 4961-4968 (1992)).

Otros ejemplos de dominios de activación incluyen, pero no se limitan a, p300, CBP, PCAF, SRC1 PvALF y ERF-2. Véase, por ejemplo, Robyr *et al.* (2000) *Mol. Endocrinol.* 14:329-347; Collingwood *et al.* (1999) *J. Mol. Endocrinol.* 23:255-275; Leo *et al.* (2000) *Gene* 245:1-11; Manteuffel-Cymborowska (1999) *Acta Biochim. Pol.* 46:77-89; McKenna *et al.* (1999) *J. Steroid Biochem. Mol. Biol.* 69:3-12; Malik *et al.* (2000) *Trends Biochem. Sci.* 25:277-283; y Lemon *et al.* (1999) *Curr. Opin. Genet. Dev.* 9:499-504. Otros ejemplos de dominios de activación incluyen, pero no se limitan a, OsGAI, HALF-1, C1, AP1, ARF-5, -6, -7 y -8, CPRF1, CPRF4, MYC-RP/GP y TRAB1. Véase, por ejemplo, Ogawa *et al.* (2000) *Gene* 245:21-29; Okanami *et al.* (1996) *Genes Cells* 1:87-99; Goff *et al.* (1991) *Genes Dev.* 5:298-309; Cho *et al.* (1999) *Plant Mol. Biol.* 40:419-429; Ulmason *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:5844-5849; Sprenger-Haussels *et al.* (2000) *Plant J.* 22:1-8; Gong *et al.* (1999) *Plant Mol. Biol.* 41:33-44; y Hobo *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:15,348-15,353.

Se desvelan dominios funcionales adicionales, por ejemplo, en la patente de Estados Unidos n.º 6.933.113 de propiedad conjunta. Adicionalmente, se describen dominios aislante, proteínas de remodelación de la cromatina tales como las que contienen dominios ISWI y proteínas de dominio de unión a metilo adecuadas para su uso en moléculas de fusión, por ejemplo, en las publicaciones internacionales WO 01/83793 y WO 02/26960 de propiedad conjunta.

En otras realizaciones, la proteína de fusión es una nucleasa con dedos de zinc (ZFN) que comprende uno o más dedos de zinc CCHC como se describe en la presente memoria y un dominio de escisión (o semidominio de escisión). Los dedos de zinc pueden modificarse por ingeniería genética para reconocer una secuencia diana en cualquier región genómica de elección y, cuando se introducen en una célula, se producirá la unión de la proteína (o proteínas) de fusión con su sitio (o sitios) de unión y la escisión dentro o cerca de dicha región genómica. Dicha escisión puede dar como resultado la alteración de la secuencia de nucleótidos de la región genómica (por ejemplo, mutación) seguido de unión de extremos no homólogos. Como alternativa, si un polinucleótido exógeno que contiene secuencias homólogas a la región genómica también está presente en dicha célula, se produce la recombinación homóloga a una alta tasa entre la región genómica y el polinucleótido exógeno, seguido de la escisión dirigida por las ZFN. La recombinación homóloga puede dar como resultado el reemplazo de la secuencia diana o la integración dirigida de secuencias exógenas, dependiendo de la secuencia de nucleótidos del polinucleótido exógeno.

Los dedos de zinc no canónicos descritos en la presente memoria proporcionan función de escisión mejorada cuando se incorporan en las ZFN. Como se describe en los ejemplos, las ZFN con 4 dedos que contienen al menos un dedo CCHC como se describe en la presente memoria, escinden al menos tan bien como las nucleasas que contienen exclusivamente dedos CCHH. En determinadas realizaciones, cuando el dedo C terminal comprende un dedo zinc CCHC no canónico, los restos entre el tercer y cuarto resto de coordinación de zinc (es decir, entre los restos de His y Cys C terminal) son diferentes de los que están presentes en un dedo de zinc CCHH canónico, y después del resto Cys C terminal, se inserta uno o más restos de glicina (por ejemplo, 1, 2, 3, 4, 5 o más).

La parte de dominio de escisión de las ZFN desveladas en la presente memoria puede obtenerse de cualquier endonucleasa o exonucleasa. Como ejemplos de endonucleasas a partir de las cuales puede obtenerse un dominio

de escisión se incluyen, pero no se limitan a, endonucleasas de restricción y endonucleasas de asentamiento (*homing*). Véase, 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 que escinden ADN (por ejemplo, la nucleasa S1; la nucleasa de judía mungo; la DNasa pancreática I; la nucleasa microcócica; la endonucleasa HO de levadura; véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o sus fragmentos funcionales) pueden usarse como una fuente de dominios y semidominios de escisión.

De manera similar, un semidominio de escisión puede proceder de cualquier nucleasa, o de parte de la misma, como se expone anteriormente, siempre que el semidominio de escisión requiera dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión dirigida de ADN genómico si las proteínas de fusión comprenden semidominios de escisión. Como alternativa, puede usarse una sola proteína que comprenda dos semidominios de escisión. Los dos semidominios de escisión pueden proceder de la misma endonucleasa, o cada semidominio de escisión puede proceder de una endonucleasa diferente. Además, los sitios diana para las dos proteínas de fusión se disponen, uno con respecto al otro, de tal manera que la unión de las dos proteínas de fusión con sus sitios diana respectivos coloca el semidominio de escisión en una orientación espacial entre sí, lo que permite que los semidominios de escisión formen un dominio de escisión funcional, por ejemplo, por dimerización. Por tanto, en determinadas realizaciones, los extremos cercanos de los sitios diana se separan por 5-8 pares de nucleótidos o por 15-18 pares de nucleótidos. En realizaciones adicionales, los sitios diana están a una distancia de diez pares de nucleótidos entre sí. Sin embargo, puede intervenir cualquier número entero de nucleótidos o pares de nucleótidos entre dos sitios diana (por ejemplo, de 2 a 50 nucleótidos o más). En general, el punto de escisión se encuentra entre los sitios diana.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y tienen la capacidad de unirse de manera específica de secuencia al ADN (en un sitio de reconocimiento), y escindir el ADN en el sitio de unión o cerca del mismo. Determinadas enzimas de restricción (por ejemplo, de tipo IIS) escinden el ADN en sitios retirados del sitio de reconocimiento y tienen dominios de escisión y unión separables. Por ejemplo, la enzima *Fok I* de tipo IIS cataliza la escisión de ADN bicatenario, a 9 nucleótidos desde su sitio de reconocimiento en una cadena y a 13 nucleótidos desde su sitio de reconocimiento en la otra. Véanse, por ejemplo, las patentes de Estados Unidos 5.356.802; 5.436.150 y 5.487.994; así como Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* 91:883-887; Kim *et al.* (1994b) *J. Biol. Chem.* 269:31,978-31,982. Por tanto, en una realización, las proteínas de fusión comprenden el dominio de escisión (o semidominio de escisión) de al menos una enzima de restricción de tipo IIS y uno más dominios de unión de dedo de zinc, que pueden o no modificarse por ingeniería genética.

Fok I es un ejemplo de una enzima de restricción de tipo IIS, cuyo dominio de escisión es separable del dominio de unión. Esta enzima particular es activa como un dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95: 10,570-10,575. Por consiguiente, para los propósitos de la presente descripción, la parte de la enzima *Fok I* usada en las proteínas de fusión desveladas se considera un semidominio de escisión. Por tanto, para la escisión bicatenaria dirigida y/o el reemplazo dirigido de secuencias celulares usando ZFN que comprenden fusiones de dedo de zinc-*Fok I*, pueden usarse dos proteínas de fusión, comprendiendo cada una de ellas un semidominio de escisión *Fok I*, para reconstituir un dominio de escisión catalíticamente activo. Como alternativa, también puede usarse una sola molécula polipeptídica que contenga un dominio de unión de dedo de zinc y dos semidominios de escisión *Fok I*. Los parámetros para la escisión dirigida y la alteración de secuencias dirigida usando fusiones de dedo de zinc-*Fok I* se proporcionan en cualquier parte en esta descripción y, por ejemplo, en la publicación de solicitud de patente de Estados Unidos n.º 2005/0064474.

En realizaciones adicionales, un semidominio de escisión de *Fok I* puede incluir una o más mutaciones en cualquier resto de aminoácido que afecte a la dimerización. Dichas mutaciones pueden ser útiles para impedir una de un par de fusiones ZFP/*Fok I* a partir de sometimiento a homodimerización que puede conducir a la escisión en secuencias no deseadas. Por ejemplo, todos los restos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de *Fok I* están suficientemente cerca de la interfaz de dimerización para influir en la dimerización. Por consiguiente, las alteraciones en las secuencias de aminoácidos en una o más de las posiciones anteriormente mencionadas pueden usarse para alterar las propiedades de dimerización del semidominio de escisión. Dichos cambios pueden introducirse, por ejemplo, construyendo una biblioteca que contenga (o que codifique) diferentes restos de aminoácidos en estas posiciones y seleccionando variantes con las propiedades deseadas, o diseñando racionalmente mutantes individuales. Además de impedir la homodimerización, también es posible que algunas de estas mutaciones puedan aumentar la eficiencia de la escisión, en comparación con la obtenida con dos semidominios de escisión de tipo silvestre.

Por lo tanto, para la escisión dirigida usando un par de fusiones ZFP/*Fok I*, una o las dos proteínas de fusión pueden comprender una o más alteraciones de aminoácidos que inhiben la autodimerización, pero permiten que se produzca la heterodimerización de las dos proteínas de fusión de tal manera que ocurra la escisión en el sitio diana deseado. En determinadas realizaciones, las alteraciones están presentes en ambas proteínas de fusión y las alteraciones tienen efectos aditivos; es decir, homodimerización de cualquier fusión, que conduce a una escisión aberrante, se minimiza o anula, mientras que la heterodimerización de las dos proteínas de fusión se facilita en comparación con la obtenida con semidominios de escisión de tipo silvestre.

En determinadas realizaciones, el dominio de escisión comprende dos semidominios de escisión, ambos de los cuales son parte de un solo polipéptido que comprende un dominio de unión, un primer semidominio de escisión y un segundo semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o secuencias de aminoácidos diferentes, siempre que actúen escindiendo el ADN.

5 Los semidominios de escisión también pueden proporcionarse en moléculas distintas. Por ejemplo, dos polipéptidos de fusión pueden expresarse en una célula, en donde cada polipéptido comprende un dominio de unión y un semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o secuencias de aminoácidos diferentes, siempre que actúen escindiendo el ADN. Adicionalmente, los dominios de unión se unen a secuencias diana que están típicamente dispuestas de tal manera que, después de la unión de los polipéptidos de fusión, los dos semidominios de escisión se presenten en una orientación espacial entre sí que permita la reconstitución de un dominio de escisión (por ejemplo, por dimerización de los semidominios), colocando de este modo los semidominios uno con respecto al otro para formar un dominio de escisión funcional, dando como resultado la escisión de la cromatina celular en una región de interés. Generalmente, la escisión mediante el dominio de escisión reconstituido se produce en un sitio localizado entre las dos secuencias diana. Una o las dos proteínas pueden modificarse por ingeniería genética para unirse a su sitio diana.

La expresión de dos proteínas de fusión en una célula puede producirse suministrando a la célula las dos proteínas; suministrando a la célula una proteína y un ácido nucleico que codifique una de las proteínas; suministrando a la célula dos ácidos nucleicos, codificando cada uno de ellos una de las proteínas; suministrando a la célula un solo ácido nucleico, que codifique ambas proteínas. En realizaciones adicionales, una proteína de fusión comprende una sola cadena polipeptídica que comprende dos semidominios de escisión y un dominio de unión de dedo de zinc. En este caso, una sola proteína de fusión se expresa en una célula, y sin desear quedar ligado a la teoría, se piensa que el ADN se escinde como resultado de la formación de un dímero intramolecular de los semidominios de escisión.

En determinadas realizaciones, los componentes de una ZFN se disponen de tal manera que el dominio de dedo de zinc esté más cerca del extremo amino de la proteína de fusión, y el semidominio de escisión esté más cerca del extremo carboxilo. Esto refleja la orientación relativa del dominio de escisión en dominios de escisión de dimerización de origen natural tales como los procedentes de la enzima *Fok I*, en donde el dominio de unión al ADN está más cerca del extremo amino y el semidominio de escisión está más cerca del extremo carboxilo. En estas realizaciones, la dimerización de los semidominios de escisión para formar una nucleasa funcional se lleva a cabo uniendo las proteínas de fusión en sitios sobre cadenas de ADN opuestas, estando los extremos 5' de los sitios de unión próximos entre sí.

En esta orientación, el dedo de zinc más C terminal es proximal al semidominio de escisión de *Fok I*. Previamente se ha determinado que las proteínas de dedo de zinc no canónicas se unen a sus dianas de ADN de un modo más eficaz cuando está presente un dedo de zinc de tipo CCHC como el dedo más C terminal. Por lo tanto, es posible que la presencia de dedos de zinc de tipo CCHC previamente descritos cerca del semidominio de escisión de *Fok I* inhiba su función. Si este fuera el caso, los dedos de zinc CCHC optimizados, descrito hasta ahora, aparentemente no exhiben esta actividad inhibitoria teórica.

En realizaciones adicionales, los componentes de las proteínas de fusión (por ejemplo, fusiones ZFP-*Fok I*) se disponen de tal manera que el semidominio de escisión está más cerca del extremo amino de la proteína de fusión, y el dominio de dedo de zinc está más cerca del extremo carboxilo. En estas realizaciones, la dimerización de los semidominios de escisión para formar una nucleasa funcional se lleva a cabo uniendo las proteínas de fusión en sitios en cadenas de ADN opuestas, estando los extremos 3' de los sitios de unión próximos entre sí.

En otras realizaciones adicionales, una primera proteína de fusión contiene el semidominio de escisión más cerca del extremo amino de la proteína de fusión, y el dominio de dedo de zinc más cerca del extremo carboxilo, y una segunda proteína de fusión se dispone de tal manera que el dominio de dedo de zinc esté más cerca del extremo amino de la proteína de fusión, y el semidominio de escisión esté más cerca del extremo carboxilo. 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 dedo de zinc más cerca del extremo carboxilo localizado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio de dedo de zinc más cerca del extremo amino. Véase también el documento WO 2005/084190; cuya descripción se incorpora por referencia.

La secuencia de aminoácidos entre el dominio de dedo de zinc y el dominio de escisión (o semidominio de escisión) se indica como "enlazador ZC". El enlazador ZC es para diferenciarse de los enlaces inter-dedo anteriormente mencionados. Véanse, por ejemplo, las publicaciones de patente de Estados Unidos 20050064474A1 y 20030232410 y la publicación de patente internacional WO 2005/084190, para detalles acerca de la obtención de enlazadores ZC que optimiza la escisión.

Vectores de expresión

Un ácido nucleico que codifica una o más ZFP o proteínas de fusión ZFP (por ejemplo, ZFN) puede clonarse en un vector para la transformación en células procariontas o eucariotas para la replicación y/o expresión. Los vectores

pueden ser vectores procariotas o eucariotas, incluyendo pero sin limitación, plásmidos, vectores lanzadera, vectores de insecto, vectores binarios (véase, por ejemplo, la patente de Estados Unidos n.º 4.940.838; Horsch *et al.* (1984) *Science* 233:496-498 y Fraley *et al.* (1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803) y similares. Un ácido nucleico que codifica una ZFP también puede clonarse en un vector de expresión, para su administración en una célula de planta.

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

El promotor que se utiliza para dirigir la expresión de un ácido nucleico que codifica una ZFP depende de la aplicación particular. Por ejemplo, para la expresión y purificación de las ZFP se usa un promotor constitutivo fuerte, adecuado para para la célula hospedadora.

En cambio, cuando una ZFP se administra *in vivo* para la regulación génica de plantas (véase, "Nucleic Acid Delivery to Plant Cells" sección más adelante), se usa un promotor constitutivo o inducible, dependiendo del uso particular de la ZFP. Como ejemplos no limitantes de promotores de plantas se incluyen secuencias promotoras procedentes de ubiquitina-3 (ubi-3) de *A. thaliana* (Callis, *et al.*, 1990, *J. Biol. Chem.* 265-12486-12493); manopina sintasa (Δ mas) de *A. tumifaciens* (Petolino *et al.*, patente de Estados Unidos n.º 6.730.824); y/o virus del mosaico de la nervadura de la yuca (CsVMV) (Verdaguer *et al.*, 1996, *Plant Molecular Biology* 31:1129-1139). Véanse, también, los ejemplos.

Además del promotor, el vector de expresión contiene típicamente una unidad de transcripción o un casete de expresión que contiene todos los elementos adicionales necesarios para la expresión del ácido nucleico en las células hospedadoras, procariotas o eucariotas. Por tanto, un casete de expresión típico contiene un promotor unido operativamente, por ejemplo, a una secuencia de ácido nucleico que codifica la ZFP, y señales necesarias, por ejemplo, para la poliadenilación eficaz del transcripto, para la terminación de la transcripción, sitios de unión al ribosoma, o terminación de la traducción. Los elementos adicionales del casete pueden incluir, por ejemplo, potenciadores y señales de corte y empalme heterólogas.

El vector de expresión particular que se usa para el transporte de la información genética en la célula se selecciona con respecto al uso deseado de la ZFP, por ejemplo, expresión en plantas, animales, bacterias, hongos, protozoos, etc. (véanse los vectores expresión descritos más adelante). En la técnica se conocen vectores de expresión bacterianos y animales convencionales y se describen con detalle, por ejemplo, en la publicación de patente de Estados Unidos 20050064474A1 y en las publicaciones de patente internacional WO05/084190, WO05/014791 y WO03/080809.

Para producir líneas de células de bacterias, de mamíferos, de levaduras o de insectos, que expresen grandes cantidades de proteínas que después pueden purificarse usando técnicas convencionales, pueden usarse métodos de transcripción convencionales (véase, por ejemplo, Colley *et al.*, *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). La transformación de células eucariotas y procariotas se realiza de acuerdo con técnicas convencionales (véase, por ejemplo, Morrison, *J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds., 1983).

Para introducir secuencias de nucleótidos extrañas en dichas células hospedadoras, puede usarse cualquiera de los procedimientos bien conocidos. Estos incluyen el uso de métodos de transfección con fosfato de calcio, polibreno, fusión de protoplastos, electroporación, ultrasonido (por ejemplo, sonoporación), liposomas, microinyección, ADN desnudo, vectores plasmídicos, vectores víricos, tanto episómicos como integrativos, y cualquiera de los otros métodos bien conocidos para la introducción de ADN genómico clonado, ADNc, ADN sintético u otro material genético ajeno en una célula hospedadora (véase, por ejemplo, Sambrook *et al.*, citado anteriormente). Solo es necesario que el procedimiento de modificación por ingeniería genética particular usado pueda introducir satisfactoriamente al menos un gen en la célula hospedadora capaz de expresar la proteína de elección.

Suministro de ácido nucleico a células de plantas

Como se ha indicado anteriormente, en un hospedador de planta deseado (por ejemplo, en el genoma) pueden introducirse construcciones de ADN mediante diversas técnicas convencionales. Para revisiones de dichas técnicas véanse, por ejemplo, Weissbach y Weissbach *Methods for Plant Molecular Biology* (1988, Academic Press, N.Y.) sección VIII, págs. 421-463 y Grierson y Corey, *Plant Molecular Biology* (1988, 2d Ed.), Blackie, Londres, Cap. 7-9.

Por ejemplo, la construcción de ADN puede introducirse en una célula de planta usando técnicas tales como electroporación y microinyección de protoplastos de células de planta, o las construcciones de ADN pueden introducirse directamente en el tejido de la planta usando métodos biolísticos, tales como bombardeo de partículas

de ADN (véase, por ejemplo, Klein *et al* (1987) *Nature* 327:70-73). Como alternativa, las construcciones de ADN pueden combinarse con regiones de ADN-T flanqueantes adecuadas e introducirse en un vector hospedador convencional de *Agrobacterium tumefaciens*. En la bibliografía científica se describen bien técnicas de transformación mediada por *Agrobacterium tumefaciens*, incluyendo la desactivación y el uso de vectores binarios. Véase, por ejemplo, Horsch *et al* (1984) *Science* 233:496-498 y Fraley *et al* (1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803.

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

Las funciones de virulencia del hospedador *Agrobacterium tumefaciens* dirigirán la inserción de la construcción y del marcador adyacente en el ADN de la célula de la planta cuando la bacteria infecte la planta usando el vector de ADN T binario (Bevan (1984) *Nuc. Acid Res.* 12:8711-8721) o el procedimiento de cocultivo (Horsch *et al* (1985) *Science* 227:1229-1231). Generalmente, el sistema de transformación con *Agrobacterium* se usa para modificar genéticamente 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 con *Agrobacterium* también puede usarse para transformar, así como transferir, ADN a plantas monocotiledóneas y a células de planta. Véase la patente de Estados Unidos n.º 5.591.616; Hemalsteen *et al* (1984) *EMBO J* 3:3039-3041; Hooykass-Van Slogteren *et al* (1984) *Nature* 311:763-764; Grimsley *et al* (1987) *Nature* 325:1677-179; Boulton *et al* (1989) *Plant Mol. Biol* 12:31-40.; y Gould *et al* (1991) *Plant Physiol.* 95:426-434.

Los métodos alternativos de transferencia y transformación de genes incluyen, pero sin limitación, transformación de protoplastos a través de la captación mediada por calcio, polietilenglicol (PEG) o electroporación de ADN desnudo (véase Paszkowski *et al.* (1984) *EMBO J* 3:2717-2722, Potrykus *et al.* (1985) *Molec. Gen. Genet.* 199:169-177; Fromm *et al.* (1985) *Proc. Nat. Acad. Sci. USA* 82:5824-5828; y Shimamoto (1989) *Nature* 338:274-276) y la electroporación de tejidos de plantas (D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505). Otros métodos para la transformación de células de plantas incluyen microinyección, captación de ADN mediada por carburo de silicio (Kaeppler *et al.* (1990) *Plant Cell Reporter* 9:415-418), y bombardeo de microproyectiles (véase Klein *et al.* (1988) *Proc. Nat. Acad. Sci. USA* 85:4305-4309; y Gordon-Kamm *et al.* (1990) *Plant Cell* 2:603-618).

Los métodos y las composiciones que se desvelan pueden usarse para insertar secuencias exógenas en una localización predeterminada en un genoma de una célula de planta. Esto es útil en tanto en cuanto que la expresión de un transgén introducido en un genoma de planta depende críticamente de su sitio de integración. Por consiguiente, pueden insertarse genes que codifiquen, por ejemplo, nutrientes, antibióticos o moléculas terapéuticas, mediante recombinación dirigida, en regiones de un genoma de planta favorable para su expresión.

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

Los ácidos nucleicos introducidos en una célula de planta pueden usarse para conferir rasgos deseados esencialmente en cualquier planta. Usando las construcciones de ácido nucleico de la presente descripción y los diversos métodos de transformación mencionados anteriormente, una amplia variedad de plantas y de sistemas de células de planta, pueden modificarse genéticamente para obtener las características fisiológicas y agronómicas deseadas descritas en la presente memoria. En determinadas realizaciones, para la modificación de las plantas y de las células de plantas diana, por ingeniería genética, incluyen, pero sin limitación, plantas monocotiledóneas y dicotiledóneas, tales como cultivos, incluyendo cultivos de cereales (por ejemplo, trigo, maíz, arroz, mijo o cebada), cultivos de frutas (por ejemplo, tomate, piña, pera, fresa, naranja), cultivos de forraje (por ejemplo, alfalfa), cultivos de verduras de raíz (por ejemplo, zanahoria, patata, remolacha azucarera, ñame), cultivos de verduras de hoja (por ejemplo, lechuga, espinaca); plantas con flor (por ejemplo, petunia, rosa, crisantemo), coníferas y pinos (por ejemplo, pino, abeto, píceas); plantas que se utilizan en fitorremediación (por ejemplo, plantas acumuladoras de metales pesados); cultivos de plantas oleaginosas (por ejemplo, girasol, colza) y plantas que se utilizan con fines experimentales (por ejemplo, *Arabidopsis*). Por tanto, los métodos y composiciones desvelados tienen uso sobre una amplia serie de plantas, incluyendo, pero sin limitación, especies de los géneros *Asparagus*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucurbita*, *Daucus*, *Glycine*, *Gossypium*, *Hordeum*, *Lactuca*, *Lycopersicon*, *Malus*, *Manihot*, *Nicotiana*, *Oryza*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Solanum*, *Sorghum*, *Triticum*, *Vitis*, *Vigna* y *Zea*.

Un experto en la técnica reconocerá que después de haber incorporado el casete de expresión de manera estable en la planta transgénica y después de haber confirmado que es operativo, este puede introducirse en otras plantas por cruzamiento sexual. Dependiendo de la especie que vaya a cruzarse, puede utilizarse cualquiera de las diversas técnicas de reproducción convencionales.

5 Una célula de planta transformada, tallo, tejido o planta puede identificarse y aislarse seleccionando o explorando el material de la planta modificado genéticamente para determinar los rasgos codificados por los genes marcadores presentes en el ADN transformante. Por ejemplo, la selección puede realizarse cultivando el material de planta modificado genéticamente en medios que contengan una cantidad inhibitoria del antibiótico o herbicida contra el cual la construcción génica transformante confiere resistencia. Además, las plantas transformadas y las células de planta también pueden identificarse explorando actividades de cualquiera de los genes marcadores visibles (por ejemplo, los genes de la β -glucuronidasa, luciferasa, B o C1) que pueden estar presentes en las construcciones de ácido nucleico recombinante. Dichas metodologías de selección y exploración son muy conocidas por los expertos en la técnica.

15 También pueden usarse métodos físicos y bioquímicos para identificar transformantes de células de planta o plantas que contengan construcciones génicas insertadas. Estos métodos incluyen, pero sin limitación: 1) análisis Southern o amplificación por PCR para detectar y determinar la estructura del inserto de ADN recombinante; 2) transferencia Northern, protección con RNasa S1, amplificación por PCR con transcriptasa inversa o cebadores-extensión para detectar y examinar transcritos de ARN de las construcciones génicas; 3) ensayos enzimáticos para detectar la actividad enzimática o riboenzimática, donde dichos productos génicos están codificados por la construcción génica; 20 4) electroforesis en gel con proteínas, técnicas de transferencia Western, inmunoprecipitación o inmunoensayos ligados a enzimas, donde los productos de la construcción génica son proteínas. También pueden usarse 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 las plantas. Los métodos para realizar todos estos ensayos son muy conocidos por los expertos en la técnica.

25 Los efectos de la manipulación génica usando los métodos desvelados en la presente memoria pueden observarse, por ejemplo, mediante transferencias Northern del ARN (por ejemplo, ARNm) aislado de los tejidos de interés. Típicamente, si la cantidad de ARNm ha aumentado, puede suponerse que el gen endógeno correspondiente se está expresando a un mayor grado que antes. Pueden usarse otros métodos de medición de genes y/o actividad CYP74B. Pueden usarse diferentes tipos de ensayos enzimáticos, dependiendo del sustrato usado y del método de detección, del aumento o disminución de un producto o subproducto de reacción. Además, los niveles de proteína CYP74B expresada pueden medirse con inmunquímica, es decir, con ensayos ELISA, RIA, EIA y con otros ensayos basados en anticuerpos muy conocidos por los expertos en la técnica, tales como mediante ensayos de detección electroforética (bien con tinción o con transferencia Western). El transgén puede expresarse selectivamente en algunos tejidos de la planta o en alguna fase del desarrollo de la planta, o el transgén puede expresarse sustancialmente en todos los tejidos de la planta, sustancialmente a lo largo de todo su ciclo de vida. Sin embargo, también es aplicable cualquier modo de expresión combinatoria.

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

Las ZFP y los vectores de expresión que codifican las ZFP se administran directamente a la planta para la escisión y/o recombinación dirigida.

45 La administración de cantidades eficaces se realiza mediante cualquiera de las vías normalmente utilizadas para la introducción de ZFP en contacto sumo con la célula de la planta que va a tratarse. Las ZFP se administran de cualquier manera adecuada. Se dispone, y son muy conocidos por los expertos en la técnica, de métodos de administración adecuados de dichas composiciones y aunque puede usarse más de una vía para administrar una composición particular, a menudo una vía particular puede proporcionar una reacción más inmediata y más eficaz que otra vía.

50 También pueden usarse transportadores y éstos se determinan, en parte, por la composición particular que vaya administrarse, así como por el método particular usado para administrar la composición. Por consiguiente, se dispone de una amplia diversidad de formulaciones de composiciones farmacéuticas adecuadas (véase, por ejemplo, Remington's Pharmaceutical Sciences, 17^a ed. 1985)).

Aplicaciones

55 Las proteínas con dedos de zinc que comprenden uno o más dedos de zinc no canónicos, como se describe en la presente memoria, son útiles para la regulación de todo el genoma y para aplicaciones de edición para las que normalmente se usan las ZFP C2H2 canónicas, incluyendo pero sin limitación: activación génica; represión génica; edición genómica (escisión, inserción, reemplazo o delección dirigida); edición epigenómica (mediante el direccionamiento de modificaciones covalentes de histonas o de ADN).

Las ZFN que comprenden dedos de zinc no canónicos, como se desvela en la presente memoria, pueden usarse para escindir ADN en una región de interés en la cromatina celular (por ejemplo, en un sitio deseado o predeterminado en un genoma, por ejemplo, en un gen, bien mutante o de tipo silvestre). Para dicha escisión de ADN dirigida, un dominio de unión de dedo de zinc se modifica genéticamente para unirse a un sitio diana o cercano al sitio de escisión predeterminado, y una proteína de fusión que comprende el dominio de unión de dedo de zinc modificado genéticamente y un dominio de escisión se expresa en una célula. Después de la unión de la parte de dedo de zinc de la proteína de fusión con el sitio diana, el ADN se escinde cerca del sitio diana por el dominio de escisión. El sitio de escisión exacto puede depender de la longitud de enlazador ZC.

Como alternativa, dos ZFN, comprendiendo cada una de ellas un dominio de unión de dedo de zinc y un semidominio de escisión, se expresan en una célula, y se unen a sitios diana que están yuxtapuestos de tal manera que se reconstituye un dominio de escisión funcional y el ADN se escinde cerca de los sitios diana. En una realización, la escisión se produce entre los sitios diana de los dos dominios de unión de dedo de zinc. Uno o los dos dominios de unión de dedo de zinc pueden modificarse genéticamente.

Para la escisión dirigida usando un polipéptido de fusión de dominio de escisión - dominio de unión de dedo de zinc, el sitio de unión puede incluir el sitio de escisión, o el borde cercano del sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25, 50 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) desde el sitio de escisión. La localización exacta del sitio de unión, con respecto al sitio de escisión, dependerá del dominio de escisión particular, y de la longitud del enlazador ZC. Para los métodos en los que se usan dos polipéptidos de fusión, comprendiendo cada uno de ellos un dominio de unión de dedo de zinc y un semidominio de escisión, los sitios de unión generalmente se extienden a ambos lados del sitio de escisión. Por tanto, el borde cercano del primer sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en un lado del sitio de escisión, y el borde cercano del segundo sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en el otro lado del sitio de escisión. Los expertos en la técnica conocen métodos para la creación de mapas de sitios de escisión *in vitro* e *in vivo*.

Una vez introducida, o expresada, en la célula diana, la proteína de fusión se une a la secuencia diana y escinde en o cerca de la secuencia diana. El sitio de escisión exacto depende de la naturaleza del dominio de escisión y/o de la presencia y/o de la naturaleza de las secuencias enlazadoras entre los dominios de unión y escisión. En los casos en los que se usan ZFN, que comprenden cada uno de ellas un semidominio de escisión, la distancia entre los bordes cercanos de los sitios de unión pueden tener 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos). Los niveles de escisión óptimos también pueden depender tanto de la distancia entre los sitios de unión de las dos ZFN (véase, por ejemplo, Smith *et al.* (2000) *Nucleic Acids Res.* 28:3361-3369; Bibikova *et al.* (2001) *Mol. Cell. Biol.* 21:289-297) como de la longitud del enlazador ZC en cada ZFN. Véase, también, la publicación de patente de Estados Unidos 20050064474A1 y las publicaciones de patente internacional WO05/084190, WO05/014791 y WO03/080809.

Dos ZFN, comprendiendo cada uno de ellas un semidominio de escisión, pueden unirse en la región de interés en la misma polaridad o en polaridades opuestas, y sus sitios de unión (es decir, sitios diana) pueden separarse mediante diversos nucleótidos, por ejemplo, de 0 a 50 pares de nucleótidos o cualquier valor entero entre ellos. En determinadas realizaciones, los sitios de unión para dos proteínas de fusión, comprendiendo cada una ellas un dominio de unión de dedo de zinc y un semidominio de escisión, pueden localizarse entre 5 y 18 pares de nucleótidos de separación, por ejemplo, 5-8 pares de nucleótidos de separación o 15-18 pares de nucleótidos de separación o 6 pares de nucleótidos de separación o 16 pares de nucleótidos de separación, o a una distancia de 10 pares de nucleótidos entre sí, medida desde el borde de cada sitio de unión más cercano al otro sitio de unión, y la escisión se produce entre los sitios de unión.

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

Como se ha indicado anteriormente, una o más proteínas de fusión pueden expresarse en una célula después de la introducción, en la célula, de polipéptidos y/o polinucleótidos. Por ejemplo, dos polinucleótidos, comprendiendo cada uno de ellos secuencias que codifican uno de los polipéptidos anteriormente mencionados, pueden introducirse en una célula, y cuando los polipéptidos se expresan y cada uno de ellos se une a su secuencia diana, la escisión se produce en o cerca de la secuencia diana. Como alternativa, un solo polinucleótido que comprende secuencias que codifican los dos polipéptidos de fusión, se introduce en la célula. Los polinucleótidos pueden ser ADN, ARN o cualquiera de las formas modificadas o análogos o ADN y/o ARN.

En determinadas realizaciones, la escisión dirigida en una región genómica por una ZFN produce de la alteración de la secuencia de nucleótidos de la región, seguido de la reparación del acontecimiento de escisión por unión de extremos no homólogos (NHEJ).

5 En otras realizaciones, la escisión dirigida en una región genómica por una ZFN también puede ser parte de un procedimiento en donde una secuencia genómica (por ejemplo, una región de interés en la cromatina celular) está reemplazada con una secuencia homóloga no idéntica (es decir, mediante recombinación dirigida) mediante mecanismos dependientes de homología (por ejemplo, inserción de una secuencia donante que comprende una secuencia exógena junto con una o más secuencias que son idénticas, u homólogas pero no idénticas, con una secuencia genómica predeterminada (es decir, un sitio diana)). Dado que las roturas bicatenarias en el ADN celular estimulan varios miles de veces mecanismos de reparación celular cerca del sitio de escisión, la escisión dirigida con las ZFN como se describe en la presente memoria, permite la alteración o el reemplazo (mediante reparación dirigida por homología) de secuencias en prácticamente cualquier sitio del genoma.

15 El reemplazo dirigido de una secuencia genómica seleccionada requiere, además de las ZFN descritas en la presente memoria, la introducción de un polinucleótido exógeno (donante). El polinucleótido donante puede introducirse en la célula antes, simultáneamente con, o posterior a, la expresión de las ZFN. El polinucleótido donante contiene suficiente homología en una secuencia genómica para dar soporte a la recombinación homóloga (o reparación dirigida por homología) entre esta y la secuencia genómica con la que guarda homología. Aproximadamente 25, 50 100, 200, 500, 750, 1.000, 1.500, 2.000 nucleótidos o más de la homología de secuencia (o cualquier valor entero entre 10 y 2.000 nucleótidos o más) dará soporte a la recombinación homóloga. La longitud del polinucleótido donante puede variar de 10 a 5.000 nucleótidos (o cualquier valor entero de nucleótidos entre estos) o más.

25 Será fácilmente obvio que la secuencia de nucleótidos del polinucleótido donante no sea típicamente idéntica a la de la secuencia genómica que reemplaza. Por ejemplo, la secuencia del polinucleótido donante puede contener una o más sustituciones, inserciones, deleciones, inversiones o reordenaciones con respecto a la secuencia genómica, mientras se presente homología suficiente con las secuencias cromosómicas. Dichos cambios de secuencia pueden ser de cualquier tamaño y pueden ser cortos como un solo par de nucleótidos. Como alternativa, un polinucleótido donante puede contener una secuencia no homóloga (es decir, una secuencia exógena, para diferenciarse de un polinucleótido exógeno) flanqueado por dos regiones de homología. Adicionalmente, los polinucleótidos donantes pueden comprender una molécula vector que contenga secuencias que no son homólogas con la región de interés en la cromatina celular. Generalmente, la región (o regiones) homóloga de un polinucleótido donante tendrá al menos 50 % de identidad de secuencia con una secuencia genómica con la que se desea la recombinación. En determinadas realizaciones, se presenta una identidad de secuencia de 60 %, 70 %, 80 %, 90 %, 95 %, 98 %, 99 % o 99,9 %. Puede presentarse cualquier valor entre 1 % y 100 % de identidad de secuencia dependiendo de la longitud del polinucleótido donante.

35 Una molécula donante puede contener diversas regiones de homología discontinuas con la cromatina celular. Por ejemplo, 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 por la secuencia en la región de interés.

40 Para simplificar los ensayos (por ejemplo, hibridación, PCR, digestión con enzimas de restricción) para determinar la inserción satisfactoria de las secuencias del polinucleótido donante, pueden presentarse determinadas diferencias de secuencia en la secuencia donante en comparación con la secuencia genómica. Preferentemente, si se localiza en una región codificante, dichas diferencias en la secuencia de nucleótidos no cambiarán la secuencia de aminoácidos, o harán cambios de aminoácidos silenciosos (es decir, cambios que no afectan a la estructura o a la función de la proteína). El polinucleótido donante puede contener opcionalmente cambios en las secuencias correspondientes a los sitios de unión del dominio de dedo de zinc en la región de interés, para impedir la escisión de secuencias donantes que se han introducido en la cromatina celular mediante recombinación homóloga.

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

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

La eficacia de la inserción de las secuencias donantes por recombinación homóloga está inversamente relacionada con la distancia, en el ADN celular, entre la rotura bicatenaria y el sitio en el cual se desea la recombinación. En otras palabras, se observan eficiencias de recombinación homóloga más alta, cuando la rotura bicatenaria está más

- cerca del sitio en donde se desea la recombinación. En los casos en los que un sitio de recombinación exacto no esté predeterminado (por ejemplo, el acontecimiento de recombinación deseado puede producirse sobre un intervalo de secuencia genómica), la longitud y la secuencia del ácido nucleico donante, junto con uno o más sitios de escisión, se seleccionan para obtener el acontecimiento de recombinación deseado. En los casos en los que el acontecimiento deseado se diseña para cambiar la secuencia de un solo par de nucleótidos en una secuencia genómica, la cromatina celular se escinde en 10.000 nucleótidos en cualquier lado de ese par de nucleótidos. En determinadas realizaciones, la escisión se produce en 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 va a cambiarse.
- 5 La inserción dirigida de secuencias exógenas en una región genómica se realiza por escisión dirigida en la región genómica usando las ZFN, junto con la provisión de un polinucleótido exógeno (donante) que contiene las secuencias exógenas. El polinucleótido donante también contiene típicamente secuencias que flanquean la secuencia exógena, que contiene suficiente homología con la región genómica para dar soporte a la preparación dirigida por homología de la rotura bicatenaria en la secuencia genómica, insertando de este modo la secuencia exógena en la región genómica. Por lo tanto, el ácido nucleico donante puede tener cualquier tamaño suficiente para dar soporte a la integración de la secuencia exógena por mecanismos de reparación dependientes de homología (por ejemplo, recombinación homóloga). Sin desear quedar ligado a ninguna teoría particular, se piensa que las regiones de homología que flanquean la secuencia exógena, proporcionan los extremos de cromosomas rotos con un molde para volver a sintetizar la información genética en el sitio de la rotura bicatenaria.
- 10 La integración dirigida de secuencias exógenas, como se ha descrito anteriormente, pueden usarse para insertar un gen marcador en una localización cromosómica seleccionada. Los genes marcadores incluyen, pero sin limitación, secuencias que codifican proteínas que actúan como mediadoras en la resistencia a antibióticos (por ejemplo, resistencia a ampicilina, resistencia a neomicina, resistencia G418 resistencia a puromicina), secuencias que codifican proteínas de color o fluorescentes o luminiscentes (por ejemplo, proteína fluorescente verde, proteína fluorescente verde potenciada, proteína fluorescente roja, luciferasa) y proteínas que actúan como mediadoras en el crecimiento celular y/o la amplificación génica (por ejemplo, dihidrofolato reductasa) potenciada. Los genes marcadores ejemplares incluyen, pero sin limitación, β -glucuronidasa (GUS), fosfinotricin N-acetil transferasa (PAT, BAR), neomicin fosfotransferasa, β -lactamasa, catecol dioxigenasa, α -amilasa, tirosinasa, β -galactosidasa, luciferasa, aequorina, EPSP sintasa, nitrilasa, acetolactato sintasa (ALS), dihidrofolato reductasa (DHFR), deshalogenasa dalapón y antranilato sintasa. En determinadas realizaciones, la integración dirigida se usa para insertar una construcción de expresión de ARN, por ejemplo, secuencias responsables de la expresión regulada de micro ARN o ARNi. En una construcción de expresión de ARN también pueden incorporarse promotores, potenciadores y secuencias reguladoras de la transcripción adicionales.
- 20 Se obtienen aumentos adicionales en cuanto a la eficacia de la recombinación dirigida, en células que comprenden una molécula de fusión de dedo de zinc/nucleasa y una molécula de ADN donante, bloqueando las células en la fase G₂ del ciclo celular, cuando los procesos de reparación activados por homología están máximamente activos. Dicha detención pueden realizarse de diversas maneras. Por ejemplo, las células pueden tratarse, por ejemplo, con fármacos, compuestos y/o moléculas pequeñas que influyen en la progresión del ciclo celular de tal manera que detienen a las células en la fase G₂. Las moléculas ejemplares de este tipo incluyen, pero sin limitación, compuestos que afectan a la polimerización de microtúbulos (por ejemplo, vinblastina, nocodazol, Taxol), compuestos que interactúan con el ADN (por ejemplo, *cis-platino* diamina dicloro (II), cisplatino, doxorubicina) y/o compuestos que afectan a la síntesis del ADN (por ejemplo, timidina, hidroxurea, L-mimosina, etopósido, 5-fluorouracilo). Aumentos adicionales en cuanto a la eficacia de la recombinación se consiguen usando inhibidores de histona desacetilasa (HDAC) (por ejemplo, butirato de sodio, trichostatina A) que alteran la estructura de la cromatina para hacer que el ADN genómico sea más accesible a la maquinaria de recombinación celular.
- 35 Como métodos adicionales para la detención de ciclo celular se incluyen la sobreexpresión de proteínas que inhiben la actividad de las quinasas CDK del ciclo celular, por ejemplo, introduciendo en la célula un ADNc que codifica la proteína o introduciendo en la célula una ZFP modificada genéticamente que activa la expresión del gen que codifica la proteína. La detección del ciclo celular también se realiza inhibiendo la actividad de ciclinas y de CDK, por ejemplo, usando métodos de ARNi (por ejemplo, patente de Estados Unidos n.º 6.506.559) o introduciendo en la célula una ZFP modificada que reprime la expresión de uno o más genes implicados en la progresión del ciclo celular, tales como, por ejemplo, geles de ciclina y/o CDK. Véase, por ejemplo, la patente de Estados Unidos n.º 6.534.261, del mismo propietario que la presente, para métodos para la síntesis de proteínas de dedo de zinc modificadas genéticamente para la regulación de la expresión génica.
- 50 Como se ha descrito anteriormente, los métodos y las composiciones desvelados para la escisión dirigida pueden usarse para inducir mutaciones en una secuencia genómica. La escisión dirigida también puede usarse para crear genes inactivados (*knockt-outs*) (por ejemplo, para genómica funcional o validación diana) y para facilitar la inserción dirigida de una secuencia en un genoma (es decir, genosustitución (*gene knock-in*)). La inserción puede ser mediante reemplazo de secuencias cromosómicas a través de recombinación homóloga o mediante integración dirigida, en la cual una nueva secuencia (es decir, una secuencia no presente en la región de interés), flanqueada por secuencias homólogas a la región de interés en el cromosoma, se inserta en un sitio diana predeterminado. Pueden usarse los mismos métodos para reemplazar una secuencia de tipo silvestre con una secuencia mutante, o para transformar un
- 60

alelo en un alelo diferente.

La escisión dirigida de fitopatógenos de infección o integrados puede usarse para tratar infecciones patógenas en un hospedador de planta, por ejemplo, escindiendo el genoma del patógeno de tal manera que su patogenicidad se reduce o elimina. De manera adicional, la escisión dirigida de genes que codifican receptores para virus de plantas puede usarse para bloquear la expresión de dichos receptores, impidiendo de este modo la infección vírica y/o la propagación vírica en la planta.

Como ejemplos de fitopatógenos se incluyen, pero sin limitación, virus de plantas, tales como, *Alfavirus*, *Alfavirus*, *Badnavirus*, *Betacriptovirus*, *Bigeminivirus*, *Bromovirus*, *Bimovirus*, *Capillovirus*, *Carlavirus*, *Carmovirus*, *Caulimovirus*, *Closterovirus*, *Comovirus*, *Cucumovirus*, *Citorhabdovirus*, *Diantovirus*, *Enamovirus*, *Fabavirus*, *Fijivirus*, *Furovirus*, *Hordeivirus*, *Hibrigeminivirus*, *Idaeovirus*, *Iarvirus*, *Ipomovirus*, *Luteovirus*, *Maclomovirus*, *Macluravirus*, *Marafivirus*, *Monogeminivirus*, *Nanavirus*, *Necrovirus*, *Nepovirus*, *Nucleorhabdovirus*, *Orizavirus*, *Ourmiavirus*, *Fitoreovirus*, *Potexvirus*, *Potivirus*, *Rimovirus*, ARN satélite, *satellivirus*, *Sequivirus*, *Sobemovirus*, *Tenuivirus*, *Tobamovirus*, *Tobravirus*, *Tombusvirus*, *Tospovirus*, *Tricovirus*, *Timovirus*, *Umbravirus*, *Varicosavirus* y *Waikavirus*; patógenos fúngicos tales como tizones (por ejemplo, *Ustilaginales*), royas (*Uredinales*), cornezuelos (*Claviceps pupurea*) y mildiu; mohos (*Oomycetes*) tales como *Phytophthora infestans* (añublo de la patata); patógenos bacterianos tales como *Erwinia* (por ejemplo, *E. herbicola*), *Pseudomonas*. (por ejemplo, *P. aeruginosa*, *P. syringae*, *P. fluorescense* y *P. putida*), *Ralstonia* (por ejemplo, *R. solanacearum*), *Agrobacterium* y *Xanthomonas*; gusanos redondos (*Nematoda*); y Phytomyxa (*Polymyxa* y *Plasmodiophora*).

Los métodos desvelados para la recombinación dirigida pueden usarse para reemplazar cualquier secuencia genómica con una secuencia homóloga, no idéntica. Por ejemplo, una secuencia genómica mutante puede reemplazarse con su homólogo de tipo silvestre, proporcionando de este modo métodos para el tratamiento de enfermedades de plantas; proporcionando resistencia a fitopatógenos; aumentando el rendimiento de los cultivos, etc. De la misma manera, un alelo de un gen puede reemplazarse por un alelo diferente usando los métodos de recombinación dirigida desvelados en la presente memoria.

En muchos de estos casos, una región de interés comprende una mutación y el polinucleótido donante comprende la secuencia de tipo silvestre correspondiente. De manera similar, una secuencia genómica de tipo silvestre puede reemplazarse por una secuencia mutante, si esto fuese deseable. De hecho, cualquier patología dependiente de una secuencia genómica particular, en alguna manera, puede corregirse o aliviarse usando los métodos y las composiciones desveladas en la presente memoria.

La escisión dirigida y la recombinación dirigida también puede usarse para alterar secuencias no codificantes (por ejemplo, secuencias reguladoras tales como promotores, potenciadores, iniciadores, terminadores, sitios de corte y empalme) para alterar los niveles de expresión de un producto génico. Dichos métodos pueden usarse, por ejemplo, para fines terapéuticos, para alteraciones en la fisiología y la bioquímica celular, genómica funcional y/o estudios de valoración diana.

Los métodos y las composiciones descritos en la presente memoria también pueden usarse para la activación y represión de la expresión génica usando fusiones entre un dominio de unión de dedo de zinc no canónico y un dominio funcional. Dichos métodos se desvelan, por ejemplo, en las patentes de Estados Unidos 6.534.261; 6.824.978 y 6.933.113, del mismo propietario que la presente.

Los métodos de represión adicionales incluyen el uso de oligonucleótidos antisentido y/o ARN de interferencia pequeño (ARNip o ARNi) dirigido a la secuencia del gen a reprimir.

En realizaciones adicionales, puede usarse una o más fusiones entre un dominio de unión de dedo de zinc y una recombinasa (o fragmento funcional de la misma), además de o en lugar de las fusiones de dominio de escisión -dedo de zinc desveladas en la presente memoria, para facilitar recombinación dirigida. Véase, por ejemplo, la patente de Estados Unidos n.º 6.534.261 del mismo propietario que la presente y Akopian *et al.* (2003) Proc. Natl. Acad. Sci. USA 100:8688-8691.

En realizaciones adicionales, los métodos y las composiciones desvelados se usan para proporcionar fusiones de dominios de unión a ZFP con dominios de activación transcripcional o de represión que requieren dimerización (bien homodimerización o heterodimerización) para su actividad. En estos casos, un polipéptido de fusión comprende un dominio de unión de dedo de zinc y un monómero de dominio funcional (por ejemplo, un monómero de un dominio de activación transcripcional o de represión dimerico). La unión de dichos dos polipéptidos de fusión en sitios diana correctamente situados permite la dimerización de tal manera que se reconstituye un dominio de represión o activación de la transcripción funcional.

Ejemplos

La presente invención se define adicionalmente en los siguientes ejemplos, en los que todas las partes y porcentajes se dan en peso y en grados Celsius, salvo que se indique otra cosa. Debe entenderse que, aunque estos ejemplos, indiquen determinadas realizaciones de la invención, se ofrecen solo como ilustrativos.

Ejemplo 1: Vectores de expresión ZFN

Los vectores de expresión que comprendían las secuencias que codificaban las ZFN con 4 dedos (denominadas "5-8" y "5-9") como se describe en los Ejemplos 2 y 14 de la publicación de patente de Estados Unidos 005/0064474 (véase el Ejemplo 2 de esa solicitud) se modificaron de la siguiente manera. Resumiendo, las ZFN 5-8 y 5-9 (que comprendían dominios de 4 dedos de zinc fusionados al dominio nucleasa de la enzima de restricción *Fok I* de tipo IIS (aminoácidos 384-579 de la secuencia de Wah *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:10564-10569) mediante un enlazador ZC de cuatro aminoácidos) se modificaron en una estructura CCHC. También se realizaron modificaciones adicionales (sustituciones e inserciones) en restos entre las estructuras de coordinación de zinc His y Cys en C terminal y/o C terminal a la Cys C terminal en el dedo 2 y/o dedo 4.

Ejemplo 2: Corrección génica de eGFP en líneas de células indicadoras

La capacidad de las ZFN que comprenden dedos de zinc CCHC como se describe en la presente memoria para facilitar la recombinación homóloga se ensayó en el sistema GFP descrito en Umov (2005) Nature 435(7042):646-51 y en la solicitud de patente de Estados Unidos 20050064474 (por ejemplo, Ejemplos 6-11). En resumen, 50 ng de cada ZFN y 500 ng del donante GFP menos el promotor (Urnov (2005) Nature) se transfectaron en 500.000 células indicadoras, usando 2 ul de Lipofectamine 2000 por muestra, siguiendo el protocolo de Invitrogen Lipofectamine 2000.

Se añadió vinblastina 24 horas después de la transfección a una concentración final de 0,2 uM, y se retiró 72 horas después de la transfección.

Las células se ensayaron con respecto a la expresión de GFP 5 días después de la transfección midiendo 40.000 células por transfección en el analizador FACS de sobremesa de Guava.

Como se muestra en la FIG. 1, la mayoría de las ZFN que comprendían dedos de zinc CCHC alterados como se muestra en las tablas 1 y 2 anteriores facilitan la recombinación homóloga en el locus indicador (GFP), dando como resultado la expresión de GFP a niveles por encima de los dedos de zinc CCHC no modificados y diversos se comportaron comparablemente a las ZFN que comprendían dedos de zinc CCHH. Cuando la variante de comportamiento óptimo se posicionó en el dedo 4 (D4) comprendía la siguiente secuencia (incluyendo y C terminal al resto de coordinación de zinc His): HAQRCGLRGSQVLV (SEQ ID NO: 53) (el dedo de zinc en la tabla 2 indicado como n.º 21 y mostrado en la FIG. 1 como "2-21"). Cuando la variante de comportamiento óptima se posicionó en el dedo 2 (D4) comprendía la siguiente secuencia (incluyendo y C terminal al resto de coordinación de zinc His): HAQRCGLRGSQVLV (SEQ ID NO: 75) (el dedo de zinc en la tabla 2 indicado como n.º 43 y mostrado en la FIG. 1 como "2-43").

Ejemplo 3: Edición de un gen IL2R γ cromosómico por recombinación dirigida

Las ZFN descritos en la presente memoria también se ensayaron en el ensayo de IL2R γ endógeno descrito en Umov (2005) Nature 435(7042):646-51 y en el Ejemplo 2 de la publicación de patente de Estados Unidos n.º 20050064474. En resumen, 2,5 microgramos de cada construcción de expresión ZFN se transfectaron en 500.000 células K562 usando un Nucleofector (Amaxa). El ADN genómico se recogió y la alteración génica se ensayó en el locus IL2R γ endógeno usando el kit de endonucleasa Surveyor.

Las ZFN se muestran en la parte izquierda superior de la FIG. 2. En particular, el dedo de zinc alterado 20 se refiere a un dedo de zinc CCHC que comprende la secuencia HTRRCGLRGSQVLV; el dedo de zinc 21 comprende la secuencia HAQRCGLRGSQVLV (SEQ ID NO: 53); el dedo de zinc 43 comprende la secuencia HAQRCGLRGSQVLV (SEQ ID NO: 75); el dedo de zinc 45 comprende la secuencia HIRTGCTGSQKP; el dedo de zinc 47 comprende la secuencia HIRTGCTGSQKP; y el dedo de zinc 48 comprende la secuencia HIRRGCTGSQKP. Los dedos de zinc 20 y 21 se usaron en el dedo 4 de las ZFN de 4 dedos y los dedos de zinc 43, 45, 47 y 48 se usaron en el dedo 2 de las ZFN de 4 dedos.

Los pares de ZFN ensayados se muestran en la FIG. 2 anterior y en la derecha del gráfico y en la tabla 5:

Tabla 5

| Muestra n.º | ZFN 5-8 | ZFN 5-9 |
|-------------|-----------------------|-----------------------|
| 1 | ninguno (GFP) | |
| 2 | tipo silvestre (CCHH) | tipo silvestre (CCHH) |
| 3 | 43 (dedo 2) | 43 (dedo 2) |
| 4 | 43 (dedo 2) | 20 (dedo 4) |
| 5 | 43 (dedo 2) | 21 (dedo 4) |
| 6 | 43 (dedo 2) | 45 (dedo 2) |
| 7 | 43 (dedo 2) | 47 (dedo 2) |

| Muestra n.º | ZFN 5-8 | ZFN 5-9 |
|-------------|-------------|-------------|
| 8 | 20 (dedo 4) | 43 (dedo 2) |
| 9 | 21 (dedo 4) | 43 (dedo 2) |
| 10 | 45 (dedo 2) | 43 (dedo 2) |
| 11 | 47 (dedo 2) | 43 (dedo 2) |
| 12 | 48 (dedo 2) | 43 (dedo 2) |

Para determinar si las mutaciones se habían inducido en el sitio de escisión, el producto de amplificación se analizó usando un ensayo Cel-1, en donde el producto de amplificación se desnaturaliza y renaturaliza, seguido de tratamiento con la nucleasa Cel-1 específica de emparejamiento erróneo. Véase, por ejemplo, Oleykowski *et al.* (1998) *Nucleic Acids res.* **26**:4597-4602; Qui *et al.* (2004) *BioTechniques* **36**:702-707; Yeung *et al.* (2005) *BioTechniques* **38**:749-758.

En la FIG. 2 se muestran resultados de dos experimentos de cada ejemplo. El experimento n.º 2 de las muestras 8 y 9 tuvieron interferencia de fondo significativa en los carriles que redujeron la eficacia aparente de estas ZFN.

Como se muestra en la FIG. 2, determinados variantes de CCHC eran esencialmente equivalentes a las ZFN C2H2 de tipo silvestre. El dedo de zinc 21 en el dedo 4 (muestras 5 y 9) produjeron mejores resultados que el dedo de zinc 20 en el dedo 4 (muestras 4 y 8). En el dedo 2, el dedo de zinc 43 produjo los mejores resultados.

Ejemplo 4: Corrección génica de eGFP en líneas de células indicadoras

Basándose en los resultados mostrados en las FIG. 1 y 2, se produjeron los dedos de zinc CCHC mostrados en las tablas 3 y 4 anteriores (indicados 1a a 10a). Estos dedos de zinc se incorporaron en las ZFN 5-8 y 5-9 y se ensayaron en el ensayo de corrección génica de GFP descrito en el Ejemplo 2 anterior. Los pares ZFN ensayados en cada muestra se muestran debajo de cada barra, donde los números de los dedos de zinc 20, 21, 43, 45, 47 y 48 son los descritos en el Ejemplo 3 y los dedos de zinc CCHC 1a a 10a comprenden la secuencia mostrada en las tablas 3 y 4 anteriores. Los dedos de zinc 20, 21, 7a, 8a, 9a y 10a se usaron en el dedo 4; los dedos de zinc 43, 45, 47, 48, 1a, 2a, 3a, 4a, 5a y 6a se usaron en el dedo 2.

En la FIG. 3 se muestran los resultados. La fila superior debajo de cada barra se refiere al dedo de zinc incorporado en ZFN 5-8 y la fila superior debajo de cada barra se refiere al dedo de zinc incorporado en ZFN 5-9. Por ejemplo, la 2ª barra de la izquierda sobre el gráfico de la FIG. 3 se refiere a una muestra transfectada con ZFN 5-8 y 5-9 en donde el D4 de ambas ZFN comprende la secuencia de dedo de zinc 20. Como se muestra, muchas de las ZFN que comprenden dedos de zinc CCHC se comportaron de un modo comparable con las ZFN de (CCHH) tipo silvestre.

Ejemplo 5: Diseño y generación del vector diana

A. Estructura general de la secuencia diana

La construcción diana para el tabaco (una dicotiledónea) incluía los 7 componentes siguientes como se muestra en las FIG. 4 y 5: i) un casete de expresión de higromicina fosfotransferasa (HPT) que comprendía un promotor de ubiquitina-3 (ubi-3) de *A. thaliana* (Callis, *et al.*, 1990, *J. Biol. Chem.* 265:12486-12493) que dirige el gen HPT de *E. coli* (Waldron *et al.*, 1985, *Plant Mol. Biol.* 18:189-200) terminado por una fase de lectura abierta 24 (orf-24) de *A. tumifaciens* en la región no traducida 3' (UTR) (Gelvin *et al.*, 1987, EP222493); ii) la secuencia-1 homóloga, que comprendía la región de unión a la matriz (MAR) RB7 de *N. tabacum* (Thompson *et al.*, 1997, WO9727207); iii) un fragmento génico de la proteína fluorescente verde (GFP) 5' (Evrogen Joint Stock Company, Moscú, Rusia) dirigido por un promotor modificado de la manopina sintasa (Δ mas) de *A. tumifaciens* (Petolino *et al.*, patente de Estados Unidos n.º 6.730.824); iv) un casete de expresión de la β -glucuronidasa (GUS) que comprendía un promotor del virus del mosaico de la nervadura de la yuca (CsVMV) (Verdaguer *et al.*, 1996, *Plant Molecular Biology* 31:1129-1139) que dirige un gen GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) terminado por la UTR 3' de la nopalina sintasa (nos) de *A. tumifaciens* (DePicker *et al.*, 1982, *J. Mol. Appl. Genet.* 1:561-573); v) un fragmento génico GFP 3' (Evrogen Joint Stock Company, Moscú, Rusia) terminado por una orf-1 en la UTR 3' de *A. tumifaciens* (Huang *et al.*, *J. Bacteriol.* 172:1814-1822); vi) la secuencia 2 homóloga, que comprende el intrón 1 de la 4-coumaroil-CoA sintasa (4-CoAS) de *A. thaliana* (Locus At3g21320, GenBank NC 003074) y; vii) un fragmento génico en 3' de la fosfinotricin fosfotransferasa (PAT) de *S. viridochromogenes* (Wohlleben *et al.*, 1988, *Gene* 70:25-37) terminado por una ORF-25/26 en la UTR3' de *A. tumifaciens* (Gelvin *et al.*, 1987, EP222493).

Se insertó un sitio de unión de la proteína de fusión dedo de zinc-FokI (IL-1-L0-FokI) (Urnov *et al.*, 2005, US 2005/0064474) aguas abajo del promotor CsVMV (Verdaguer *et al.*, 1996, *Plant Molecular Biology* 31:1129-1139) y se fusionó con la secuencia codificante GUS (Jefferson, 1987, *Plant Mol. Biol. Rep.* 5:387-405) en el extremo N. Dos copias de un segundo sitio de unión de la proteína de fusión dedo de zinc-FokI (Scd27-L0-FokI) (Urnov *et al.*, 2005, US 2005/0064474) flanqueaba los fragmentos del gen GFP en 5' y 3' (Evrogen Joint Stock Company, Moscú, Rusia). Cada sitio de unión contenía cuatro repeticiones en tándem de la secuencia de reconocimiento de la proteína de fusión dedo de zinc-FokI particular, de tal manera que cada sitio de unión tenía un tamaño de ~ 200 pb (FIG. 6A).

Esto se diseñó para garantizar que las secuencias de reconocimiento fuesen accesibles a la proteína de fusión dedo de zinc-FokI en el entorno de cromatina complejo. Cada secuencia de reconocimiento incluía una secuencia de repetición invertida a la que se unía una sola proteína de fusión de dedo de zinc-FokI como un homodímero y escindía el ADN bicatenario (FIG. 6B). Los fragmentos del gen GFP 5' y 3' solapaban en 540 pb proporcionando homología dentro de la secuencia diana y se insertó un codón de terminación en el extremo 3' del fragmento GFP 5' para garantizar la traducción no funcional de GFP de la secuencia diana.

El vector de transformación que comprendía la secuencia diana se generó a través de un proceso de clonación multietapa como se describe más adelante.

B. Construcción del vector binario HPT (pDAB1584)

El vector pDAB1400, que contenía un casete de expresión GUS, comprendía un promotor de ubi-3 de *A. thaliana* (Callis, *et al.*, 1990, J. Biol. Chem. 265:12486-12493) que dirige el gen GUS (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) terminado por una orf-1 UTR de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822), se usó como la construcción base de partida (FIG. 7).

Para evitar cualquier elemento regulador repetido innecesario en la construcción diana, la orf-1 UTR de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822) en pDAB1400 se reemplazó con una orf-24 UTR de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493), que se escindió de pDAB782 (FIG. 8) como un fragmento SacI/XbaI y se clonó en los mismos sitios en pDAB1400. La construcción resultante contenía un promotor de ubi-3 de *A. thaliana* (Callis, *et al.*, 1990, J. Biol. Chem. 265:12486-12493) que dirige al gen GUS (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) terminado por una orf-24 UTR de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493) y se denominó pDAB1582 (FIG. 9).

La secuencia codificante HPT (Waldron *et al.*, 1985, Plant Mol. Biol. 18:189-200) se amplificó por PCR a partir del plásmido pDAB354 (FIG. 10) usando los cebadores (*primers*) P1 y P2. Se añadió un sitio de BbsI en el extremo 5' del cebador P1 y el sitio SacI se mantuvo en el extremo 3' del cebador P2. El fragmento de PCR HPTII se sometió a digestión con BbsI/SacI y se clonó en pDAB1582 que se sometió a digestión con NcoI-SacI para reemplazar el gen GUS con el gen HPT del fragmento PCR. El plásmido resultante se denominó pDAB1583 (FIG. 11).

El fragmento de ubi-3 de *A. thaliana*/HPT/ orf-24 de *A. tumefaciens* se escindió después de pDAB1583 por digestión con NotI y se trató con ADN polimerasa de T4 para generar extremos romos. El casete de expresión tratado con HPT de extremos romos se clonó en pDAB2407 (FIG. 12), un vector base binario, en el sitio PmlI dando como resultado el plásmido DAB1584 (FIG. 13).

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

La orf-1 UTR de *A. tumefaciens* orf-1 (Huang *et al.*, J. Bacteriol. 172:1814-1822) en pDAB2418 (FIG. 14) se reemplazó con la orf25/26 UTR de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493) para impedir secuencias reguladoras repetidas en el vector diana. Para hacer el intercambio de la UTR, la orf25/26 UTR de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493) se amplificó por PCR a partir del plásmido pDAB4045 (FIG. 15) usando los cebadores P3 y P4. Se añadieron los sitios SmaI y AgeI en el extremo 3' del fragmento PCR y el sitio SacI se mantuvo en el extremo 5'. El ADN plasmídico pDAB2418 que contenía un casete de expresión del gen PAT que comprendía el promotor de ubiquitina-10 (ubi-10) de *A. thaliana* (Callis, *et al.*, 1990, J. Biol. Chem. 265:12486-12493) que dirige el gen PAT (Wohlleben *et al.*, 1988, Gene 70:25-37) terminado por la orf-1 UTR de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822) y una secuencia MAR RB7 de *N. tabacum* (Thompson *et al.*, 1997, WO9727207), se sometió a digestión con SacI y AgeI y se recuperaron los dos fragmentos más grandes. Estos fragmentos se ligaron con el producto de la PCR orf-25/26 UTR de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493) digerido con SacI y AgeI. El plásmido resultante se denominó pDAB1575 (FIG. 16). La región de unión a la matriz (MAR) RB7 de *N. tabacum* (Thompson *et al.*, 1997, WO9727207) sirve como secuencia-1 homóloga en el vector diana.

Se seleccionó el intrón-1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBankNC 003074) que servía como secuencia-2 homóloga en el vector diana. La secuencia codificante del gen PAT (Wohlleben *et al.*, 1988, Gene 70:25-37) se analizó y los 299/300 pb aguas abajo del codón de inicio se identificaron como el sitio de inserción del intrón de tal manera que se formarían los sitios de corte y empalme 5' y 3' apropiados. Después, el intrón de longitud completa se fusionó con 253 pb de la secuencia codificante PAT 3' parcial por síntesis de ADN (Picoscript Ltd., LLP, Houston, Texas). Se añadieron los sitios NotI y SacI en el extremo 5' y 3' del fragmento de ADN, respectivamente. El fragmento de ADN sintetizado se digirió después con NotI/SacI y se insertó en pDAB1575 en los mismos sitios para reemplazar la secuencia codificante PAT de longitud completa. La construcción resultante se denominó pDAB1577 (FIG. 17).

Se sintetizó un fragmento de ADN de 241 pb que contenía 4 repeticiones en tándem del sitio de reconocimiento Scd27-L0-FokI (FIG. 6) (Picoscript Ltd., LLP, Houston, Texas) con un sitio SmaI añadido en ambos extremos 5' y 3' del fragmento. El fragmento que contenía el sitio de unión FokI-dedo de zinc sintetizado se digirió después con SmaI y se insertó en pDAB1577 en el sitio MscI. El vector resultante se denominó pDAB1579 (FIG. 18). Después, se

insertó un segundo fragmento que contenía el sitio de unión FokI-dedo de zinc digerido con SmaI en el pDAB1579 en el sitio SmaI. La construcción resultante se denominó pDAB1580 (FIG. 19). Este vector contiene secuencias homólogas 1 y 2 (MAR RB7 de *N. tabacum* e intrón 1 de 4-CoAS de *A. thaliana* 4, respectivamente) y dos sitios de unión de dedo de zinc-FokI Scd27 sintetizados, conteniendo cada uno de ellos 4 repeticiones en tándem de sitios de reconocimiento Scd27-L0-FokI.

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

El gen GFP, CopGFP, se adquirió en Evrogen Joint Stock Company (Moscú, Rusia) y la secuencia codificante de longitud completa se amplificó por PCR usando los cebadores P5 y P6. Se añadieron los sitios BbsI y SacI en los extremos 5' y 3' del producto de la PCR, respectivamente. Después el producto de la PCR CopGFP se digirió con BbsI/SacI y se clonó en pDAB3401 (FIG. 20) que comprendía el promotor Δ mas de *A. tumefaciens* modificado (Petolino *et al.*, US6730824) que dirige el gen GUS (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) y termina por una orf-1 UTR 3' de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822) en los sitios NcoI/SacI para reemplazar el gen GUS. El vector resultante se denominó pDAB1570 (FIG. 21).

Para construir los dos fragmentos GFP no funcionales, parcialmente duplicados, un fragmento de ADN, que contenía la mayor parte de la secuencia codificante de CopGFP con una delección de 47 pb en el extremo 5', se amplificó por PCR usando los cebadores P9 y P10. Se añadió un sitio ApaI en ambos extremos 5' y 3' y se añadió un sitio StuI adicional en el extremo 5' aguas abajo del sitio ApaI. Después, el producto de la PCR se digirió con ApaI y se insertó en pDAB1570 en el sitio ApaI, creando de este modo dos fragmentos GFP no funcionales en el mismo vector con una secuencia duplicada de 540 pb. La construcción resultante se denominó pDAB1572 (FIG. 22).

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

Un fragmento de ADN de 233 pb que contenía 4 repeticiones en tándem del sitio de reconocimiento IL-1_L0-FokI (FIG. 6) se sintetizó por Picoscript Ltd., LLP, (Houston, Texas) añadiéndose los sitios NcoI y AflIII en los extremos 5' y 3', respectivamente. Después, el fragmento sintetizado se digirió con NcoI/AflIII y se insertó en pDAB4003 (FIG. 23), que contenía un gen GUS (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405) dirigido por un promotor de CsVMV (Verdaguer *et al.*, 1996, Plant Molecular Biology 31:1129-1139) terminado por la orf-1 en la UTR 3' de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822) en el sitio NcoI. Después, se generó una fusión N-terminal entre el sitio de unión IL-1_Lo-FokI y la secuencia codificante GUS. El vector resultante se denominó pDAB1571 (FIG. 24).

Para impedir elementos de repetición en la UTR 3' en el vector diana, la UTR 3' de *nos* de *A. tumefaciens* (DePicker *et al.*, 1982, J. Mol. Appl. Genet. 1:561-573) se escindió de pDAB7204 (FIG. 25) como un fragmento SacI/PmeI y se clonó en pDAB1571, que se digirió con SacI/NaeI, para reemplazar la orf-1 UTR 3' de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822). El plásmido resultante se denominó pDAB1573 (FIG. 26).

F. Construcción del vector diana final (pDAB1585)

Para construir el vector diana final, el casete de expresión GUS con la inserción del sitio diana de la proteína de fusión IL-1-FokI se escindió de pDAB1573 por digestión con NotI, se trataron los extremos romos y se insertó en pDAB1572 en el sitio StuI. El vector intermedio resultante se denominó pDAB1574 (FIG. 27). Todo el casete que contenía el promotor Δ mas modificado (Petolino *et al.*, US6730824), una secuencia GFP parcialmente duplicada en 5' (Evrogen Joint Stock Company, Moscú, Rusia), el promotor CsVMV (Verdaguer *et al.*, 1996, Plant Molecular Biology 31:1129-1139), una secuencia diana de proteína de fusión IL-1-FokI, la región codificante del gen GUS (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405), una UTR 3' *nos* de *A. tumefaciens* (DePicker *et al.*, 1982, J. Mol. Appl. Genet. 1:561-573), una GFP parcialmente duplicada en 3' (Evrogen Joint Stock Company, Moscú, Rusia) y una orf-1 UTR 3' de *A. tumefaciens* (Huang *et al.*, J. Bacteriol. 172:1814-1822) se escindió de pDAB1574 y se insertó en pDAB1580 en el sitio NotI. El plásmido resultante se denominó pDAB1581 (FIG. 28). Después, el fragmento AgeI de pDAB1581 se insertó en pDAB1584 en el sitio AgeI creando de este modo la construcción diana final, pDAB1585 (FIG. 4 y 5).

Ejemplo 6: Generación de líneas celulares transgénicas con secuencias diana integradas

Se usó un cultivo en suspensión de células de tabaco, denominado BY2, en donde las secuencias diana del Ejemplo 5 se integraron de manera estable mediante transformación con *Agrobacterium*. La línea celular base, BY2, se obtuvo en Jun Ueki de Japan Tobacco, Iwata, Shizuoka, Japón. Este cultivo prolifera como células de 5-10 μ de diámetro en grupos de 100-150 células con un tiempo de generación de aproximadamente 18 horas. Los cultivos de suspensión de células BY2 se mantuvieron en medios que contenían sales basales LS (Phyto Technology Labs L689), KH_2PO_4 170 mg/l, sacarosa 30 g/l, 2,4-D 0,2 mg/l y tiamina-HCL 0,6 mg/l a un pH de 6,0. Las células BY2 se subcultivaron cada 7 días añadiendo 50 ml de medio basado en LS a un VCC 0,25 ml. El cultivo de suspensión de células BY2 se mantuvo en matraces de 250 ml en un agitador rotatorio a 25 °C y a 125 RPM.

Para generar el cultivo de células BY2 transgénico con secuencias diana integradas, un matraz de una suspensión de tabaco post subcultivo de cuatro días se dividió en 10-12 alícuotas de cuatro ml que se cultivaron conjuntamente en placas de Petri de 100 x 25 mm con 100 μ l de la cepa LBA4404 de *Agrobacterium* que llevaba pDAB1585 en

crecimiento durante una noche a una $DO_{600} \sim 1,5$. Las placas se envolvieron con *parafilm* y se incubaron a 25 °C sin agitación durante 3 días, después de los cuales se retiró el líquido sobrante y se reemplazó con 11 ml de medio basal basado en LS que contenía carbenicilina 500 mg/l.

5 Después de la resuspensión de las células de tabaco, se dispensó 1 ml de suspensión sobre placas de 100 x 25 mm de medio base apropiado que contenía carbenicilina 500 mg/l e higromicina 200 mg/l solidificado con agar TC 8 g/l y se incubó sin envolver a 28 °C en la oscuridad. Esto produjo 120-144 placas de selección para un solo tratamiento. Los aislados individuales resistentes a higromicina aparecieron 10-14 días después de la siembra en las placas y se transfirieron a placas individuales de 60x20 mm (un aislado por placa) donde se mantuvieron como callos en un programa de subcultivo de 14 días hasta necesitarse para análisis y experimentos de retransformación posteriores

10 **Ejemplo 7: Exploración y caracterización en acontecimientos transgénicos diana**

Los acontecimientos transgénicos resistentes a higromicina generados de la transformación del vector diana en cultivos de células de tabaco BY2, como se describe en el Ejemplo 6, se analizaron de la siguiente manera.

15 Los análisis iniciales realizados para la exploración de estos acontecimientos transgénicos incluyeron análisis de expresión de GUS para indicar la accesibilidad de la secuencia diana, análisis PCR de la secuencia diana de longitud parcial y completa para confirmar la presencia y la integridad del vector diana y análisis de transferencia Southern para determinar el número de copias de la secuencia diana integrada. Un subconjunto de los acontecimientos transgénicos que mostró expresión de GUS contenía una sola copia de la secuencia diana de longitud completa; estos se seleccionaron para volver a establecer cultivos en suspensión para generar las líneas diana para una retransformación posterior. Estas líneas diana restablecidas también se sometieron a caracterización adicional, lo que incluyó un análisis de transferencia Southern más minucioso, confirmación de la secuenciación de todo el inserto diana, y análisis de la secuencia genómica flanqueante.

20 Los cultivos en suspensión o de tejido de callo de tabaco transgénico, iniciados a partir de los acontecimientos seleccionados, se analizaron con respecto a la actividad de GUS incubando muestras de 50 mg en 150 µl de tampón de ensayo durante 24-48 horas a 37 °C. El tampón de ensayo consistía en fosfato de sodio 0,2 M, pH 8,0, 0,1 mM de cada una de ferricianuro de potasio y ferrocianuro de potasio, EDTA sódica 1,0 mM, 5-bromo-4-cloro-3-indoilo-β-glucuronido 0,5 mg/ml y Tritón X-100 al 0,6 % (v/v) (Jefferson, 1987, Plant Mol. Biol. Rep. 5:387-405). La aparición de colonias de color azul se usó como el indicador de expresión del gen GUS, que indicó que la inserción de la secuencia diana era activa desde el punto de vista de la transcripción y por lo tanto accesible en el entorno genómico local.

30 Los acontecimientos transgénicos que expresaban GUS se ensayaron por PCR usando el par de cebadores P15/P16 que condujo a la amplificación de un fragmento de ADN de 10 kb que se extendía desde la UTR 3' del casete de expresión de HTP en el extremo 5' de la secuencia diana hasta la UTR 3' del casete del gen PAT parcial en el extremo 3' de la secuencia diana. Dado que todos los acontecimientos se obtuvieron con selección de higromicina, se supuso que el casete de expresión de HPT estaba intacto en todos los acontecimientos diana. Por lo tanto, únicamente la UTR 3' del casete de expresión de HPT se incluyó en el análisis PCR de longitud completa. Un subconjunto de acontecimientos también se ensayó por PCR usando los pares de cebadores P15/P17 y P18/P19 para determinar la integridad de los extremos 5' y 3' de la secuencia diana, respectivamente. Todos los acontecimientos diana confirmados con análisis PCR se ensayaron adicionalmente mediante análisis de transferencia Southern para determinar el número copias de la secuencia diana integrada.

40 Se realizó análisis de transferencia Southern en todos los acontecimientos diana que pasaron la exploración de la expresión de GUS y la PCR de longitud completa. Diez µg de ADN genómico se digirieron con NsiI, que era un cortador exclusivo dentro de la secuencia diana. El ADN genómico digerido se separó en un gel de agarosa al 0,8 % y se transfirió sobre una membrana de nailon. Después del entrecruzamiento, el ADN transferido sobre la membrana se hibridó con una sonda génica HPT para determinar el número copias del extremo 5' de la secuencia diana. Después, la misma transferencia se separó y volvió a hibridarse con una sonda del gen PAT para determinar el número de copias del extremo 3' de la secuencia diana.

50 Se seleccionaron acontecimientos múltiples que mostraron expresión de GUS y que contenían una sola copia de la secuencia diana de longitud completa para caracterización adicional, lo que incluyó un análisis de transferencia Southern más minucioso, la confirmación de toda secuencia diana y el análisis de la secuencia genómica flanqueante. Se seleccionó un acontecimiento, denominado BY2-380, basándose en la caracterización molecular. El cultivo de suspensión se restableció a partir de este acontecimiento para la retransformación posterior con vectores que comprendían ADN donante y genes de la proteína de fusión dedo de zinc no C2H2 - FokI.

55 Para garantizar el cultivo en suspensión establecido a partir del acontecimiento diana BY2-380 contenido en la secuencia diana intacta como se esperaba, la secuencia diana principal de la UTR 3' del casete de expresión HPT en el extremo 5' de la secuencia diana en la UTR 3' del casete del gen PAT parcial en el extremo 3' de la secuencia diana, se amplificó por PCR usando el par de cebadores P15/P16 y se clonó en el vector pCR2.1 TOPO (Invitrogen, Carlsbad, California). Los productos de la PCR insertados en el vector TOPO se secuenciaron por tecnología Lark, Inc. (Houston, Texas). Los resultados de la secuencia indicaron que BY2-380 tenía las secuencias diana completas

como se esperaba.

La línea celular BY2-380 se analizó adicionalmente para obtener las secuencias genómicas flanqueante usando el kit Universal GenomeWalker (Clontech, Mountain View, California). Resumiendo, 2,5 µg de ADN genómico se digirió con tres enzimas de restricción de extremo romo, EcoRV, DraI y StuI en reacciones distintas. El ADN digerido se purificó a través de extracción con fenol/cloroformo y se ligó con adaptador BD Genome Walker. Se realizó amplificación con PCR anidada con el ligamiento como molde y el cebador P20 (marchando aguas arriba del extremo 5' de la inserción de la secuencia diana) y P21 (marchando aguas abajo del extremo 3' de la inserción de la secuencia diana) para la reacción PCR primaria, y el cebador P22 (marchando aguas arriba del extremo 5' de la inserción de la secuencia diana) y P23 (marchando aguas arriba del extremo 3' de la inserción de la secuencia diana) para la reacción de PCR anidada secundaria. Los fragmentos amplificados de las reacciones de PCR secundarias se clonaron en el vector pCR2.1 TOPO o pCR Blunt II TOPO (Invitrogen, Carlsbad, California) y se secuenciaron usando un kit de secuenciación Dye Terminator Cycle (Beckman Coulter, Fullerton, CA). Las secuencias genómicas flanqueantes se obtuvieron a partir de la línea diana BY2-380 mediante este proceso. Después, se diseñaron cebadores basándose en las secuencias genómicas flanqueantes y se usaron para amplificar toda la secuencia diana.

Los fragmentos amplificados obtenidos de esta línea diana tenían el tamaño esperado. Los dos extremos de los fragmentos amplificados se confirmaron por secuenciación.

Ejemplo 8: Diseño y generación del vector de ADN donante

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

Para construir el vector donante, los 299 pb de la secuencia codificante del gen PAT parcial 5' se fusionaron con el intrón-1 de 4-CoAS de *A. thaliana* de longitud completa (Locus At3g21320, GenBank NC 003074) mediante síntesis de ADN por Picoscript Ltd., LLP, (Houston, Texas). Se añadieron sitios NotI y XhoI en el extremo 5' y 3' del fragmento, respectivamente. Este fragmento de ADN sintetizado se digirió después con NcoI/XhoI y se insertó en pDAB1575 en los mismos sitios para reemplazar la secuencia codificante del gen PAT de longitud completa y su UTR 3'. La construcción resultante se denominó pDAB1576 (FIG. 29).

Después, la construcción pDAB1576 se digirió con AgeI y todo el fragmento que contenía el casete de expresión de PAT parcial 5' flanqueado por la secuencia-1 homóloga y la secuencia-2 homóloga se insertó en pDAB2407, el vector base binario, en el mismo sitio. La construcción resultante se denominó pDAB1600 (FIG. 30) y fue la versión binaria del vector donante para la retransformación de las células de la planta.

Ejemplo 9: Diseño y generación de vectores de expresión de nucleasas con dedos de zinc

El gen de la proteína de fusión dedo zinc-FokI se condujo mediante un promotor de CsVMV y una UTR 5' (Verdaguer *et al.*, 1996, Plant Molecular Biology 31:1129-1139). En el casete de también se incluían una fase de lectura abierta-24 (orf-24) en la región no traducida (UTR) 3' de *A. tumifaciens* (Gelvin *et al.*, 1987, EP222493).

Para crear estos vectores, los controles de C2H2 y sus variantes C3H de las secuencias codificantes de IL-1-FokI y Scd27-FokI descritas en los Ejemplos 1 a 4 anteriores PCR se amplificaron por PCR a partir de sus diseños originales con los sitios BbsI o NcoI y SacI añadidos en el extremo 5' y 3' de los fragmentos PCR, respectivamente y se clonaron en el pDAB3731 (FIG. 31) digerido con NcoI-SacI. Los plásmidos resultantes se denominaron pDAB4322 (FIG. 32), pDAB4331 (FIG. 33), pDAB4332 (FIG. 34), pDAB4333 (FIG. 35) pDAB4334 (FIG. 36), pDAB4336 (FIG. 37) y pDAB4339 (FIG. 38). Todos estos vectores contenían los sitios attL1 y attL2 flanqueando el casete de expresión de ZFN y eran compatibles con el sistema de clonación Gateway™ (Invitrogen, Carlsbad, California).

Se construyeron dos conjuntos de vectores de versión binaria para la proteína de fusión IL-1-FokI. Uno contenía el gen marcador de selección PAT y el otro no contenía el gen marcador de selección PAT. Para la proteína de fusión SCd27-FokI, solamente se construyó la versión binaria del vector sin el gen marcador de selección PAT. Para crear los vectores binarios con el gen marcador de selección PAT, el casete de expresión de la proteína de fusión IL-1-FokI en pDAB4322, pDAB4331, pDAB4332, pDAB4333, pDAB4334 y pDAB4336 se clonaron en pDAB4321 (FIG. 39) a través de reacción de recombinación LR usando la mezcla enzimática LR Clonase™ (Invitrogen, Carlsbad, California). Los plásmidos resultantes se denominaron pDAB4323 (FIG. 40), pDAB4341 (FIG. 41), pDAB4342 (FIG. 42), pDAB4343 (FIG. 43), pDAB4344 (FIG. 44), pDAB4346 (FIG. 45). Para crear los vectores binarios sin el gen marcador de selección PAT, el casete de expresión C2H2 IL-1-FokI, C3H IL-1-FokI y Scd27-FokI en el pDAB4331, pDAB4336 y pDAB4339, respectivamente, se clonaron en pDAB4330 (FIG. 46) a través de reacción de recombinación LR usando la mezcla enzimática LR Clonase™ (Invitrogen, Carlsbad, California). Los plásmidos resultantes se denominaron pDAB4351 (FIG. 47), pDAB4356 (FIG. 48) y pDAB4359 (FIG. 49), respectivamente.

Para crear el control C2H2 de SCD27-FokI, el fragmento HindIII/SacI, que comprendía el promotor de CsVMV y la UTR 5' que dirige el gen PAT en pDAB7002 (FIG. 50), se reemplazó con un fragmento que comprendía el promotor de CsVMV y la UTR 5' y la UTR 5' de *N. tabacum que dirige a GUS*, que se escindió de pDAB7025 (FIG. 51) con HindIII/SacI. El plásmido resultante se denominó pDAB1591 (FIG. 52). Las secuencias codificantes DE Scd27-L0-FokI se amplificaron por PCR a partir de sus vectores originales pCDNA3.1-SCD27a-L0-FokI (FIG. 53) usando el par de cebadores P13/P14. Se añadieron los sitios BbsI y SacI en los extremos 5' y 3' de los fragmentos PCR, respectivamente. El gen PAT en pDAB1591 se reemplazó con el fragmento PCR del gen de la proteína de fusión de dedo de zinc a través de clonación CON SacI/NcoI. El plásmido resultante se denominó pDAB1594 (FIG. 54). La versión binaria de este vector se construyó escindiendo el casete de expresión génica de la proteína de fusión de dedo de zinc de pDAB1594 como una fragmento PmeI/XhoI, rellenando en los extremos y clonando en pDAB2407 en el sitio PmeI. El plásmido resultante se denominó pDAB1598 (FIG. 55). Los detalles de todos los vectores binarios usados para la transformación de plantas se resumen en la tabla 6.

Tabla 6: Vectores de expresión de nucleasas con dedos de zinc

| Vector | ZFN | Tipo de ZFP | Posición del dedo de zinc | Secuencia de aminoácidos | SEQ ID NO: |
|----------|-------------|-------------|---------------------------|--------------------------|------------|
| pDAB4323 | IL 1-FokI | C2H2 | D4 | C...C...HTKIH | 94 |
| pDAB4341 | IL 1-FokI. | C2H2 | D4 | C...C...HTKIH | 95 |
| pDAB4342 | IL 1-FokI* | C3H | D4 | C...C...HTKIC | 96 |
| pDAB4343 | IL 1-FokI* | C3H | D4 | C...C...HTKRCGGG | 97 |
| pDAB4344 | IL 1-FokI* | C3H | D4 | C...C...HAQRCG | 98 |
| pDAB4346 | IL 1-FokI* | C3H | D2 | C...C...HIRTGC | 99 |
| pDAB4351 | IL 1-FokI* | C2H2 | D4 | C...C...HTKIH | 100 |
| pDAB4356 | IL 1-FokI* | C3H | D2 | C...C...HIRTGC | 101 |
| pDAB1598 | Scd27-FokI | C2H2 | D4 | C...C...HTKIH | 102 |
| pDAB4359 | Scd27-FokI* | C3H | D4 | C...C...HAQRCGG | 103 |

*El dominio *Fok I* era un sesgo codónico de plantas.

15 Ejemplo 10: Diseño y generación del vector de control positivo

Para estimar la frecuencia de recombinación ilegítima y servir como un control positivo, se usó un vector que contenía el casete de expresión del gen PAT. Para ser comparable con los recombinantes finales, se insertó el intrón -1 de 4-CoAs de *A. thaliana* (Locus At3g21320, GenBank NC 003074) en el 299/300 pb de la secuencia codificante de PAT (Wohlleben *et al.*, 1988, Gene 70:25-37). Para crear esta construcción, el fragmento SwaI/ClaI de 2559 pb de pDAB1576 se ligó con el fragmento estructural de pDAB1577 (FIG. 56) que se digirió con las mismas enzimas de restricción. El vector resultante contenía el casete de expresión del gen PAT con los 1743 pb de la inserción del intrón-1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) en el centro de la secuencia codificante de PAT (Wohlleben *et al.*, 1988, Gene 70:25-37). Este vector se denominó pDAB1578 (FIG. 57).

Para crear el vector binario de pDAB1578, el casete de expresión del gen PAT con el intrón-1 de *A. thaliana* (Locus At3g21320, GenBank NC 003074) se escindió de pDAB1578 con PmeI/XhoI. Después de tratar el extremo romo del extremo 3' del fragmento, este se insertó en pDAB2407, el vector base binario, en el sitio PmeI. El vector resultante se denominó pDAB1601 (FIG. 58) que comprendía el gen PAT (Wohlleben *et al.*, 1988, Gene 70:25-37) que contenía la secuencia intrón-1 de 4-CoAS de *A. thaliana* (Locus At3g21320, GenBank NC 003074) dirigida por el promotor ubi 10 de *A. thaliana* (Callis, *et al.*, 1990, J. Biol. Chem. 265-12486-12493) y terminada por la orf25/26 en la UTR 3' de *A. tumefaciens* (Gelvin *et al.*, 1987, EP222493).

30 Ejemplo 11: Demostración de la recombinación homóloga intracromosómica por retransformación de cultivos célula diana con genes de nucleasas con dedos de zinc C3H

Para validar la funcionalidad de las nucleasas con dedos de zinc C3H en la estimulación de la recombinación homóloga intracromosómica, se incluyeron dos fragmentos de GFP no funcionales con secuencias solapantes de 540 pb en el vector diana como se muestra en la FIG. 59. Entre medias de estos dos fragmentos estaba un casete de expresión del gen GUS. La secuencia de unión de la proteína de fusión IL-1-FokI se fusionó con la secuencia codificante de GUS en su extremo N. Sin quedar ligado a ninguna teoría, se llegó a la hipótesis de que en presencia de la proteína de fusión IL-1-FokI, las secuencias de unión IL-1 ZFN podrían reconocerse e inducirse una rotura de ADN bicatenario, que estimularía el proceso de reparación de ADN endógeno. Sin la presencia de ADN donante, los dos fragmentos de GFP parcialmente homólogos experimentarían un proceso de recombinación homóloga intracromosómica y se reconstituiría un gen de GFP funcional.

La línea celular transgénica BY2-380 que contenía una sola copia integrada de longitud completa de la secuencia diana, se usó para reiniciar cultivos en suspensión colocando ~ 250-500 mg de tejido de callo en 40-50 ml de medio basal basado en LS que contenía higromicina 100 mg/l y subcultivando cada 7 días como se ha indicado anteriormente. Antes de la retransformación, los cultivos en suspensión se transfirieron a medio basal sin higromicina durante al menos dos pases.

La transformación mediada por *Agrobacterium* de los cultivos de células diana se realizó como se ha descrito anteriormente. Para cada experimento, se generaron 8 placas de cocultivo de la siguiente manera: una placa comprendía células cocultivadas con 300 μ l de la cepa base LBA4404 de *Agrobacterium*; una placa comprendía células cocultivadas con 300 μ l de una cepa de *Agrobacterium* que llevaba pDAB1590 (construcción de GFP funcional); seis placas cada una de ellas comprendía células cocultivadas con 300 μ l de una cepa de *Agrobacterium* que llevaba pDAB4323, pDAB4341, pDAB4342, pDAB4343, pDAB4344 y pDAB4346, respectivamente. Después del cocultivo usando los métodos descritos anteriormente, las células se sembraron en 8 placas que contenían el medio basal basado en LS complementado con carbenicilina 500 mg/l sin reactivo de selección. La expresión aparente del gen GFP funcional constituido produjo fluorescencia visible aproximadamente 5-8 días después de la transformación. El número de locus fluorescentes verdes por campo se contó observando 5 campos 'al azar' al microscopio por placa, 8 placas por construcciones en cada experimento y se calculó el promedio de 6 experimentos independientes.

Como se resume en la tabla 7, se observó un promedio de 9,50 y 7,57 locus fluorescentes verdes por campo de dos nucleasas con dedos de zinc C3H, pDAB4346 y pDAB4343, respectivamente. Estos dos diseños C3H de IL-1-FokI se comportaron mejor que sus controles C2H2, pDAB4341 (6,37 locus por campo) y pDAB4323 (5,53 locus por campo). Mientras tanto, en comparación con los controles C2H2, la función de otras dos variantes C3H de la proteína de fusión IL-1-FokI, pDAB4344 (4,39 locus por campo) y pDAB4342 (0,25 locus por campo) se deterioró de un modo significativo, en particular pDAB4342, en donde la conversión de C3H se realizó simplemente reemplazando la segunda cisteína con histidina en el cuarto dedo. No se observó fluorescencia apreciable más allá de un fondo ligero en los controles negativos transformados con la cepa base de *Agrobacterium*, LBA4404.

Tabla 7

Construcción de GFP funcional a través de recombinación homóloga intracromosómica estimulada por la proteína de fusión IL-1-FokI de dedo de zinc

| Vector | Tipo de ZFP | Expresión de GFP | Ensayo Tukey** |
|-----------|-------------|------------------|----------------|
| pDAB4346 | C3H | 9,50 | A |
| pDAB4343 | C3H | 7,57 | B |
| pDAB4341 | C2H2 | 6,37 | C |
| pDAB4323* | C2H2 | 5,53 | D |
| pDAB4344 | C3H | 4,39 | E |
| pDAB4342 | C3H | 0,25 | F |

*contiene el dominio FokI de sesgo codónico que no es de planta

**los medios no relacionados con la misma letra son significativamente diferentes al nivel de 0,05

Ejemplo 12: Demostración de recombinación homóloga intercromosómica por retransformación de cultivos de células diana con genes de nucleasas con dedos de zinc C3H y secuencias de ADN donantes

Para validar la funcionalidad de la proteína de fusión dedo de zinc C3H - FokI en la simulación de la recombinación homóloga intercromosómica en el sistema de tabaco ejemplar, se desarrollaron y ensayaron dos estrategias.

En la estrategia 1, el sitio de unión para la proteína de fusión dedo de zinc-FokI (IL-1-L0-FokI), se incluyó en el centro de la construcción diana (FIG. 61). En esta estrategia, el sitio de unión estaba flanqueado por ~3 kb de secuencias no homólogas en ambos lados seguido por la secuencia-1 homóloga (MAR RB7 de *N. tabacum*) y la secuencia-2 homóloga (intrón-1 de 4-CoAS de *A. thaliana*) aguas arriba y aguas abajo, respectivamente. Como se ha demostrado anteriormente (por ejemplo, en la publicación de patente de Estados Unidos n.º 20050064474) en presencia de la proteína de fusión dedo de zinc C2H2 IL-1-FokI, la secuencia de unión IL-1-L0-FokI se reconoció y se indujo una rotura de ADN bicatenario en este sitio específico, que estimuló el proceso de reparación del ADN endógeno. En presencia de ADN donante, que contenía secuencias homólogas idénticas a las de la secuencia diana, el gen PAT parcial 5' junto con su promotor, reemplazó todo el fragmento de ADN de ~6 kb entre las secuencias homólogas y la diana a través de recombinación homóloga. A través de este proceso, las dos secuencias del gen PAT parciales, con el intrón-1 de 4-CoAS de *A. thaliana* interpuesto entre ambas, reconstituyó un gen PAT funcional dando como resultado la expresión de PAT y un fenotipo de resistencia a herbicidas.

En la estrategia 2, se incluyeron dos sitios de unión dedo de zinc-FokI (Scd27-L0-FokI) en el vector diana: uno directamente aguas abajo de MAR RB7 de *N. tabacum* y el otro directamente aguas arriba del intrón 1 de 4-CoAS de

A. thaliana (FIG. 62). Entre los dos sitios de unión de la proteína de fusión dedo de zinc - FokI había ~6 kb de secuencia, que incluyó el fragmento GFP 5', un casete de expresión GUS y el fragmento GFP 3'. Como se ha demostrado anteriormente (por ejemplo, publicación de patente de Estados Unidos n.º 20050064474), en presencia de la proteína de fusión Scd27 dedo zinc-FokI, las dos secuencias de unión reconocidas y las roturas de ADN bicatenario se indujeron en ambas localizaciones, lo que eliminó el fragmento de ADN de ~6 kb entre estas dos secuencias de unión, y estimuló el proceso de reparación de ADN endógeno. Similar a la estrategia 1, en presencia de ADN donante, que contenía secuencias homólogas idénticas a las de la secuencia diana, el gen PAT parcial 5' junto con su promotor, se insertó en la secuencia diana a través de recombinación homóloga en el sitio en donde se indujeron las roturas de ADN bicatenario. A través de este proceso, las dos secuencias del gen PAT parciales, con el intrón-1 de 4-CoAS de *A. thaliana* interpuesto entre ambas, reconstituyó un gen PAT funcional, dando como resultado la expresión de PAT y un fenotipo de resistencia a herbicidas.

La transformación mediada por *Agrobacterium* del cultivo de células *diana* BY2-380 se realizó como se ha descrito anteriormente. Para cada experimento, se generaron 12 placas de cocultivo de la siguiente manera: una placa comprendía células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1600 (ADN donante) y 250 µl de la cepa base de *Agrobacterium*, LBA4404; una placa comprendía células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1601 (marcador de selección PAT) y 250 µl de la cepa base de *Agrobacterium*, LBA4404; dos placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1600 (ADN donante) y 250 µl de una cepa de *Agrobacterium* que llevaba pDAB4351 (C2H2 IL-1 ZFP-FokI); tres placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1600 (ADN donante) y 250 µl de una cepa de *Agrobacterium* que llevaba pDAB4356 (C3H IL-1 ZFP-FokI); dos placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1600 1600 (ADN donante) y 250 µl de una cepa de *Agrobacterium* que llevaba pDAB1598 (C2H2 Scd 27a ZFP-FokI); tres placas comprendían células cocultivadas con 50 µl de una cepa de *Agrobacterium* que llevaba pDAB1600 (ADN donante) y 250 µl de una cepa de *Agrobacterium* que llevaba pDAB4359 (C3H Scd27aZFP-FokI). Después de cocultivo usando los métodos descritos anteriormente, las células se sembraron en placas en el medio basal basado en LS que contenía carbenicilina 500 mg/l y Bialaphos® 15 mg/l. Los aislados individuales resistentes a Bialaphos® aparecieron 2-4 semanas después de la siembra en placa y se transfirieron a placas individuales de 60 x 20 mm (un aislado por placa) donde se mantuvieron como callos en un programa de subcultivo de 14 días hasta necesitarse para análisis.

Los aislados múltiples resistentes a Bialaphos® se obtuvieron de la nucleasa con dedo de zinc IL-1 C3H (pDAB4356) y de la nucleasa con dedo de zinc C3H Scd27 (pDAB4359). Estos aislados se analizaron por PCR usando el par de cebadores P24/25, que amplificó un fragmento de ADN que abarcaba el gen PAT reconstituido. El cebador P24 era homólogo al extremo 5' de la secuencia codificante PAT en el ADN donante y el cebador P25 era homólogo al extremo 3' de la secuencia codificante PAT en el ADN diana. Si las dos secuencias codificantes PAT parciales se uniesen a través de recombinación homóloga podría obtenerse un fragmento de PCR de 2,3 kb. Como se muestra en la FIG. 63, se obtuvo un producto de PCR de 2,3 kb a partir de aislados múltiples analizados. Estos aislados se obtuvieron tanto de la cotransformación del gen de la proteína de fusión C3H IL-1 dedo de zinc-FokI/ADN donante y como del gen de la proteína de fusión C3H Scd27 dedo de zinc-FokI/ADN donante. Los productos de PCR de 2,3 kb de aislados independientes múltiples representativos de los derivados de ambas transformaciones del gen de la proteína de fusión C3H IL-1 dedo de zinc-FokI y C3H Scd27 dedo de zinc-FokI se purificaron de geles de agarosa y se clonaron en el vector pCR2.1 TOPO (Invitrogen, Carlsbad, California). El producto de PCR de 2,3 kb insertado en el vector TOPO se secuenció después usando el kit de secuenciación Dye Terminator Cycle (Beckman Coulter). Los resultados de la secuenciación confirmaron que todos los productos de PCR clonados en el vector TOPO contenían la secuencia recombinada como se esperaba, incluyendo las secuencias del gen PAT parciales 5' y 3' con el intrón-1 de 4-CoAS de *A. thaliana* interviniente. Estos resultados confirmaron la recombinación intercromosómica esperada para ambas estrategias ensayadas e ilustran el direccionamiento génico mediante la expresión de genes de la proteína de fusión C3H dedo de zinc-FokI.

Ejemplo 13: Identificación de secuencias de genes diana en cultivo de células de maíz

A. Identificación de secuencias

En este ejemplo, se seleccionaron secuencias de ADN para un gen de maíz endógeno de función conocida como dianas para la edición genómica usando nucleasas con dedos de zinc modificadas genéticamente. La estructura y la secuencia genómica de este gen, denominado IPP2-K, que procede de una línea endogámica de maíz patentada 5XH751, se había descrito en el documento WO2006/029296; cuya descripción se incorpora por referencia.

En particular, se usó la secuencia genómica de IPP2-K para consultar la base de datos del genoma de maíz del TIGR (*The Institute for Genomic Research*) (disponible en internet en la <http://www.tigr.org/tdb/tgi/maize/>) usando algoritmos BLAST. Se identificaron diversos fragmentos genómicos adicionales con segmentos de homología solapante con IPP2-K, incluyendo, pero sin limitación, los accesos AZM515213 y TC311535. Basándose en la secuencia de estos accesos, así como en la secuencia de IPP2-K, se diseñaron oligonucleótidos cortos múltiples para su uso como cebadores de la PCR usando el programa Primer3 (Rozen, S. y Skaletsky, H.J. (2000) *Primer3* en la WWW para usuarios en general y para programadores biólogos. En: Krawetz S, Misener S (eds.) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, págs. 365-386; también

disponible en internet). Estos cebadores incluyen, pero sin limitación, los siguientes oligonucleótidos en orientación directa:

1. 5'-ATGGAGATGGATGGGGTTCTGCAAGCCGC-3' (SEQ ID NO: 104)
2. 5'-CTTGCAAGGTACTGCGGCTCAAGAAGATTC-3' (SEQ ID NO: 161)
- 5 3. 5'-ATGAAGAAAAGACAGGGAATGAAGGAC-3' (SEQ ID NO: 162)
4. 5'-ATGAAGAAAAGACAGGGAATGAAGGACCGCCAC-3' (SEQ ID NO: 163)
5. 5'-CATGGAGGGCGACGAGCCGGTGTAGCTG-3' (SEQ ID NO: 164)
6. 5'-ATCGACATGATTGGCACCCAGGTGTTG-3' (SEQ ID NO: 165)

Además, los cebadores incluyen, pero sin limitación, los siguientes oligonucleótidos en orientación inversa:

- 10 7. 5'-TTTCGACAAGCTCCAGAAAATCCCTAGAAAC-3' (SEQ ID NO: 166),
8. 5'-ACAAGCTCCAGAAAATCCCTAGAAACAC-3' (SEQ ID NO: 167)
9. 5'-TTCGACAAGCTCCAGAAAATCCCTAGAAACAC-3' (SEQ ID NO: 168)
10. 5'-TGCTAAGAACATTCTTTTCGACAAGCTCC-3' (SEQ ID NO: 169)
11. 5'-GAACATTCTTTTCGACAAGCTCCAGA.AA.ATCC-3' (SEQ ID NO: 170)

- 15 Todos los cebadores oligonucleotídicos se sintetizaron y adquirieron en Integrated DNA Technologies (IDT, Coralville, IA).

B. Cultivo de células de maíz Hi II

- 20 Para obtener embriones inmaduros para el inicio de cultivos de callos, se realizaron cruces F₁ entre padres A y B Hi-II cultivados en invernadero (Armstrong, C., Green, C. y Phillips, R. (1991) Maize Genet. Coop. News Lett. 65: 92-93). Se recogieron embriones de aproximadamente 1,0-1,2 mm de tamaño (~ 9-10 días después de la polinización), de mazorcas sanas y se esterilizó la superficie limpiando con jabón Liqui-Nox®, sumergido en etanol a 70 % durante 2-3 minutos, después sumergido en lejía comercial al 20 % (hipoclorito de sodio al 0,1 %) durante 30 minutos.

- 25 Las mazorcas se aclararon en agua destilada, estéril y los embriones cigóticos inmaduros se cortaron asépticamente y se cultivaron en medio 15Ag10 (medio N6 (Chu C.C., Wang C.C., Sun C.S., Hsu C., Yin K.C., Chu C.Y., y Bi F.Y. (1975) Sci. Sinica 18:659-668), 2,4-D 1,0 mg/l, sacarosa 20 g/l, hidrolizado de caseína 100 mg/l (digestión enzimática), L-prolina 25 mM, AgNO₃ 10 mg/l, Gelrite 2,5 g/l, pH 5,8) durante 2-3 semanas con el escutelo orientado lejos del medio. Los tejidos que mostraban la morfología esperada (Welter, ME, Clayton, DS, Miller, MA, Petolino, JF. (1995) Plant Cell Rep: 14:725-729) se transfirieron selectivamente, a intervalos quincenales, a medio reciente 15Ag10 durante aproximadamente 6 semanas, después se transfirieron a medio 4 (medio N6, 2,4-D 1,0 mg/l, sacarosa 20 g/l, hidrolizado de caseína 1,00 mg/l (digestión enzimática), L-prolina 6 mM, Gelrite 2,5 g/l, pH 5,8) a intervalos quincenales durante aproximadamente 2 meses.
- 30

- 35 Para iniciar los cultivos en suspensión embriogénicos, aproximadamente un volumen celular comprimido (VCC) de 3 ml de tejido de callo, originado de un solo embrión, se añadieron a aproximadamente 30 ml de medio H9CP+ líquido (mezcla salina basal MS (Murashige T., & Skoog F. (1962) Physiol. Plant. 15:473-497), vitaminas MS modificadas que contenían 10 veces menos ácido nicotínico y 5 veces más tiamina-HCl, 2,4-D 2,0 mg/l, ácido α -naftalenacético (NAA) 2,0 mg/l, sacarosa 30 g/l, hidrolizado de caseína 200 mg/l (digestión ácida), mioinositol 100 mg/l, L-prolina 6 mM, agua de coco al 5 % v/v (añadida justo antes de subcultivo), pH 6,0). Los cultivos en suspensión se mantuvieron en la oscuridad en matraces Erlenmeyer de 125 ml en un agitador con temperatura controlada configurado a 125 rpm a 28 °C. Durante el establecimiento de la línea celular (2-3 meses), las suspensiones se subcultivaron cada 3,5 días añadiendo un VCC de 3 ml de células y 7 ml de medio acondicionado a 20 ml de medio reciente H9CP+ líquido usando una pipeta de calibre ancho. Después alcanzar la madurez, según el tiempo de generación del crecimiento, las suspensiones se mejoraron y se mantuvieron en matraces de 500 ml a través de los cuales 12 ml de un VCC de células y 28 ml de medio acondicionado se transfirieron a medio H9CP+ 80 ml. Después de establecimiento completo del cultivo en suspensión, las alícuotas se crioconservaron para su uso futuro. Véase, el documento WO 2005/107437.
- 40
- 45

C. Aislamiento y amplificación de ADN

- Los cultivos de células de maíz Hill, como se ha descrito anteriormente, se desarrollaron en matraces de 250 ml en medio GN6 convencional (medio N6, 2,4-D 2,0 mg/l, sacarosa 30 g/l, Gelrite 2,5 g/l, pH 5,8) y se extrajo ADN genómico usando el kit de extracción de Qiagen (Valencia, CA) Plant DNeasy siguiendo las recomendaciones del fabricante. Las reacciones de amplificación por PCR, usando los cebadores descritos anteriormente en todas las posibles combinaciones, se realizaron con las siguientes condiciones: volumen de reacción 25 μ l que contenía molde de ADNg 20 ng, 20 pmol de cada cebador, DMSO 1 % y 10 unidades de Pf polimerasa Accuprime™ (Invitrogen, Carlsbad, CA) en el tampón enzimático del fabricante. Los productos de amplificación, cuyo tamaño variaba de 500 pb a 2 kb, resultaron de los ciclos de amplificación que consistían en 95 °C-1', (95°C-30", 57-62 °C-30", 72 °C-1') X 30, 72 °C-5', mantenido a 4 °C. Los fragmentos amplificados se clonaron directamente en el vector pCR2.1 (Invitrogen, Carlsbad, CA) usando el kit de clonación TA de Invitrogen (Carlsbad, CA) siguiendo las recomendaciones del fabricante.
- 50
- 55

D. Análisis de secuencia

5 Análisis previos del gen IPP2-K en cultivo de células 5XH751 endogámicas y Hill de maíz, habían indicado la presencia de 2-3 genes distintos que comprendían una familia de genes pequeña (Sun *et al.*, in press, Plant Physiology; documento WO2006029296). Por lo tanto, los fragmentos clonados aislados se secuenciaron con el kit de secuenciación CEQ Dye Terminator Cycle de Beckman Coulter (Fullerton, CA) siguiendo las recomendaciones del fabricante. En análisis de secuencia de clones múltiples reveló que 2 fragmentos génicos distintos, procedentes de 2 locus distintos y previamente caracterizados del genoma del maíz, se habían aislado de células Hill.

10 La comparación de las 2 secuencias aisladas de células Hill cultivas indicó que, en las regiones codificantes previstas, existían pequeñas diferencias, tales como polimorfismos mononucleotídicos (SNP, acrónimo de las siglas inglesas *single nucleotide polymorphisms*) entre los 2 parálogos, mientras que las regiones intrónicas y no codificantes variaban significativamente a nivel nucleotídico. Estas diferencias entre los 2 parálogos se observó porque presentaban regiones de secuencia que podían discriminarse mediante una proteína de unión a ADN dependiente de secuencia, tal como un dominio con dedos de zinc. Un experto en la técnica puede diseñar dominios de unión a ADN con dedos de zinc que se unen a una secuencia génica y no a otra secuencia génica muy similar. La secuencia génica parcial de 1,2 kb que corresponde al parólogo de interés (FIG. 66) se seleccionó como molde para el diseño de una proteína nucleasa con dedos de zinc y posteriormente se sometió a análisis de dominio de unión a ADN con dedos zinc descrito anteriormente.

Ejemplo 14: Diseño de dominios de unión a ADN con dedos de zinc de IPP2-K

20 Usando sitios diana identificados para IPP2-K, se seleccionaron hélices de reconocimiento para dedos de zinc de IPP2-K. Los diseños de dedos de zinc se muestran a continuación en la tabla 8:

Tabla 8: Diseños de dedos de zinc de IPP2-K

| Nombre de ZFN | D1 | D2 | D3 | D4 | D5 | D6 |
|----------------|-------------------------|-------------------------|--------------------------|-------------------------|-------------------------|-------------------------|
| IPP2-K-1072a1 | DRSALSR (SEQ ID NO:105) | RNDDRKK (SEQ ID NO:106) | RSDNLST (SEQ ID NO:107) | HSHARIK (SEQ ID NO:108) | RSDVLSE (SEQ ID NO:109) | QSGNLAR (SEQ ID NO:110) |
| IPP2-K-1072b1 | DRSALSR (SEQ ID NO:105) | RNDDRKK (SEQ ID NO:106) | RSDNLAR (SEQ ID NO:111) | TSGSLTR (SEQ ID NO:112) | RSDVLSE (SEQ ID NO:109) | QSGNLAR (SEQ ID NO:110) |
| IPP2-K-1072c11 | DRSALSR (SEQ ID NO:105) | RNDDRKK (SEQ ID NO:106) | TSGNLTR (SEQ ID NO:113) | TSGSLTR (SEQ ID NO:112) | RSDVLSE (SEQ ID NO:109) | QSGNLAR (SEQ ID NO:110) |
| IPP2-K-r1065a1 | RSDHLSE (SEQ ID NO:114) | QSATRKK (SEQ ID NO:115) | ERGLTLAR (SEQ ID NO:116) | RSDALTQ (SEQ ID NO:117) | NINGUNO | NINGUNO |
| IPP2-K-r1149a2 | RSDSLSA (SEQ ID NO:118) | RSAALAR (SEQ ID NO:119) | RSDNLSE (SEQ ID NO:120) | ASKTRTN (SEQ ID NO:121) | DRSHLAR (SEQ ID NO:122) | NINGUNO |
| IPP2-K-1156a2 | RSDHLST (SEQ ID NO:123) | QSGSLTR (SEQ ID NO:124) | RSDHLSE (SEQ ID NO:114) | QNHHRIN (SEQ ID NO:125) | TGSNLTR (SEQ ID NO:126) | DRSALAR (SEQ ID NO:127) |

Los sitios diana de los diseños de dedos de zinc se muestran a continuación en la tabla 9:

Tabla 9: Sitios diana de dedos de zinc de IPP2-K

| Nombre de ZFN | Sitio diana (5' a 3') |
|----------------|------------------------------------|
| IPP2-K-1072a1 | GAAGTGGTTGAGTCGGTC (SEQ ID NO:128) |
| IPP2-K-1072b1 | GAAGTGGTTGAGTCGGTC (SEQ ID NO:129) |
| IPP2-K-1072c1 | GAAGTGGTTGAGTCGGTC (SEQ ID NO:129) |
| IPP2-K-r1065a1 | ATGGCCCCACAG (SEQ ID NO: 130) |
| IPP2-K-r1149a2 | GGCACCCAGGTGTTG (SEQ ID NO: 131) |
| IPP2-K-1156a2 | GTCGATGGTGGGGTATGG (SEQ ID NO:132) |

Los diseños de IPP2-K se incorporaron en vectores de expresión de dedos de zinc que codifican una proteína que tiene una estructura CCHC. Véanse, las tablas 1 a 4 anteriores. Después, las secuencias codificantes de dedos de zinc no canónicos de fusionaron con el dominio nucleasa de la enzima de restricción de tipo IIS *FokI* (aminoácidos 384-579 de la secuencia de Wah *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:10564-10569 mediante un enlazador ZC de cuatro aminoácidos) para formar las ZFN de IPP2-K.

Ejemplo 15: Corrección génica usando nucleasas con dedos de zinc de IPP2-K

La capacidad de las ZFN de IPP2-K como se describe en la presente memoria para facilitar la recombinación homóloga se ensayó en el sistema GFP descrito en Umov (2005) Nature 435(7042):646-51 y en la publicación de patente de Estados Unidos n.º 20050064474 (por ejemplo, Ejemplos 6-11). En resumen, 50 ng de cada ZFN y 500 ng del donante de GFP menos el promotor (Umov (2005) Nature) se transfectoron en 500.000 células indicadoras, usando 2 ul de Lipofectamine 2000 por muestra, siguiendo el protocolo de Invitrogen Lipofectamine 2000.

Se añadió vinblastina 24 horas después de la transfección a una concentración final de 0,2 uM, y se retiró 72 horas después de la transfección.

Las células se ensayaron con respecto a la expresión de GFP 5 días después de la transfección midiendo 40.000 células por transfección en el el analizador FACS de sobremesa de Guava. Los resultados se muestran en la Figura 69.

Ejemplo 16: Expresión de ZFN C3H1 en células de maíz Hill

A. Diseño de vectores

Se construyeron vectores plasmídicos para la expresión de proteínas ZFN en células de maíz. Para minimizar la expresión y la estequiometría relativa de las 2 proteínas distintas necesarias para formar un heterodímero de nucleasa con dedo de zinc funcional, se adoptó una estrategia de expresión que daba como resultado la inserción de las fases de lectura abierta de ambos monómeros ZFN en un solo vector, conducido por un solo promotor. Esta estrategia aprovecha la funcionalidad de una secuencia 2A (Mattion, N.M., Harnish, E.C., Crowley, J.C. & Reilly, P.A. (1996) J. Virol. 70, 8124-8127) procedente del virus *Thesoa assigna*, una señal de localización nuclear (NLS, acrónimo de las siglas en inglés de *nuclear localitation signal*) del maíz del gen *opaco-2* (*op-2*) (Maddaloni, M., Di Fonzo, N., Hartings, H., Lazzaroni, N., Salamini, F., Thompson, R., & Motto M. (1989) Nucleic Acids Research Vol. 17(18):7532), y un promotor derivado del gen *ubiquitina-1* del maíz (Christensen A.H., Sharrock R.A., & Quail P.H. (1992) Plant Mol Biol. 18(4):675-89). Se diseñó un esquema de clonación modular gradual para desarrollar estos vectores de expresión para cualquier par dado de genes que codifican ZFN seleccionados del archivo de la biblioteca o sintetizados de nuevo.

En primer lugar, un vector pVAX (véase, por ejemplo, la publicación de patente de Estados Unidos 2005-0267061; cuya descripción se incorpora por referencia) se modificó para que incluyera el dominio de expresión N terminal como se muestra en la FIG. 65, paneles A a E. Las características de este plásmido modificado (pVAX-N2A-NLSop2-EGFP-FokMono) (FIG. 65A) incluyen un segmento rediseñado y sintetizado que codifica una NLS procedente del gen *op-2* del maíz (RICRKESNRESARRSRYRK, SEQ ID NO: 133), y un segmento rediseñado y sintetizado que codifica el dominio de nucleasa FokI utilizando el sesgo codónico del maíz. Adicionalmente, una inserción mononucleotídica (C) aguas abajo del sitio único XhoI creó un sitio SacI extra por comodidad en la clonación.

En segundo lugar, también se modificó un vector pVAX (véase, por ejemplo publicación de patente de Estados Unidos 2005-0267061), que incluía el dominio de expresión C terminal: Las características de este plásmido modificado (pVAX-C2A-NLSop2-EGFP-FokMono) (FIG. 65B) incluyen un segmento rediseñado y sintetizado que codifica una NLS procedente del gen *op-2* del maíz (RKRKESNRESARRSRYRK, SEQ ID NO: 133), y un segmento rediseñado y sintetizado que codifica el dominio nucleasa FokI utilizando el sesgo codónico del maíz. Adicionalmente, la secuencia 2A del virus *Thesoa assigna* (EGRGSLTTCGDVEENPGP, SEQ ID NO: 134) se introdujo en el extremo N de la ORF de ZFN con el fin de ligar posteriormente los 2 dominios que codifican la proteína.

Los casetes génicos que codifican las ORF de proteínas individuales de dedos de zinc se clonaron en el vector N2A o C2A por ligamiento usando las enzimas de restricción KpnI y BamHI para crear extremos compatibles. A continuación, el fragmento BglII/XhoI del vector C2A se insertó en el vector N2A mediante los mismos sitios de restricción, produciendo una construcción intermedia que contenía un casete que incluía los 2 dominios que codifican la ZFN flanqueados por los sitios de restricción NcoI y SacI.

Finalmente, el casete NcoI/SacI a partir de esta construcción intermedia (FIG. 65C), que contenía ambos genes de ZFN, se escindió por restricción usando aquellas enzimas y se ligó en el esqueleto del plásmido pDAB3872 (FIG. 65D). Los plásmidos resultantes incluían los genes de ZFN más el promotor en cuestión y las secuencias terminadoras, más marcadores de selección para el mantenimiento del plásmido.

En las construcciones finales, un ejemplo de lo cual se muestra en la FIG. 65E, el casete de expresión ZFN (que incluye elementos promotores y terminadores) está flanqueado por sitios attL para una manipulación conveniente usando el sistema Gateway de Invitrogen (Carlsbad, CA). Cada una de las construcciones ZFN generadas usando este esquema de clonación, se transformó en células DH5a de *E. coli* (Invitrogen, Carlsbad, CA) y posteriormente se mantuvo en la selección apropiada.

B. Suministro de ADN y expresión transitoria

Se generaron preparaciones de plásmidos de los vectores de expresión ZFN construidos como se define en la FIG. 65E, a partir de cultivos de 2 l de células de *E. coli* que se cultivaron en medio LB más antibióticos usando un kit Gigapred sin endonucleasa de Qiagen (Valencia, CA) siguiendo las recomendaciones del fabricante. El ADN plasmídico se suministró directamente a células de cultivo Hill de maíz usando diversos métodos.

En un ejemplo, las células de maíz se sometieron a suministro de ADN mediante Whiskers™. Aproximadamente 24 horas antes del suministro de ADN, 3 ml VCC de células en suspensión de maíz Hill más 7 ml de medio acondicionado se subcultivaron en 20 ml de medio GN6 líquido (medio GN6 sin Gelrite) en un matraz Erlenmeyer de 125 ml y se colocó en un agitador a 125 rpm a 28 °C durante 24 horas. Se retiraron 2 ml VCC y se añadieron a 12 ml de medio osmótico GN6 S/M (medio N6, 2,4-D 2,0 mg/l, sacarosa 30 g/l, sorbitol 45,5 g/l, manitol 45,5 g/l, mioinositol 100 mg/l, pH 6,0) en un matraz Erlenmeyer de 125 ml. El matraz se incubó en la oscuridad durante 30-35 minutos a 28 °C con agitación moderada (125 rpm). Durante este tiempo, se preparó una suspensión de 50 mg/ml de whiskers de carburo de silicio (Advanced Composite Materials, Inc., Eureka Springs, AK) añadiendo el volumen apropiado de medio líquido GN6 S/M a whiskers (fibras cortas) estériles previamente pesados. Después de la incubación en GN6 S/M, el contenido de cada matraz se vertió en un tubo de centrifuga cónico de 15 ml.

Después de la sedimentación de las células, se extrajo todo el medio GN6 S/M líquido menos 1 ml y se recogió en el matraz de 125 ml para su uso futuro. La suspensión de whiskers, previamente hidratada, se agitó vorticialmente durante 60 segundos a velocidad máxima, se añadieron 160 µl al tubo de centrifuga usando una punta de pipeta de filtro de calibre ancho y se añadieron 20 µg de ADN. El tubo se 'agitó vorticialmente con los dedos' y se colocó inmediatamente en un amalgamador deltaul Caulk 'Vari-Mix II', modificado para sujetar un tubo de cultivo de 17 x 100 mm, y después se agitó durante 60 segundos a velocidad media. Después de la agitación, el cóctel de células, los medios, los whiskers y el ADN volvieron a llevarse al matraz Erlenmeyer junto con 18 ml de medio GN6 líquido adicional. Se dejó que las células se recuperasen en un agitador a 125 RPM durante 2 horas a 28 °C en la oscuridad.

Aproximadamente 5-6 ml de la suspensión dispersada se filtraron sobre un papel de filtro Whatman n.º 4 (5,5 cm) usando una unidad recogedora de células de vidrio conectada a una línea de vacío interna de tal manera que se obtuvieron 5-6 filtros por muestra. Los filtros se colocaron sobre placas de 60 x 20 mm de medio GN6 y se cultivaron a 28 °C en la oscuridad. Después de 24, 48 o 72 horas, las células de los 2-5 papeles de filtro se rasparon, se recogieron en un tubo, se colocaron en hielo seco y después se congelaron a -80 °C.

En otro ejemplo de suministro de ADN, las preparaciones de plásmido sin endonucleasa purificadas se suministraron directamente a células de maíz usando técnicas de bombardeo con microproyectil adaptadas del protocolo del instrumento del fabricante. Todos los bombardeos se realizaron con el sistema Biolistic PDS-1000/He™ (Bio-Rad Laboratories, Hercules, CA). Para el recubrimiento de las partículas, 3 mg de partículas de oro de 1,0 micras de diámetro se lavaron una vez con etanol al 100 %, dos veces con agua destilada estéril y se resuspendieron en 50 µl de agua en un tubo de Eppendorf siliconizado. Se añadieron cinco microgramos de ADN plasmídico, espermidina 20 µl (0,1 M) y cloruro de calcio 50 µl (2,5 M) a la suspensión de oro. La mezcla se incubó a temperatura ambiente durante 10 minutos, se sedimentó a 10K rpm durante 10 s, se resuspendió en 60 µl de etanol al 100 % frío y se distribuyeron 8-9 µl sobre cada microtransportador. Para preparar las células para el bombardeo, grupos de células se retiraron del cultivo líquido 3 días después del cultivo y se colocaron en un círculo de 2,5 cm de diámetro de medio osmótico que consistía en medio de crecimiento más 0,256 M de cada uno de manitol y sorbitol en una placa de Petri. Las células se incubaron en medio osmótico durante 4 horas antes del bombardeo. El bombardeo tuvo lugar en el instrumento descrito anteriormente, que se utiliza colocando el tejido en el anaquel central al vacío a 1100 psi y 27 pulgadas de Hg y siguiendo el manual operativo. En un intervalo de 24 horas posterior al tratamiento, los grupos de células bombardeadas se recogieron, se congelaron en N₂ líquido y se conservaron a -80 °C.

Otro ejemplo de suministro de ADN y de expresión transitoria de las ZFN en las células de maíz implicó la utilización de preparaciones de protoplastos. Usando métodos modificados de Mitchell y Petolino (1991) J. Plant. Physiol. 137: 530-536 y Lyznik *et al.* (1995) Plant J. 8(2): 177-186, se prepararon protoplastos de cultivo de células de maíz Hill. Los cultivos en suspensión se recogieron 48 horas después del cultivo (crecimiento semilogarítmico) por centrifugación a 1000 rpm durante 5 minutos. El medio de cultivo se retiró y 5 ml de VCC concentrado se lavaron suavemente en 10 ml de medio W5 (NaCl₂ 154 mM; CaCl₂H₂O 125 mM; KCl₂ 5 mM; glucosa 5 mM; pH 5,8).

Las células lavadas se recogieron por centrifugación a 100 rpm durante 5 minutos y posteriormente se incubaron en un cóctel enzimático que contenía Cellulase Y-C 3 % + pectoliasa Y23 0,3 % (Karlan Research Products Corp., Cottonwood, AZ) en 25 ml de medio K3 esterilizado por filtración (KNO₃ 2,5 g; NH₄NO₃ 250 mg; CaCl₂ 900 mg (dihidrato); Mg₂SO₄ 250 mg; NH₄SO₄ 250 mg; NaPO₄ 150 mg (monobásico); xilosa 250 mg; reserva de sulfato

ferroso/quelato 10 ml (F318); micronutriente B5 1 ml (reserva 1000X - 750 mg de yoduro potásico; ácido molíbdico deshidratado 250 mg (sal sódica); cloruro de cobalto 25 mg; sulfato de cobre 25 mg); vitaminas K3 10 ml (reserva 100X - mioinositol 1 g; piridoxina HCl 10 mg; tiamina HCl 100 mg; ácido nicotínico 10 mg);+ manitol 0,6 M; pH = 5,8]. Las células se incubaron a 25 °C durante 5-6 horas con agitación suave (50 rpm) para digerir la pared celular secundaria de la planta.

Después de la degradación de la pared celular, la mezcla de célula-enzima se filtró a través de un cedazo para células de 100 micras y el flujo de paso, que contenía protoplastos y los desechos celulares, se lavaron con el mismo volumen de medio K3 + manitol 0,6 M. Los protoplastos se centrifugaron a 800 rpm durante 5 minutos, el sobrenadante se desechó y se repitió el lavado. El sedimento de protoplastos se lavó resuspendido en solución de K3 20 ml + manitol 0,6 M + Ficoll 400 al 9 %. Diez ml de esta solución se dispensaron en 2 tubos de plástico estériles y 2 ml de medio TM (MES 19,52 g; manitol 36,45 g; reserva de CaCl₂·H₂O 2 M 40 ml; pH=5,5) se extendieron suavemente sobre la suspensión, formando un gradiente discontinuo.

Los protoplastos viables se separaron de los no viables, de los desechos celulares y de las células intactas en suspensión mediante centrifugación a 800 rpm durante 5 minutos. La banda diferenciada de protoplastos formada en la interfaz del gradiente se retiró con una pipeta y se lavó solución TM reciente 10 ml, seguido de centrifugación a 800 rpm durante 5 minutos. El sedimento de protoplastos resultante se resuspendió en 1 ml de medio TM y se cuantificó el número de protoplastos viables con tinción de diacetato de fluoresceína (FDA) 25 mg/mg en un hemocitómetro. La solución de protoplastos se ajustó a una concentración final de 1×10^7 protoplastos/ml en medio TM.

Aproximadamente 1×10^6 protoplastos (100 μ l) se transfirieron a un tubo Eppendorf de 2 ml que contenía ADN plasmídico purificado 10-80 μ g. Gradualmente, se añadieron 100 μ l de una suspensión de PEG-3350 al 40 % (Sigma Chemical Co., St. Louis, MO) y la suspensión se mezcló suavemente. La mezcla de protoplastos/ADN se incubó durante 30 minutos a temperatura ambiente, seguido de una dilución por goteo con medio de crecimiento GN6 1 ml. Los protoplastos diluidos se incubaron en este medio durante 24 horas a 25 °C y posteriormente se recogieron, se congelaron en N₂ líquido y se conservaron a -80°C.

Ejemplo 17: Funcionalidad de ZFN *in vivo*

La funcionalidad de una ZFN en este ejemplo se entiende que incluye (pero sin limitación) la capacidad de una ZFN para expresarse en células de una especie de cultivo, y para que la ZFN actúe como mediadora en una rotura bicatenaria en el genoma endógeno de ese cultivo a través del reconocimiento de, unión a y escisión de, su diana deseada. También se entiende que, en este ejemplo, la diana de la ZFN es un gen en un locus endógeno y la conformación dentro del genoma del cultivo.

Para evaluar si las ZFN modificadas genéticamente tenían funcionalidad contra el gen diana predicho en un contexto genómico, se desarrollaron ensayos basados en secuencias de ADN. Se predijeron roturas de ADN bicatenario inducidas por ZFN que inducían mecanismos de preparación tales como unión de extremos no homólogos (NHEJ) (revisado por Cahill *et al.*, (2006) *Mechanisms Front Biosci.* 1(11): 1958-76). Un resultado de NHEJ es que una proporción de las cadenas de ADN rotas se repararán de una manera imperfecta dando como resultado pequeñas deleciones, inserciones o sustituciones en el sitio de escisión. Un experto en la técnica puede detectar estos cambios en la secuencia del ADN a través de diversos métodos.

A. Secuenciación y clonación basada en PCR

En un ejemplo, células cultivadas Hill de maíz que expresan proteínas ZFN se aislaron a las 24 horas postransformación, se congelaron y sometieron a extracción de ADN genómico usando el kit de extracción de Qiagen (Valencia, CA) Plant DNeasy siguiendo las recomendaciones del fabricante. La amplificación por PCR se realizó usando cebadores oligonucleotídicos específicos para el gen diana y flanqueando el sitio de escisión predicho de la ZFN. Se usó un cebador PCR de orientación directa (5'-GGAAGCATTATTCCAATTTGATGATAATGG-3') (SEQ ID NO:135), y un cebador PCR de orientación inversa (5'-CCCAAGTGTCGAGGTTGTCAATATGTTAC-3') (SEQ ID NO:136) específico para el parólogo del gen IPP2-K diana en combinación para amplificar el ADN genómico purificado en las siguientes condiciones: volumen de reacción 25 μ l que contenía molde de ADNg 20 ng, 20 pmol de cada cebador, DMSO al 1 % y 10 unidades de Pf polimerasa Accuprime (Invitrogen, Carlsbad, CA) en el tampón enzimático del fabricante. Los productos de amplificación del tamaño esperado resultaron de ciclos tipos de amplificación que consistían en 95°C-1', (95°C-30", 61°C-30", 72°C-1') X 30, 72°C-5', mantenido a 4 °C.

Los fragmentos amplificados se clonaron directamente en el vector pCR2.1 (Invitrogen, Carlsbad, CA) usando el kit de clonación TA Invitrogen (Carlsbad, CA). Los fragmentos clonados aislados se secuenciaron con el kit de secuenciación CEQ Dye Terminator Cycle de Beckman Coulter (Fullerton, CA) siguiendo las recomendaciones del fabricante en un formato de 96 pocillos. En este experimento, se predecía que las proteínas ZFN se unían a 2 secuencias específicas cortas del gen IPP2-K para crear una nucleasa heterodimérica que escinde el ADNbc como se muestra en la FIG. 66.

El análisis de los resultados de secuenciación de clones múltiples reveló que el clon n.º 127 contenía una pequeña delección exactamente en el sitio de escisión predicho de la ZFN, indicando que el mecanismo de unión de extremos no homólogos, NHEJ, había actuado como mediador en una reparación imperfecta de la secuencia de ADN en ese sitio (FIG. 67).

- 5 Estos resultados demuestran la capacidad de estas ZFN modificadas genéticamente para inducir roturas bicatenarias diana, de una manera específica en un locus génico endógeno dentro de especies de cultivo.

B. Análisis de secuenciación masiva en paralelo

En otro ejemplo, se aplicó una combinación de PCR y métodos de pirosecuenciación en paralelo masivos para investigar los genomas de muestras de células múltiples que expresaban diferentes proteínas ZFN dirigidas contra la misma secuencia. Se sintetizaron tres variantes de un cebador de PCR de orientación directa (5'-XXXCACCAAGTTGTATTGCCTTCTCA-3') (SEQ ID NO: 137) en donde XXX = GGG, CCC, o GGC y tres variantes de un cebador de PCR de orientación inversa (5'-XXXATAGGCTTGAGCCAAGCAATCTT-3') (SEQ ID NO: 138) en donde XXX=GCC, CCG o CGG (IDT, Coralville, IA). Las etiquetas de 3 pb en el extremo 5' de cada cebador sirven como un identificador clave e indican la muestra celular de la que se origina el amplicón. Se usaron pares de cebadores con las mismas etiquetas (claves) identificadoras en combinación para amplificar el ADN genómico purificado procedente de muestras de células de maíz en las siguientes condiciones: volumen de reacción 50 ul que contenía molde de ADNg 40 ng, 20 pmol de cada cebador, DMSO 1 % y 10 unidades de Pf polimerasa Accuprime (Invitrogen, Carlsbad, CA) en el tampón enzimático del fabricante. Los productos de amplificación del tamaño esperado resultaron de ciclos de amplificación que consistían en 95°C-1', (95 °C-30", 65 °C-30", 72 °C-1') X 30, mantenido a 4 °C y se purificaron usando el kit de purificación de PCR MinElute de Qiagen (Valencia, CA) siguiendo las recomendaciones del fabricante.

Las reacciones de pirosecuenciación masiva en paralelo (también conocidas como secuenciación 454) se realizaron directamente sobre los productos de la PCR como se describe en (Margulies *et al.* (2005) Nature 437: 376-380) por 454 Life Sciences (Branford, CT). En los análisis de los resultados de la secuenciación 454 se realizó identificando lecturas de secuencias que contenían delecciones del tamaño esperado y la posición dentro de la molécula de ADN.

Los resultados de estos análisis indican la presencia de delecciones pequeñas múltiples en el sitio de escisión esperado para estas ZFN, como se muestra en la FIG. 68. Estas delecciones se localizan exactamente en el sitio diana de ZFN e indican que las roturas bc, inducidas por la ZFN, se generaron en el genoma y posteriormente se repararon por unión de extremos no homólogos, NHEJ. Estos resultados muestran además la capacidad de estas ZFN modificadas genéticamente para inducir roturas bicatenarias dirigidas, de una manera específica en un locus génico endógeno dentro de especie de cultivo.

Ejemplo 18: Diseño de ADN donante para la integración dirigida

En este ejemplo, se entiende que el ADN donante incluye moléculas de ADN bicatenario que se suministran en células de plantas y se incorporan en el genoma nuclear. El mecanismo mediante el cual se produce esta incorporación se realiza mediante unión de extremos no homólogos (NHEJ; revisado por Cahill *et al.*, (2006) Mechanisms Front Biosci. 1: 1958-76) de homología independiente u otro mecanismo similar en el sitio de una rotura bicatenaria en el ADN nuclear. Dicha incorporación de tipo ligamiento, dirigida por NHEJ de ADN donante en el genoma se denomina integración al azar, dado que la posición de la integración del ADN donante se determina principalmente por la presencia de una rotura de ADN bicatenario. En este mecanismo, la integración de ADN donante en el genoma no depende ni de la secuencia de nucleótidos del genoma en el sitio de la rotura ni de la secuencia de nucleótidos del propio donante. Por lo tanto, durante la integración al azar, la "dirección" en el genoma en el cual se incorpora el ADN donante, no está especificada ni prevista basándose en la secuencia del ADN donante. La integración al azar es el mecanismo primario mediante el cual se produce la transgénesis de ADN donante durante la transformación de plantas convencional mediante el suministro de ADN mediado por *Agrobacterium* o biolístico en células de plantas vivas.

A diferencia de la integración al azar, también puede incorporarse ADN donante en el genoma mediante integración dirigida. Se entiende que la integración dirigida se produce en el sitio (posición) de una rotura bicatenaria mediante mecanismos dependientes de homología, tales como hibridación de cadena sencilla dependiente de homología o recombinación homóloga (revisado en van den Bosch *et al.* (2002) Biol Chem. 383(6): 873-892). En el caso de reparación de rotura de ADN dependiente de homología, ADN donante que contiene la secuencia de nucleótidos con identidad o similitud al ADN en el sitio de la rotura puede incorporarse en este sitio. Por lo tanto, la "dirección" en la cual se integra el ADN donante en el genoma depende de la identidad de secuencia o de la similitud de secuencia de nucleótidos entre el genoma y las moléculas de ADN donante. En sistemas de plantas, se sabe que la reparación de roturas bicatenarias en el ADN, utiliza rutas tanto de unión de extremos no homólogos, NHEJ, como dependientes de homología (revisado por Puchta (2005) J. Exp. Bot. 56: 1-14).

En este ejemplo, se describe el diseño y la construcción de moléculas de ADN donante que van a integrarse en el genoma mediante integración dirigida en el sitio de una rotura bicatenaria inducida por proteínas ZFN específicas de secuencia. Diferentes proteínas ZFN pueden inducir roturas bicatenarias en diferentes nucleótidos en la secuencia

génica diana; el sitio específico de la rotura bicatenaria inducida recibe el nombre de posición.

Como se describe en el Ejemplo 13, se ha caracterizado la secuencia de nucleótidos de un gen diana, IPP2K, de maíz. Posteriormente, se diseñaron proteínas ZFN que se unen a bases específicas de ese gen diana (Ejemplo 14) y se validó su actividad de unión/escisión en esa secuencia dentro del gen diana en ambos sistemas heterólogos y contra el gen endógeno en células de maíz (Ejemplos 15-17). En esta memoria, se describe la construcción de diversas moléculas donantes diseñadas para incorporar en el genoma del maíz en la posición de la rotura bicatenaria mediada por ZFN en el gen IPP2K mediante integración dirigida. Un experto en la técnica puede construir una molécula de ADN donante diseñada para incorporar en una rotura bicatenaria inducida por ZFN mediante integración dirigida conducida por homología en cualquier posición en cualquier genoma del que se conoce su secuencia de nucleótidos y se espera que esa secuencia contenga una rotura bicatenaria.

En una realización descrita en la presente memoria, la molécula de ADN donante comprende un casete de expresión génica de tolerancia a herbicidas, autónomo, delimitado por segmentos de secuencia de nucleótidos idénticos a los del gen diana, IPP2K en la posición diana. En esta realización, se entiende que el casete de tolerancia a herbicidas autónomo, incluye una unidad transcripcional promotora (UTP) completa que contiene un promotor, un gen de tolerancia a herbicidas y una secuencia terminadora que se sabe que es funcional en las células de la planta. Un experto en la técnica puede seleccionar cualquier combinación de promotor, gen y terminador para constituir la UTP autónoma. También se incluye en esta construcción de plásmido, fragmentos de ADN con identidad de secuencia con el gen diana en maíz (IPP2K) en la posición indicada. Estos fragmentos sirven como "flancos de homología" del ADN donante y dirigen la incorporación de este donante en el gen diana en la posición específica mediante integración dirigida. Los flancos de homología se colocan tanto aguas arriba como aguas abajo de la UTP en la orientación 5' a 3' correcta con respecto a la unidad transcripcional promotora UTP. Un experto en la técnica puede imaginar flancos de homología de diverso tamaño y orientación en una construcción de ADN donante.

En otra realización descrita en la presente memoria, la molécula de ADN donante comprende una construcción de plásmido que contiene un casete de expresión génica de tolerancia a herbicidas no autónomo delimitado por segmentos de secuencia de nucleótidos idénticos a los del gen IPP2K en la posición diana. En esta realización, se entiende que el casete de tolerancia a herbicidas no autónomo incluye una unidad transcripcional promotora (UTP) incompleta que carece de un promotor funcional. La UTP no autónoma contiene un gen de tolerancia a herbicidas y una secuencia terminadora que se sabe que es funcional en células de plantas. Un experto en la técnica puede seleccionar cualquier combinación de gen y secuencia terminadora para constituir una UTP no autónoma. En este ejemplo de un donante no autónomo, la expresión del gen de tolerancia a herbicidas depende de la incorporación del segmento donante en una localización genómica próxima a un promotor funcional que puede dirigir la expresión de ese gen. Se puede imaginar la situación relativamente inusual en la que el donante se incorpore, mediante integración al azar, en un locus genético donde reside un promotor fortuito y que está disponible para dirigir la expresión del gen de tolerancia a herbicidas. Como alternativa, basándose en la presencia de flancos de homología de fragmentos de ADN de la longitud apropiada con la identidad de secuencia en el gen diana en una posición específica en el maíz dentro de la construcción de ADN donante, puede producirse la integración dirigida exacta del ADN donante en el gen diana en la posición específica (como se describe para el donante autónomo) y por lo tanto aprovechar el promotor endógeno de dicho gen diana. En esta realización, los flancos de homología se colocan tanto aguas arriba como aguas abajo de la UTP en la orientación 5' a 3' correcta con respecto a la UTP. Un experto en la técnica puede imaginar flancos de homología de diverso tamaño y orientación en una construcción de ADN donante.

En las dos realizaciones descritas en la presente memoria (diseño de donante autónomo y no autónomo), las construcciones de plásmido contienen típicamente elementos adicionales que permiten la clonación, expresión del gen de tolerancia a herbicidas y posterior análisis. Dichos elementos incluyen orígenes de replicación bacteriana, sitios de restricción modificados genéticamente, etc. y se describen más adelante. Un experto en la técnica puede imaginar la utilización de los diferentes elementos que comprenden una molécula de ADN donante.

A. Cepas bacterianas y condiciones de cultivo

Se desarrollaron cepas de *Escherichia coli* (One Shot[®] Top 10, células químicamente competentes; MAX Efficiency[®] DH5 α [™], células químicamente competentes, Invitrogen Life Technologies, Carlsbad, CA), a una temperatura de 37 °C, durante 16 horas usando caldo de Luria-Bertani (LB: Bactotripton 10 g/l, NaCl 10 g/l, extracto de Bactolevadura 5 g/l), agar LB (caldo LB más Bactoagar 15 g/l) o caldo Terrific (TB: Bactotripton 12 g/l, extracto de Bactolevadura 24 g/l, glicerol a 0,4 % v/v, KH₂PO₄ 17 mM, K₂HPO₄ 72 mM). Los cultivos líquidos se agitaron a 200 rpm. Al medio se añadió, según necesidad, cloramfenicol (50 μ g/ml), kanamicina (50 μ g/ml), o ampicilina (100 μ g/ml). Todos los antibióticos, medios de cultivo y reactivos de tampón usados en este estudio se adquirieron en Sigma-Aldrich Corporation (St. Louis, MO) o Difco Laboratories (Detroit, MI).

B. Posición 1 del esqueleto del plásmido

Mediante ingeniería genética se diseñó un esqueleto del plásmido que contenía flancos de homología para la posición 1 de IPP2K, para permitir la integración de cualquier secuencia de ADN donante en el sitio diana

correspondiente del gen IPP2K. Un experto en la técnica puede imaginar esqueletos plasmídicos usando diversos sitios de clonación, elementos de diseño modular y secuencias homólogas en cualquier secuencia diana dentro del genoma de interés. El esqueleto plasmídico ilustrado en la presente memoria, se originó con el vector plásmido base pBC SK(-) fagémido (3,4 Kbp) (Stratagene, La Jolla, CA). Para construir la posición -1 del esqueleto plasmídico, se usó una síntesis de cuatro etapas, como se describe más adelante.

En la etapa n.º 1, se preparó el plásmido base. Tres μg de pBC SK(-) se linealizaron usando 10 unidades de endonucleasas de restricción *Spe* I y 10 unidades de *Not* I (New England Biolabs, Beverly, MA), durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE (Tris-acetato 0,04 M, EDTA 0,002 M) a 1,0 % complementado con bromuro de etidio al 0,5 % (Sigma-Aldrich Corporation, St. Louis, MO). Los fragmentos de ADN se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). El vector de subclonación digerido con *Spe* I/*Not* I de 3,4 Kpb, pBC SK(-), se escindió del gel y se purificó de acuerdo con las directrices del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA).

En la etapa n.º 2, se aislaron flancos de homología 5' y 3' de la posición 1 de IPP2K. A través de Integrated DNA Technologies, Inc. (Coralville, IA) se sintetizaron los siguientes cebadores oligonucleotídicos, en condiciones de desalinización convencionales y se diluyeron con agua a una concentración de 0,125 $\mu\text{g}/\mu\text{l}$:

5'-GCGGCCGCGTCTCACCGCGGCTTGGGGATTGGATACGGAGCT -3' (SEQ ID NO: 143)

5'- ACTAGTGATATGGCCCCACAGGAGTTGCTCATGACTTG -3' (SEQ ID NO: 144)

5'- ACTAGTCCAGAACTGGTTGAGTCGGTCAAACAAGATTGCT -3' (SEQ ID NO: 145)

5'- GTCGACCTTGATGCTACCCATTGGGCTGTTGT -3' (SEQ ID NO: 146)

Las reacciones de amplificación por PCR se realizaron usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y consistían en lo siguiente: cinco μl de tampón II 10X LA PCR™ (Mg^{2+}), molde de ADN bicatenario 20 ng (maíz Hill), 10 pmol de cebador oligonucleotídico directo, 10 pmol de cebador oligonucleotídico inverso, mezcla de dNTP 8 μl (2,5 mM cada uno), H_2O 33,5 μl , ADN polimerasa *TaKaRa LA Taq*™ 0,5 μl (2,5 unidades), 1 gota de aceite mineral. Las reacciones de la PCR se realizaron usando un ciclador térmico de ADN de 48 muestras de 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/mantenido. Quince μl de cada reacción PCR se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE al 1,0 % complementado con bromuro de etidio al 0,5 %. Los fragmentos amplificados se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los productos de amplificación esperados se diagnosticaron por la presencia de fragmentos de ADN de 0,821 Kpb (flanco de homología 5') o 0,821 Kpb (flanco de homología 3').

Estos fragmentos se escindieron del gel y se purificaron de acuerdo con las instrucciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Después, los fragmentos purificados se clonaron en el plásmido pCR2.1 usando el kit TOPO TA Cloning® (con el vector pCR®2.1) y células de *E. coli* químicamente competentes One Shot® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 $\mu\text{l}/\text{ml}$, se inocularon colonias individuales y se incubaron durante 16 horas a 37 °C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μg de plásmido aislado de plásmidos del clon de flanco de homología 5' se digirieron con 10 unidades de *Spe* I y *Not* I. Tres plásmidos del clon del flanco de homología principal se digirieron con 10 unidades de *Spe* I y 20 unidades de *Sal* I (New England Biolabs, Beverly, MA). Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE al 1,0 % complementado con bromuro de etidio al 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de plásmidos esperados se diagnosticaron por la presencia de fragmentos de ADN insertados de 0,821 Kpb (flanco de homología 5') o 0,821 Kpb (flanco de homología 3') además del vector pCR®2.1 de 3,9 Kpb.

Las reacciones de secuenciación bicatenarias de los clones de plásmidos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN de Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI). La secuencia del fragmento de 0,821 Kpb correspondiente a la posición-1 del flanco de homología 5' procedente de IPP2K se muestra en la Fig. 87 (SEQ ID NO: 171). La secuencia del

fragmento de 0,821 Kpb correspondiente a la posición-1 del flanco de homología 3' procedente de IPP2K se muestra en la Fig. 88 (SEQ ID NO: 172).

En la etapa n.º 3 la posición-1 de los flancos homología 5' se ligó en el plásmido base. Los fragmentos restringidos correspondientes a los clones que contenían la posición-1 correcta de la secuencia del flanco de homología 5' se escindieron del gel y se purificaron de acuerdo con las directrices del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Los fragmentos correspondientes al flanco de homología 5' de la posición-1 (0,821 Kpb) se ligaron después con el plásmido base purificado digerido con *Spe I/NotI* (etapa n.º 1) a una proporción de vector:inserto de 1:5 usando 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 μ l en condiciones de incubación de 16 horas en un baño con agua a 16 °C. Cinco μ l de la reacción de ligamiento se transformaron posteriormente en células *E. coli* químicamente competentes One Shot® Top 10, (Invitrogen Life Technologies, Carlsbad, CA) y se sembraron en placas en las condiciones de selección descritas por el fabricante. En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 μ l/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm.

Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de ADN plasmídico aislado se digirieron con 10 unidades de *Spe I* y *Not I*. (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE al 1,0 % complementado con bromuro de etidio al 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de plásmidos esperados se diagnosticaron por la presencia de un fragmento de ADN insertado de 0,821 Kpb (flanco de homología 5') además del plásmido base de 3,4 Kpb.

En la etapa n.º 4, los flancos de homología 3' en la posición-1 se ligaron en el producto de la etapa n.º 3. Tres μ g del producto modificado genéticamente, descrito en la etapa n.º 3, se linealizaron usando 10 unidades de *Spe I* y 20 unidades de *Sal I* (New England Biolabs, Beverly, MA), endonucleasas de restricción, durante 1 hora a 37 °C. El ADN restringido se sometió a 100 V durante 1 hora en un gel de agarosa TAE (Tris-acetato 0,04 M, EDTA 0,002 M) a 1,0 % complementado con bromuro de etidio a 0,5 % (Sigma-Aldrich Corporation, St. Louis, MO). Los fragmentos de ADN se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). El producto digerido con *Spe I/Sal I* de ~4,2 Kpb de la etapa n.º 3 se escindió con gel y se purificó de acuerdo con las instrucciones del fabricante usando el kit de extracción de gel QIAquick Gel (QIAGEN Inc., Valencia, CA).

Los fragmentos aislados del donante del flanco de homología 3' (0,821 Kpb) generados en la etapa n.º 2 se combinaron posteriormente con el producto de la etapa n.º 3 que se digirió con *Spe I/Sal I* y se purificaron como se describe anteriormente en una reacción de ligamiento de 20 μ l usando una proporción de vector:inserto de 1:5 y 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA). Las reacciones de ligamiento se incubaron durante 16 horas en un baño con agua a 16 °C. Después del ligamiento, 5 μ l de la reacción de ligamiento se transformaron en células químicamente competentes MAX Efficiency® DH5 α ™ (Invitrogen Life Technologies, Carlsbad, CA) siguiendo las recomendaciones del fabricante. En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con cloramfenicol 50 μ l/ml, se inocularon colonias individuales.

Los cultivos se incubaron durante 16 horas a 37 °C agitando a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de ADN plasmídico aislado se digirieron con 10 unidades de *Sal I* y *Not I*. (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones esperados se diagnosticaron por la presencia de los fragmentos de ADN de 1,64 Kpb (inserto) y 3,33 Kpb (plásmido base). El plásmido resultante recibió el nombre de pDAB7471 (FIG. 70).

C. Posición -2 del esqueleto del plásmido

Para permitir la integración de cualquier secuencia de ADN donante en el sitio diana correspondiente del gen IPP2K, un esqueleto plasmídico, que contenía los flancos de homología para la posición -2 de IPP2K, se modificó genéticamente. Un experto en la técnica puede imaginar esqueletos plasmídicos usando diversos sitios de clonación, elementos de diseño modular y homólogos de secuencia en cualquier secuencia diana dentro del genoma de interés. El esqueleto plasmídico ilustrado en la presente memoria se originó con el vector plasmídico base pBC SK(-) fagémido (3,4 Kpb) (Stratagene, La Jolla, CA). Se usó una síntesis de cuatro etapas, como se describe más adelante, para construir la posición -2 del esqueleto del plásmido.

En la etapa n.º 1, se preparó el plásmido base. Tres μg de pBC SK(-) se linealizaron usando 10 unidades de *Spe* I y 10 unidades de *Not* I (New England Biolabs, Beverly, MA), endonucleasas de restricción, durante 1 hora a 37 °C. El ADN restringido se sometió a 100 V durante 1 hora en un gel de agarosa TAE (Tris-acetato 0,04 M, EDTA 0,002 M) a 1,0 % complementado con bromuro de etidio a 0,5 % (Sigma-Aldrich Corporation, St. Louis, MO). Los fragmentos de ADN se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). El vector de subclonación, digerido con *Spe* I/*Not* I, de 3,4 Kpb, pBC SK(-) se escindió del gel y se purificó siguiendo las indicaciones del fabricante usando el kit de extracción en gel QIAquick (QIAGEN Inc., Valencia, CA).

En la etapa n.º 2, se aislaron los flancos de homología 5' y 3' de la posición -2 de IPP2K. A través de Integrated DNA Technologies, Inc. (Coralville, IA) se sintetizaron los siguientes cebadores oligonucleotídicos en condiciones de desanilización convencionales y se diluyeron con agua a una concentración de 0,125 $\mu\text{g}/\mu\text{l}$:

5'-GCGGCCGCTAGATAGCAGATGCAGATTGCT-3' (SEQ ID NO: 147)

5'-ACTAGTATTGGCACCCAGGTGTTGGCTCA-3' (SEQ ID NO: 148)

5'-ACTAGTCATGTTCGATGGTGGGGTATGGTTCAGATTCAG-3' (SEQ ID NO: 149)

5'-GTCGACGTACAATGATTTTCAGTTACGGCCTCAGGAC-3' (SEQ ID NO: 150)

Las reacciones de amplificación por PCR se realizaron usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y consistían en lo siguiente: 5 μl de tampón II 10X LA PCR™ (Mg²⁺), molde de ADN bicatenario 20 ng (Hill de maíz), 10 pmol de cebador oligonucleotídico directo, 10 pmol de cebador oligonucleotídico inverso, mezcla de dNTP 8 μl (2,5 mM de cada uno), H₂O 33,5 μl , ADN polimerasa *TaKaRa LA Taq*™ 2,5 μl (2,5 unidades), 1 gota de aceite mineral. Las reacciones de PCR se realizaron usando un ciclador térmico de ADN de 48 muestras de Perkin-Elmer Cetus (Norwalk, CT), usando las siguientes condiciones de ciclo: 94 °C, 4 min/1 ciclo; 98 °C 20 s, 55 °C 1 min, 68 °C 1 min/30 ciclos; 72 °C, 5 min/1 ciclo; 4°C/mantenido. Quince μl de cada reacción PCR se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos amplificados se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los productos de amplificación esperados se diagnosticaron por la presencia de un fragmento de ADN de 0,855 Kpb (flanco de homología 5') o de 0,845 Kpb (flanco de homología 3'). Estos fragmentos se escindieron del gel y se purificaron de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Después, los agentes purificados se clonaron en el plásmido pCR2.1 usando el kit TOPO TA Cloning® (con el vector pCR®2.1) y células de *E. coli* químicamente competentes TOP One Shot® (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 $\mu\text{l}/\text{ml}$, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μg de plásmido aislado de plásmidos del clon del flanco de homología 5' se digirieron con 10 unidades de *Spe* I y *Not* I. Tres plásmidos del clon del flanco de homología- cebador se digirieron con 10 unidades de *Spe* I y 20 unidades de *Sal* I (New England Biolabs, Beverly, MA). Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C.

El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de los plásmidos esperados se diagnosticaron por la presencia de fragmentos de ADN insertados de 0,855 Kpb (flanco de homología 5') o de 0,845 Kpb (flanco de homología 3') además del vector pCR®2.1 de 3,9 Kpb.

Las reacciones de secuenciación bicatenarias de los clones del plásmidos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN de Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI). La secuencia del fragmento de 0,855 Kpb correspondiente a la posición-2 del flanco de homología 5' procedente de IPP2K se muestra en la Fig. 89 (SEQ ID NO: 139). La secuencia del fragmento de 0,845 Kpb correspondiente a la posición-2 del flanco de homología 3' procedente de IPP2K se muestra en la Fig. 90 (SEQ ID NO: 140).

En la etapa n.º 3 la posición-1 de los flancos homología 5' se ligaron en el plásmido base. Los fragmentos restringidos correspondientes a los clones que contenían la secuencia correcta del flanco de homología 5' de la posición-1 se escindieron del gel y purificaron de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Los fragmentos correspondientes al flanco de homología 5'

de la posición-1 (0,855 Kpb) se ligaron después al plásmido base purificado digerido con *Spe I/Not I* (etapa n.º 1) a una proporción de vector:inserto de 1:5 usando 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 μ l en condiciones de incubación de 16 horas en un baño con agua a 16 °C.

5 Después, cinco μ l de la reacción de ligamiento se transformaron en células de *E. coli* químicamente competentes One Shot® Top 10, (Invitrogen Life Technologies, Carlsbad, CA) y se sembraron en placas usando las condiciones de selección descritas por el fabricante. En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 μ l/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de ADN plasmídico aislado se digirieron con 10 unidades de *Spe I* y *Not I* (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de plásmidos esperados se diagnosticaron por la presencia de un fragmento de ADN insertado de 0,855 Kpb (flanco de homología 5') además del plásmido base de 3,4 Kpb.

20 En la etapa n.º 4, los flancos de homología 3' en la posición-2 se ligaron en el producto de la etapa n.º 3. Tres μ g del producto modificado genéticamente, descrito en la etapa n.º 3, se linealizaron usando 10 unidades de *Spe I* y 20 unidades de *Sal I* (New England Biolabs, Beverly, MA), endonucleasas de restricción, durante 1 hora a 37 °C. El ADN restringido se sometió a 100 V durante 1 hora en un gel de agarosa TAE a 1,0% (Tris-acetato 0,04 M, EDTA 0,002 M) a 1,0 % complementado con bromuro de etidio a 0,5 % (Sigma-Aldrich Corporation, St. Louis, MO). Los fragmentos de ADN se visualizaron usando luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). El producto digerido con *Spe I/Sal I* de 4,25 Kpb de la etapa n.º 3 se escindió del gel y se purificó de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick Gel (QIAGEN Inc., Valencia, CA).

30 Los fragmentos aislados del donante del flanco de homología 3' (0,845 Kpb) generados en la etapa n.º 2 se combinaron posteriormente con el producto de la etapa n.º 3 que se digirió con *Spe I/Sal I* y se purificaron como se describe anteriormente en una reacción de ligamiento de 20 μ l usando una proporción de vector:inserto de 1:5 y 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA). Las reacciones de ligamiento se incubaron durante 16 horas en un baño con agua a 16 °C. Después del ligamiento, 5 μ l de la reacción de ligamiento se transformaron en células químicamente competentes MAX Efficiency® DH5 α ™ (Invitrogen Life Technologies, Carlsbad, CA) siguiendo las recomendaciones del fabricante. En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con cloramfenicol 50 μ l/ml, se inocularon colonias individuales. Los cultivos se incubaron durante 16 horas a 37 °C agitando a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de plásmido aislado se digirieron con 10 unidades de *Sal I* y *Not I*. (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones esperados se diagnosticaron por la presencia de dos fragmentos de ADN de ~1,7 Kpb (inserto) y 3,33 Kpb (plásmido base). El plásmido resultante recibió el nombre de pDAB7471 (FIG. 71).

D. Construcción del casete de expresión del gen de tolerancia a herbicidas autónomo

Se construyó un casete de expresión del gen de tolerancia a herbicidas autónomo que comprendía un promotor que contenía una unidad transcripcional promotora (UTP) completa, el gen de tolerancia a herbicidas y secuencias de terminación de poliadenilación (poliA) (FIG. 72). En esta realización, la secuencia promotora procede de actina-1 de *O. sativa* [McElroy *et al.* (Plant Cell 2, 163-171; 1990); Registro GenBank S44221 y Registro GenBank X63830]. El gen de tolerancia a herbicidas comprende el gen PAT (fosfotricin acetil transferasa), que confiere resistencia al herbicida bialafos (una versión modificada de la región que codifica PAT originalmente procedente de *Streptomyces viridochromogenes* (Registro GenBank R M22827; Wohlleben *et al.* Gene 70,25-37; 1988). Las modificaciones en la secuencia original de la fase de lectura abierta más larga de M22827 son sustanciales e incluyen la alteración del patrón de utilización de codones para optimizar la expresión en plantas. Excepto por la sustitución de la metionina por la valina, como el primer aminoácido codificado, y la adición de la alanina como el segundo aminoácido codificado, la proteína codificada a partir de fase de lectura abierta PAT de pDAB3014, es idéntica a la codificada por la fase de lectura abierta más larga del número de registro M22827. La versión reconstruida de PAT se encuentra con el número de registro GenBank I43995. Las secuencias terminadoras proceden de la lipasa de *Z. mays* L. [clon de ADNc de lipasa de maíz con número de registro GenBank L35913, excepto que, en la posición 1093 de L35913, una C se reemplazó con una G en la posición 2468 en pDAB3014. Esta secuencia de maíz comprende la región no traducida 3'/región terminadora de la transcripción para el gen PAT].

A través de Integrated DNA Technologies, Inc. (Coralville, IA) se sintetizaron los siguientes cebadores oligonucleotídicos, en condiciones de desanilización convencionales y se diluyeron con agua a una concentración de 0,125 µg/µl:

5'- ACTAGTAACTGACCTCACTCGAGGTCATTCATATGCTTGA -3' (SEQ ID NO: 151)

5 5'- ACTAGTGTGAATTCAGCACTTAAAGATCT -3' (SEQ ID NO: 152)

Las reacciones de amplificación por PCR se realizaron usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y consistían en lo siguiente: 5 µl 10X de tampón II LA PCR™ (Mg²⁺), 20 ng de molde bicatenario [ADN plasmídico pDAB3014], 10 pmol de cebador oligonucleotídico directo, 10 pmol de cebador oligonucleotídico inverso, 8 µl de mezcla de dNTP (2,5 mM cada uno), H₂O 33,5 µl, ADN polimerasa TaKaRa LA Taq™ 0,5 µl (2,5 unidades), 1 gota de aceite mineral. La reacciones de PCR se realizaron usando un ciclador térmico de ADN de 48 muestras de Perkin-Elmer Cetus (Norwalk, CT) con las siguientes condiciones de ciclo: 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/mantenido. Quince µl de cada reacción de PCR se sometieron a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %.

15 Los fragmentos amplificados se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los productos de amplificación esperados se diagnosticaron por la presencia de un fragmento de ADN de 2,3 Kpb. Este fragmento se escindió del gel y se purificó de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). El fragmento purificado se clonó después en el plásmido pCR2.1 usando el kit de clonación TOPO TA Cloning® y se transformó en células de *E. coli* químicamente competentes One Shot® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

25 En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 µl/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37 °C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres µg de plásmido aislado se digirieron con 10 unidades de *Spe* I y *Not* I. Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C.

30 El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de plásmidos esperados se diagnosticaron por la presencia de un fragmento de ADN insertado de 2,325 Kpb además del vector pCR@2.1 de 3,9 Kpb. Las reacciones de secuenciación bicatenarias de los clones plasmídicos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN de Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI).

40 E. Inserción del casete génico de tolerancia a herbicidas autónomo en el donante autónomo en el esqueleto del plásmido

Para crear un plásmido donante, el casete génico de tolerancia a herbicidas autónomo descrito en el Ejemplo 18D se insertó en las construcciones del esqueleto del plásmido descritas en los Ejemplos 18B y 18C. El fragmento restringido procedente un clon que contenía la secuencia esperada de 2,325 Kpb descrito anteriormente (FIG. 72) se escindió del gel y se purificó de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA).

Después, este fragmento se combinó en una reacción de ligamiento con cualquiera de pDAB7471 purificado (posición-1 del esqueleto del plásmido, FIG. 70) o pDAB 7451 (posición-2 del esqueleto del plásmido FIG. 71) que se había digerido con la enzima de restricción *Spe* I y posteriormente se desfosforiló. El ligamiento se realizó en las siguientes condiciones: proporción vector:inserto de 1:5 y 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 µl en condiciones de 16 horas de incubación en un baño con agua a 16 °C. Después, cinco µl de la reacción de ligamiento se transformaron en 50 µl de células de *E. coli* químicamente competentes MAX Efficiency® DH5α™ (Invitrogen Life Technologies, Carlsbad, CA) y se sembraron en placas en las condiciones de selección descritas por el fabricante.

55 En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con cloramfenicol 50 µl/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD

Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μg de ADN plasmídico aislado se digirieron con 10 unidades de *Spe* I y *Not* I (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de los plásmidos esperados se diagnosticaron por la presencia de fragmentos de ADN de 2,325 Kpb y ~4,9 Kpb (vector pDAB7471) o 2,325 Kpb y ~5,0 Kpb (vector pDAB7451).

Los plásmidos resultantes se denominaron pDAB7422 (donante autónomo posición-1) (FIG. 73) y pDAB7452 (donante autónomo posición-2) (FIG. 74), respectivamente.

10 F. Construcción del casete de expresión del gen de tolerancia a herbicidas no autónomo

Se construyó un casete de expresión del gen de tolerancia a herbicidas no autónomo que comprendía una unidad transcripcional promotora (UTP) incompleta (FIG. 75). En esta realización, se usó una estrategia que aprovechaba la funcionalidad de una secuencia 2A (Mattion, N.M., Harnish, E.C., Crowley, J.C. & Reilly, P.A. (1996) *J. Virol.* 70, 8124-8127) procedente del virus *Thesoa assigna*, un gen de tolerancia a herbicidas y secuencias de terminación de poliadenilación (poliA), pero no promotoras. En esta realización, la secuencia señal 2A de terminación de la traducción, se había modificado genéticamente para estar traducionalmente en fase con el gen de tolerancia a herbicidas. Además, la secuencia codificante 2a/herbicida se había modificado genéticamente para que coincidiera con la fase de lectura de la traducción del gen IPP2K diana. El gen de tolerancia a herbicidas comprende el gen PAT (fosfinotricin acetil transferasa), que confiere resistencia al herbicida bialafos (una versión modificada de la región que codifica PAT originalmente procedente de *Streptomyces viridochromogenes* (Registro GenBank R M22827; Wohlleben et al. *Gene* 70:25-37; 1988). Las modificaciones en la secuencia original de la fase de lectura abierta más larga de M22827 son sustanciales, e incluyen la alteración del patrón de utilización de codones para optimizar la expresión en plantas. Excepto por la sustitución de la metionina por la valina, como el primer aminoácido codificado, y la adición de alanina, como el segundo aminoácido codificado, la proteína codificada a partir de fase de lectura abierta PAT de pDAB3014 es idéntica a la codificada por la fase de lectura abierta más larga de M22827 (que comienza con GTG en la posición 244 de M22827). La versión reconstruida de PAT se encuentra con el número de registro GenBank I43995. Las secuencias terminadoras proceden de la lipasa de *Z. mays* L. [clon de ADNc de lipasa de maíz de número de registro GenBank L35913, excepto que se reemplazó una C en la posición 1093 de L35913 con una G en la posición 2468 en pDAB3014]. Esta secuencia del maíz comprende la región no traducida 3'/terminadora de la transcripción para el gen PAT.

A través de Integrated DNA Technologies, Inc. (Coralville, IA) se sintetizaron los siguientes cebadores oligonucleotídicos: en condiciones de desanilización convencionales y se diluyeron con agua a una concentración de 0,125 $\mu\text{g}/\mu\text{l}$:

5'- ACTAGTGGCGGGGAGAGGGCAGAGGAAGTCTTCTAACATGC

GGTGACGTGGAGGAGAATCCCGGCCCTAGGATGGCTTCTCCGGAGAGGAGAC

CAGTTGA -3 (SEQ ID NO:153)

35 5'- ACTAGTATGCATGTGAATTCAGCACTTAAAGATCT-3' (SEQ ID NO: 154)

Las reacciones de amplificación de PCR se realizaron usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y consistían en lo siguiente: 5 μl de tampón II LA PCR™ 10X (Mg^{2+}), 20 ng de molde bicatenario (ADN plasmídico pDAB3014), 10 pmol de cebador oligonucleotídico directo, 10 pmol de cebador oligonucleotídico inverso, 8 μl de mezcla de dNTP (2,5 mM cada uno), H_2O 33,5 μl , ADN polimerasa *TaKaRa LA Taq*™ 0,5 μl (2,5 unidades), 1 gota de aceite mineral. La reacciones de la PCR se realizaron usando un ciclador térmico de ADN de 48 muestras Perkin-Elmer Cetus (Norwalk, CT) en las siguientes condiciones de ciclo: 94 °C, 4 min/1 ciclo; 98 °C 20 s, 55 °C 1 min, 68 °C 2 min/30 ciclos; 72 °C, 5 min/1 ciclo; 4°C/mantenido. Quince μl de cada reacción PCR se sometieron a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos amplificados se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los productos de amplificación esperados se diagnosticaron por la presencia de un fragmento de ADN de ~1 Kpb. Este fragmento se escindió del gel y se purificó de acuerdo con las instrucciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). El fragmento purificado se clonó después en el plásmido pCR2.1 usando el kit de clonación TOPO TA Cloning® y se transformó en células de *E. coli* químicamente competentes One Shot® TOP10 (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 $\mu\text{l}/\text{ml}$, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmido NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μg de plásmido aislado se digirieron con 10 unidades de

Spe I. Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de los plásmidos esperados se diagnosticaron por la presencia de un fragmento de ADN insertado de ~1,0 Kpb y 3,9 Kpb (vector pCR@2.1). Las reacciones de secuenciación bicatenarias de los clones del plásmidos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI).

G. Inserción del casete génico de tolerancia a herbicidas no autónomo en el donante no autónomo en el esqueleto del plásmido

Para crear un plásmido donante, el casete del gen de tolerancia a herbicidas no autónomo descrito en el Ejemplo 18F se insertó en las construcciones del esqueleto del plásmido descritas en los Ejemplos 18B y 18C. El fragmento restringido correspondiente a un clon que contenía la secuencia correcta de 1 Kpb se escindió del gel y purificó de acuerdo con las indicaciones del fabricante usando el kit de extracción QIAquick (QIAGEN Inc., Valencia, CA). Después, este fragmento se combinó en una reacción de ligamiento con cualquiera de pDAB7471 purificado (posición-1 del esqueleto del plásmido) (FIG. 70) o pDAB 7451 (posición-2 del esqueleto del plásmido) (FIG. 71) que se había digerido con la enzima de restricción Spe I y posteriormente se desfosforiló. El ligamiento se realizó en las siguientes condiciones: proporción de vector:inserto de 1:5 y 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 µl en condiciones de 16 horas de incubación en un baño con agua a 16 °C. Después, cinco µl de la reacción de ligamiento se transformaron en 50 µl de células de *E. coli* químicamente competentes MAX Efficiency® DH5α™ (Invitrogen Life Technologies, Carlsbad, CA) y se sembraron en placas en las condiciones de selección descritas por el fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con cloramfenicol 50 µl/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrifuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres µg de ADN plasmídico aislado se digirieron con 10 unidades de Spe I y Not I (New England Biolabs, Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones de los plásmidos esperados se diagnosticaron por la presencia de fragmentos de ADN de 1,0 Kpb y 4,96 Kpb (vector pDAB7471) o 1,0 Kpb y ~5,0 Kpb (vector pDAB7451). Los plásmidos resultantes se denominaron pDAB7423 (donante no autónomo posición-1) (FIG. 76) y pDAB7454 (donante no autónomo posición-2) (FIG. 77), respectivamente.

H. ZFN posición 1 + secuencias donantes RH: Plásmido de combinación.

Como una estrategia alternativa al suministro de dos plásmidos distintos en una célula de planta (por ejemplo, conteniendo un plásmido elementos ZFN que conteniendo un segundo plásmido las secuencias donantes de tolerancia a herbicidas), se modificaron genéticamente plásmidos sencillos, que contenían todos los elementos necesarios ilustrados en esta patente. Los plásmidos de combinación descritos en este ejemplo contienen tanto las ZFN diseñadas para dirigirse y generar roturas bicatenarias en el locus IPP2K específico, así como la UTP (unidad transcripcional promotora) de PAT autónoma y/o la UTP 2A/PAT no autónoma y flancos donantes diseñados para integrarse en estos sitios de rotura.

Se utilizó tecnología Gateway®, que usa recombinación específica de sitio basada en el fago lambda (Landy, A. (1989) Ann. Rev. Biochem. 58:913), para transformar los vectores pDAB7422 y pDAB7423 (descritos en los ejemplos 6E y 6G) en vectores de destino Gateway®. Una vez transformados, los plásmidos que mantienen los casetes de expresión ZFN (alojados en vectores de entrada Gateway®) pueden movilizarse fácilmente al vector de destino creando un plásmido de combinación ZFN/ donante. Un µg de cada uno de dichos plásmidos se digirió con 10 unidades de Not I (New England Biolabs, Beverly, MA) durante 1 hora a 37 °C. La endonucleasa de restricción Not I se termoinactivó a 65 °C durante 15 minutos y los extremos de los fragmentos se defosforilaron posteriormente a 37 °C durante 1 hora usando 3 unidades de fosfatasa alcalina de camarón (SAP, *shrimp alkaline phosphatase*) (Roche Diagnostics GmbH, Mannheim, Alemania). El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos de los vectores (pDB7422 = 7,317 Kpb, pDAB7423 = 5,971 Kpb) se visualizaron con luz UV, el tamaño se calculó por comparación con una escalera de ADN de 1 Kpb, se escindieron del gel y posteriormente se purificaron de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA).

Este fragmento de vector se combinó después un fragmento *Not* I de 2,274 Kpb que contenía elementos de la tecnología Gateway® attR1, *ccdB*, *Cm^R*, y attR2 en una reacción de ligamiento realizada en las siguientes condiciones: proporción de vector:inserto de 1:5 y 500 unidades de ADN ligasa de T4 (Invitrogen Life Technologies, Carlsbad, CA) en un volumen de reacción de 20 μ l en condiciones de 16 horas de incubación en un baño con agua a 16 °C. Después, cinco 5 μ l de la reacción de ligamiento se transformaron en 50 μ l de células de *E. coli* químicamente competentes One Shot® *ccdB* Survival™ (Invitrogen Life Technologies, Carlsbad, CA) y se sembraron en placas en las condiciones de selección descritas por el fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con cloramfenicol 50 μ g/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de ADN plasmídico aislado se digirieron con 10 unidades de *Eco*RI I (New England Biolabs Inc., Beverly, MA) y se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron con luz UV y el tamaño de los fragmentos se calculó por comparación con una escalera de ADN de 1 Kpb. Los clones plasmídicos esperados se diagnosticaron por la presencia de fragmentos de ADN de 1,448 Kpb, 1,946 Kpb y 6,197 Kpb para el donante HR posición-1 UTP PAT autónomo y de 5,807 Kpb y 2,438 Kpb para el donante HR posición-1 PAT no autónomo. Los plásmidos resultantes se denominaron pDAB7424 (donante autónomo posición-1 adaptado a Gateway®) (FIG. 78) y pDAB7425 (donante no autónomo posición-1 adaptado a Gateway®) (FIG. 79), respectivamente.

Como resultado de estas manipulaciones de clonación, los plásmidos pDAB7424 y pDAB7425 se denominaron vectores de destino Gateway®. El vector pDAB7412 tenía funcionalidad como un vector de entrada Gateway® que contenía los siguientes elementos: *ZmUbi1v.2/ZFN12/Zm Per5* UTR 3'. Para transferir un casete de expresión ZFN (vector de entrada Gateway®) en cualquier molécula donante autónoma o no autónoma (vector de destino Gateway®), se realizó una reacción LR Clonase™ II (Invitrogen Life Technologies, Carlsbad, CA) como indicó el fabricante a una relación de 50 ng (vector de entrada): 150 ng/ μ l (vector de destino). Los plásmidos de combinación positivos resultantes se denominaron pDAB7426 (donante HR autónomo posición-1/ZFN12) (FIG. 80) y pDAB7427 (donante HR no autónomo/ZFN12) (FIG. 81).

30 Ejemplo 19: Suministro de ADN codificante de ZFN y donante en células de plantas

Para permitir la integración de ADN donante, mediada por ZFN, en el genoma de la planta a través de integración dirigida, se entiende que se requiere el suministro de ADN que codifica ZFN seguido de la expresión de una proteína ZFN funcional en la célula de la planta. También se requiere el suministro simultáneo de ADN donante en dicha célula de planta, de tal manera que la proteína ZFN funcional puede inducir roturas bicatenarias en el ADN diana que después se repara mediante integración dirigida por homología del ADN donante en el locus diana. Un experto en la técnica puede imaginar que la expresión de una proteína ZFN funcional puede lograrse mediante diversos métodos, incluyendo, pero sin limitación, transgénesis de la construcción que codifica la ZFN, o la expresión transitoria de la construcción que codifica la ZFN. En estos dos casos, la expresión de la proteína ZFN funcional y el suministro de ADN donante en la célula de la planta se realizaron simultáneamente para dirigir la integración diana

En los ejemplos dictados en la presente memoria, se demuestran métodos para el suministro simultáneo de ADN codificante de ZFN y donante en células de plantas. Un experto en la técnica usaría cualquiera de una diversidad de métodos de suministro de ADN apropiados para células de plantas, incluyendo, pero sin limitación, transformación mediada por *Agrobacterium*, suministro de ADN basado en biolística o suministro de ADN mediado por Whiskers™. En una realización descrita en la presente memoria, se realizaron experimentos de suministro de ADN mediado por Whiskers™ usando diversas combinaciones de ADN donante con construcciones de ADN codificantes de ZFN. Estas combinaciones incluyen 1) un solo plásmido que contiene tanto la secuencia codificante de ZFN como ADN donante y 2) dos plásmidos distintos, uno que contiene la secuencia codificante de ZFN y el otro que contiene el ADN donante. En otra realización, se realizó el suministro de ADN basado en biolística usando diversas combinaciones de construcciones de ADN donante con ADN codificante de ZFN. Un experto en la técnica puede deducir que estas combinaciones podrían incluir 1) un solo plásmido que contenga tanto la secuencia codificante de ZFN como el ADN donante y 2) dos plásmidos distintos, conteniendo uno la secuencia codificante de ZFN y el otro ADN donante.

A. Suministro de ADN mediado por Whiskers™

Como se ha descrito anteriormente en la presente memoria, se produjeron cultivos de células Hi-II embriogénicas de maíz, y se usaron como fuente de células de plantas vivas en las que se demostró la integración dirigida. Un experto en la técnica puede imaginar la utilización de cultivos de células procedentes de diversas especies de plantas, o de tejidos de plantas diferenciados, procedentes de una diversidad de especies de plantas, como la fuente de células de plantas vivas en las que se demostró la integración dirigida.

En este ejemplo, se subcultivaron 12 ml de VCC a partir de una línea celular previamente crioconservada más 28 ml de medio acondicionado en 80 ml de medio GN6 líquido (medio N6 (Chu *et al.*, 1975), 2,4-D 2,0 mg/l, sacarosa 30 g/l, pH 5,8) en un matraz Erlenmeyer de 500 ml, y se colocó en un agitador a 125 rpm a 28 °C. Esta etapa se repitió 2 veces usando la misma línea celular de tal manera que se distribuyó un total de 36 ml de VCC en 3 matraces.

Después de 24 horas los medios GN6 líquidos se retiraron y se reemplazaron con 72 ml de medio osmótico GN6 S/M (medio N6, 2,4-D 2,0 mg/l, sacarosa 30 g/l, sorbitol 45,5 g/l, manitol 45,5 g/l, mioinositol 100 mg/l, pH 6,0). El matraz se incubó en la oscuridad durante 30-35 minutos a 28 °C con agitación moderada (125 rpm). Durante el periodo de incubación, se preparó una suspensión de 50 mg/ml de whiskers (fibras cortas) de carburo de silicio (Advanced Composite Materials, LLC, Greer, SC) añadiendo 8,1 ml de medio líquido GN6 S/M a 405 mg de whiskers estériles de carburo de silicio.

Después de la incubación en medio osmótico GN6 S/M, el contenido de cada matraz se agrupó en un frasco de centrifuga de 250 ml. Después de que todas las células del matraz se sedimentasen en el fondo, se extrajo un volumen de contenido sobrante de aproximadamente 14 ml de medio líquido GN6 S/M y se recogió en un matraz estéril de 1 l para su uso futuro. La suspensión de whiskers, previamente hidratada, se mezcló a velocidad máxima en un agitador vorticial durante 60 segundos, y después se añadió al frasco de centrifuga.

En un ejemplo, en donde se suministraba un solo plásmido, que contenía tanto la secuencia codificante de ZFN como el ADN donante, en las células de la planta, se añadieron al frasco 170 μ g de ADN plasmídico circular purificado. En un ejemplo alternativo, en donde se suministraban conjuntamente dos plásmidos distintos, uno que contenía la secuencia codificante de ZFN y el otro el ADN donante, se ensayaron estrategias múltiples con cantidades de ADN. Una estrategia utilizó 85 μ g de ADN donante y 85 μ g de ADN codificante de dedos de zinc. Otras modificaciones utilizan proporciones molares de factor 10, 5 o 1 de ADN donante con respecto a un factor 1 de ADN de dedo de zinc, basándose en el tamaño (en pares de kilobases) de los plásmidos individuales de tal manera que se añadieron un total 170 μ g de ADN por frasco. En todos los casos de suministro conjunto, el ADN se agrupó previamente en un tubo antes de añadirse al frasco de centrifuga. Una vez añadido el ADN, el frasco se colocó inmediatamente en un mezclador de pintura comercial Red Devil 5400 modificado (Red Devil Equipment Co., Plymouth, MN) y se agitó durante 10 segundos. Después de la agitación, el cóctel de células, los medios, los whiskers y el ADN se añadió al contenido de un matraz de 1 l junto con medio líquido GN6 reciente de 125 ml para reducir el osmoticante. Se dejó que las células se recuperasen en un agitador, configurado a 125 rpm durante 2 horas. Seis ml de la suspensión dispersada se filtraron sobre papel de filtro Whatman n.º 4 (5,5 cm) usando una unidad recogedora de células de vidrio conectada a una línea de vacío interna de tal manera que se obtuvieron 60 filtros por frasco. Los filtros se colocaron en placas de 60 x 20 mm de medio sólido GN6 (igual que el medio líquido GN6 excepto que contiene con agente gelificante Gelrite 2,5 g/l) y se cultivaron a 28 °C en la oscuridad durante 1 semana.

B: Suministro de ADN mediado por biolística

En los ejemplos citados en la presente memoria, se subcultivaron suspensiones embrionarias de maíz en medio líquido GN6 aproximadamente 24 horas antes del experimento, como se describe anteriormente con la presente memoria. El medio líquido sobrante se retiró y se esparcieron aproximadamente 0,4 VCC de células en un círculo de 2,5 cm de diámetro sobre el centro de una placa de Petri de 100 x 15 mm que contenía medio GN6 S/M solidificado con Gelrite 2,5 g/l. Las células se cultivaron en la oscuridad durante 4 horas. Para revestir las partículas biolísticas con ADN, 3 mg de partículas de oro de 1,0 micras de diámetro se lavaron una vez con etanol a 100%, dos veces con agua destilada estéril y se resuspendieron en 50 μ l de agua en un tubo Eppendorf sicolinizado. Por separado, a la suspensión de oro se añadió un total de 5 μ g de ADN plasmídico, espermidina 20 μ l (0,1 M) y cloruro de calcio 50 μ l (2,5 M) y se mezcló en un agitador vorticial. La mezcla se incubó a temperatura ambiente durante 10 minutos, se sedimentó a 10.000 rpm en una microcentrifuga de sobremesa durante 10 segundos, se resuspendió en etanol a 100 % frío 60 μ l y se distribuyeron 8-9 μ l sobre cada microtransportador.

El bombardeo se realizó usando el sistema Biolistic PDS-1000/He™ (Bio-Rad Laboratories, Hercules, CA). Las placas que contenían las células se colocaron en el anaquel central al vacío a 1100 psi y 27 pulgadas de Hg, y se bombardearon siguiendo el manual operativo. Dieciséis horas después del bombardeo, el tejido se transfirió en grupos pequeños a medio GN6 (1H) y se cultivó durante 2-3 semanas a 28 °C en la oscuridad. Las transferencias continuaron cada 2-4 semanas hasta que aparecieron los supuestos aislados transgénicos resultantes de la integración de ADN donante. La identificación, el aislamiento y la regeneración de supuestos acontecimientos de integración de ADN donante generados a través de suministro de ADN mediado por biolística son idénticos al proceso utilizado para los supuestos acontecimientos de integración de ADN donante generados mediante suministro de ADN mediado por Whiskers™ y descrito más adelante.

C. Identificación de aislamiento de supuestos acontecimientos transgénicos de integración dirigida

Una semana después del suministro de ADN, los papeles de filtro se transfirieron a placas de 60 x 20 mm de medio de selección GN6 (1H) (medio N6, 2,4-D 2,0 mg/l, sacarosa 30 g/l, mioinositol 100 mg/l, bialafos 1,0 mg/l de Herbiace (Meiji Seika, Japón), Gelrite 2,5 g/l, pH 5,8). Esas placas de selección se incubaron a 28 °C durante una semana en la oscuridad.

Después de 1 semana de selección en la oscuridad, el tejido se embebió en medio reciente raspando la mitad de las células de cada placa en un tubo que contenía 3,0 ml de medio agarosa GN6 mantenido a 37-38 °C (medio N6, 2,4-D 2,0 mg/l, sacarosa 30 g/l, mioinositol 100 mg/l, agarosa SeaPlaque® 7 g/l, pH 5,8, esterilizado en autoclave durante solo 10 minutos a 121 °C) y bialafos 1 mg/l de Herbiace.

- 5 La mezcla de agarosa/tejido se fraccionó con una espátula y posteriormente 3 ml de mezcla de agarosa/tejido se vertieron uniformemente sobre la superficie de una placa de Petri de 100 x 15 mm que contenía medio GN6 (1H). Este proceso se repitió en ambas mitades de cada placa. Una vez embebido todo el tejido, las placas se sellaron individualmente con Nescofilm® o Parafilm M®, y se cultivaron a 28 °C en condiciones de oscuridad durante hasta 10 semanas. Los aislados supuestamente transformados que crecieron en estas condiciones de selección se retiraron de las placas embebidas y se transfirieron a medio de selección reciente en placas de 60 x 20 mm. Se consideró que un acontecimiento era resistente al herbicida aplicado (bialafos), si después de aproximadamente 2 semanas, el crecimiento sostenido era evidente, y posteriormente se recogió una alícuota de las células en tubos Eppendorf de 2 ml para el análisis del genotipo.

- 15 Un experto en la técnica podría utilizar un gen que codificase cualquier marcador de selección apropiado en el ADN donante y aplicaría condiciones de selección comparables a las de las células vivas. Por ejemplo, un gen marcador seleccionable alternativo, tal como AAD-1, como se describe en el documento WO 2005/107437 A2, podría implementarse como un donante para la selección y recuperación de acontecimientos integrados en células de maíz como se describe en la presente memoria.

Ejemplo 20: Exploración de acontecimientos de integración dirigida mediante genotipificado por PCR

- 20 En este ejemplo, se entiende que el genotipificado por PCR incluye, pero sin limitación, amplificación, por reacción en cadena de la polimerasa (PCR), de ADN genómico procedente de tejido de callo de maíz aislado, que se prevé que contiene ADN donante embebido en el genoma, seguido de clonación y análisis de secuencia convencional de productos de amplificación por PCR. Se han descrito bien métodos de genotipificado por PCR (por ejemplo, Rios, G. *et al.* (2002) *Plant J.* 32:243-253) y pueden aplicarse a ADN genómico procedente de cualquier especie de planta o cualquier tipo de tejido, incluyendo cultivos de células.

- 25 Un experto en la técnica puede imaginar estrategias para el genotipificado por PCR que incluyen (pero sin limitación) amplificación de secuencias específicas en el genoma de la planta, amplificación de secuencias específicas múltiples en el genoma de la planta, amplificación de secuencias no específicas en el genoma de la planta o combinaciones de las mismas. La amplificación puede ir seguida de clonación y secuenciación, como se describe en este ejemplo, o mediante análisis directo de secuencias de productos de amplificación. Un experto en la técnica puede imaginar métodos alternativos para el análisis de los productos de amplificación generados en la presente memoria.

- 30 En una realización descrita en la presente memoria, se emplean cebadores oligonucleotídicos específicos para el direccionamiento génico en amplificaciones por PCR. En otra realización descrita en la presente memoria, se emplean cebadores oligonucleotídicos específicos de secuencias de ADN donante en amplificaciones por PCR. Otra realización incluye una combinación de cebadores oligonucleotídicos que se unen tanto a una secuencia génica diana como a una secuencia de ADN donante. Un experto en la técnica puede imaginar combinaciones adicionales de cebadores y reacciones de amplificación para el investigar el genoma.

A. Extracción de ADN genómico

- 40 Se extrajo ADN genómico (ADNg) de células de maíz aisladas, tolerantes a herbicidas, descritas en el Ejemplo 19 y se utilizó como un molde para experimentos de genotipificado por PCR. Se extrajo ADNg de un volumen celular concentrado (VCC) de aproximadamente 100-300 μ l de callos Hill tolerantes a herbicidas que se aislaron como se describe anteriormente de acuerdo con el protocolo del fabricante, detallado en el kit DNeasy® 96 Plant (QIAGEN Inc., Valencia, CA). El ADN genómico se eluyó en 100 μ l de tampón de elución proporcionado en el kit produciendo una concentración final de 20-200 ng/ μ l y posteriormente se analizó mediante métodos de genotipificado basados en PCR indicados más adelante.

B. Diseño de cebadores para el genotipificado por PCR

- 50 Un experto en la técnica podría usar una diversidad de estrategias para el diseño e implementación de genotipificado basado en PCR. Son factibles cebadores oligonucleotídicos diseñados para hibridar con el gen diana, con secuencias de ADN donantes y/o con combinaciones de ambas. Para diseñar cebadores oligonucleotídicos que puedan hibridar con la diana génica IPP2K en regiones no incluidas por los flancos de homología construidos en las moléculas de ADN donante, los clones plasmídicos que contenían datos de secuencias diana génicas adicionales se caracterizaron mediante secuenciación de ADN. Se realizaron reacciones de secuenciación bicatenarias de los clones plasmídicos como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI). Estas secuencias

corresponden a regiones del gen IPP2K aguas arriba (5'-) y aguas abajo (3'-) de las regiones diana de ZFN y se describen en la Fig. 91 (SEQ ID NO: 141) y Fig. 92 (SEQ ID NO: 142).

5 En los ejemplos presentados en la presente memoria, todos los cebadores oligonucleotídicos se sintetizaron a través de Integrated DNA Technologies, Inc. (Coralville, IA) en condiciones de desalinización convencionales y se diluyeron con agua a una concentración de 100 μ M. Se diseñó el siguiente conjunto de cebadores oligonucleotídicos directo e inverso para hibridar con secuencias de ADN específicas para la diana génica IPP2K que se encuentra fuera de los límites de las secuencias de ADN donante. Estos oligonucleótidos son los siguientes:

5'-TGGACGGAGCGAGAGCCAGAATTCGACGCT G-3' (SEQ ID NO: 153)

5'-GTGCAAGAATGTATTGGGAATCAACGTGAT G-3' (SEQ ID NO: 154)

10 También se diseñó un segundo conjunto de cebadores oligonucleotídicos directo e inverso para hibridar con la secuencia de ADN específica para la diana génica IPP2K diana que se encuentra fuera de los límites de las secuencias de ADN donante, incluso anidada dentro del primer par:

5'-CTGTGGTACCAGTACTAGTACCAGCATC-3' (SEQ ID NO: 155)

5'-TCT TGGATCAAGGCATCAAGC ATTCCAATCT-3' (SEQ ID NO: 156)

15 Adicionalmente, se diseñaron cebadores oligonucleotídicos directo e inverso para hibridar específicamente con ADN donante correspondiente a la región codificante del gen tolerancia a herbicidas:

5'-TGGGTAAGTGGCCTAACTGG-3' (SEQ ID NO: 157)

5'-TGGAAGGCTAGGAACGCTTA-3' (SEQ ID NO: 158)

5'-CCAGTTAGGCCAGTTACCCA-3' (SEQ ID NO: 159)

20 5' TAAGCGTTCCTAGCCTTCCA -3' (SEQ ID NO: 160)

C. ADN donante - Amplificación específica por PCR

25 Se realizaron reacciones primarias de amplificación por PCR usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y que consistían en lo siguiente: 2,5 μ l 10X de tampón Ex TaqPCR™, 40-200 ng de molde de ADN genómico bicatenario, 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones de PCR se realizaron usando un ciclador térmico Peltier, Bio-Rad, de ADN de 96 muestras Engine Tetrad2 (Hercules, CA) con las siguientes condiciones de ciclo: 94 °C, 3 min/ciclo; 94 °C 30 s, 64 °C 30 s, 72 °C 5 min/35 ciclos; 72 °C, 10 min/l ciclo; 4 °C/mantenido.

30 Los productos de amplificación de la reacción primaria por PCR se reamplificaron posteriormente en una reacción secundaria por PCR que comprendía lo siguiente: 2,5 μ l 10X de tampón Ex TaqPCR™, 2 μ l de molde (dilución 1:100 de 1ª reacción PCR en H₂O), 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones PCR se realizaron usando un ciclador térmico Peltier, Bio-Rad, de ADN de 96 muestras Engine Tetrad2 (Hercules, CA) con las siguientes condiciones de ciclo: 95 °C, 1 min/l ciclo; 94 °C 15 s, 61 °C 30 s, 72 °C 30 s/30 ciclos; 72 °C, 1 min/l ciclo; 4 °C/mantenido. Diez μ l de cada producto amplificado se sometieron a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos amplificados se visualizaron con luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN Plus de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). Los productos de PCR que contenían el fragmento esperado se diagnosticaron por la presencia de un fragmento de ADN de 0,317 Kpb, como se muestra en la FIG. 82.

Ejemplo 21: Detección de acontecimientos de integración dirigida

35 Se espera que, de los acontecimientos tolerantes a herbicidas que contenían una molécula de ADN donante integrada que codificaba un casete génico de tolerancia a herbicidas, alguna proporción de dichos acontecimientos fuesen el producto de la integración dirigida de ADN donante en el sitio de la rotura bicatenaria inducida por ZFN. Para diferenciar estos acontecimientos de integración dirigida de los procedentes de integración al azar del casete génico de tolerancia a herbicidas, se utilizó una estrategia de genotipificado basada en PCR que usaba una combinación de cebadores de PCR específicos de genoma y posterior específica genoma y de donante.

A. Amplificación específica de genoma y posterior específica de genoma/donante

50 En esta realización, para las reacciones de PCR primarias se utilizaron cebadores oligonucleotídicos específicos para las regiones de la diana génica IPP2K aguas arriba y aguas abajo de la región de integración donante (por ejemplo, Figs. 92 y 93). Las reacciones de amplificación por PCR primarias se realizaron usando reactivos

proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón y que consistían en lo siguiente: 2,5 μ l 10X de tampón Ex TaqPCR™, 40-200 ng de molde de ADN_g de maíz bicatenario, 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM de cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones de PCR se realizaron usando un ciclador térmico Bio-Rad, de ADN de 96 muestras Entine Tetrad2, Peltre (Hercules, CA) con las siguientes condiciones de ciclo: 94 °C, 3 min/ciclo; 94 °C 30 s, 64 °C 30 s, 72 °C 5 min/35 ciclos; 72 °C, 10 min/ciclo; 4 °C/mantenido.

El producto de reacción de PCR primaria se diluyó posteriormente a 1:100 en H₂O y se usó como ADN molde para dos reacciones de PCR secundarias distintas. En esta realización, las reacciones secundarias utilizan cebadores que se unen en la región genómica de IPP2K y en la molécula donante, dando lugar a un amplicón que abarca el límite de integración entre el genoma y el donante. La primera reacción se enfocó en el límite 5' entre el genoma y el donante. La segunda reacción se enfocó en el límite 3' entre el donante y el genoma. Ambas reacciones consistían en lo siguiente: 2,5 μ l 10X de tampón Ex Taq PCR™, 2 μ l de molde [dilución 1:100 de la reacción PCR 1ª], 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones de PCR se realizaron usando un ciclador térmico Peltier, Bio-Rad, de ADN de 96 muestras Engine Tetrad2 (Hercules, CA) con las siguientes condiciones de ciclo: 94 °C, 3 min/ciclo; 94 °C 30 s, 60 °C 30 s, 72 °C 2 min/35 ciclos; 72 °C, 10 min/ciclo; 4 °C/mantenido. Veinte μ l de cada reacción PCR secundaria se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %.

Los fragmentos amplificados se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN Plus de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). Los productos de la PCR procedentes de la integración diana del donante en el gen IPP2K se diagnosticaron por la presencia de fragmentos de ADN de 1,65 Kpb (límite 5') (FIG. 83) o 1,99 Kpb (límite 3') (FIG. 84). Estos fragmentos se escindieron del gel y se purificaron de acuerdo con las indicaciones del fabricante usando el kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA). Después, los fragmentos purificados se clonaron en el plásmido pCR2.1 usando el kit TOPO TA Cloning® (con el vector pCR®2.1) y células de *E. coli* químicamente competentes TOP One Shot® (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 μ l/ml, se inocularon colonias individuales y se incubaron durante 16 horas a 37 °C con agitación a 200 rpm. Después de la incubación, 1,5 ml de las células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de plásmido aislado se digirieron con 10 unidades de Eco RI (New England Biolabs, Beverly, MA). Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN Plus de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA).

Los clones plasmídicos esperados se diagnosticaron por la presencia de fragmentos de ADN insertados del tamaño apropiado además del vector pCR®2.1 de 3,9 Kbp. Las reacciones de secuenciación bicatenarias de los clones plasmídicos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN de Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI). Las alineaciones de los nucleótidos se realizaron usando la versión 10.1 del vector NTi (Invitrogen Life Technologies, Carlsbad, CA).

El análisis de los datos de secuencia de un acontecimiento de integración dirigido (acontecimiento n.º 073) se realizó de la siguiente manera. Los productos de PCR primaria que abarcaban todo el sitio de integración del genoma se sometieron a amplificación secundaria enfocada en cualquiera de los límites 5' o 3' entre el genoma y el donante. El alineamiento de los fragmentos clonados correspondientes a estos productos de amplificación secundaria con la secuencia genómica de IPP2K de tipo silvestre así como la secuencia esperada de un acontecimiento de integración dirigido claramente indicó que se producía la integración exacta del ADN donante en el sitio diana.

La secuencia de nucleótidos del locus genómico de IPP2K, el límite del genoma/donante, la secuencia de nucleótidos de las regiones donantes correspondientes a los flancos de homología de IPP2K y la secuencia de nucleótidos del casete de tolerancia a herbicidas se preservaron todos en productos de PCR clonados múltiples procedentes de este acontecimiento. Por lo tanto, este acontecimiento representa un genoma en el que aparece la reparación, dirigida por homología, de una rotura bicatenaria mediada por ZFN y la integración dirigida de un ADN donante en una diana génica específica. Se han obtenido acontecimientos transformados adicionales que representan apariciones exclusivas de integración dirigida, demostrando que los métodos enseñados en la presente memoria son reproducibles en callos de maíz. Un experto en la técnica podría aplicar estos métodos a cualquier diana génica en cualquier especie de planta en la que se considere que es deseable la integración dirigida.

B. Amplificación específica de genoma/donante anidada

En esta realización, las reacciones de PCR tanto primaria como después secundaria utilizaron cebadores oligonucleotídicos específicos para las regiones de la diana génica IPP2K aguas arriba o aguas abajo de la región de integración donante (apéndices V y VI) en combinación con cebadores oligonucleotídicos específicos para la secuencia donante. En este ejemplo, las reacciones de amplificación por PCR primarias se realizaron usando reactivos proporcionados por TaKaRa Biotechnology Inc., Seta 3-4-1, Otsu, Shiga, 520-2193, Japón que consistían en lo siguiente: 2,5 μ l 10X de tampón Ex TaqPCR™, 40-200 ng de molde de ADN_g de maíz bicatenario, 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones de PCR se incubaron usando un ciclador térmico Bio-Rad, de ADN de 96 muestras Entine Tetrad2, Peltre (Hercules, CA) con las siguientes condiciones de ciclo: 94 °C, 3 min/ciclo; 94 °C 30 s, 52 °C o 64 °C 30 s, 72 °C 2 min/35 ciclos; 72 °C, 10 min/l ciclo; 4 °C/mantenido.

La reacción de PCR primaria se diluyó después a 1:100 en H₂O y se usó como ADN molde para una reacción de PCR secundaria. En esta realización, las reacciones secundarias también utilizan cebadores que se unen en la región genómica de IPP2K y en la molécula donante, dando lugar a un amplicón que abarca el límite de integración entre el genoma y el donante. Los cebadores específicos usados determinan si la amplificación está enfocada en cualquiera de los límites 5' o 3' entre el genoma y el donante. Los reactivos para estas reacciones consistían en lo siguiente: 2,5 μ l 10X de tampón Ex Taq PCR™, 2 μ l de molde [dilución 1:100 de reacción PCR 1°], 10 μ M de cebador oligonucleotídico directo, 10 μ M de cebador oligonucleotídico inverso, 2 μ l de mezcla de dNTP (2,5 mM cada uno), H₂O 16 μ l, 0,5 μ l (2,5 unidades) de ADN polimerasa Ex Taq™. Las reacciones de PCR se realizaron usando un ciclador térmico Bio-Rad, de ADN de 96 muestras Entine Tetrad2, Peltre (Hercules, CA) con las siguientes condiciones de ciclo: 94 °C, 3 min/ciclo; 94 °C 30 s, 54 °C o 60 °C 30 s, 72 °C 2 min/35 ciclos; 72 °C, 10 min/l ciclo; 4 °C/mantenido. Veinte μ l de cada reacción de PCR secundaria se sometieron a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %.

Los fragmentos amplificados se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN Plus de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA). Los productos de la PCR procedentes de la integración dirigida del donante en el gen IPP2K se diagnosticaron por la presencia de fragmentos de ADN de 1,35 Kpb (límite 5') (FIG. 85) o 1,66 Kpb (límite 3') (FIG. 86). Estos fragmentos se escindieron del gel y se purificaron de acuerdo con las indicaciones del fabricante usando el kit de extracción en gel QIAquick (QIAGEN Inc., Valencia, CA). Después, los fragmentos purificados se clonaron en el plásmido pCR2.1 usando el kit TOPO TA Cloning® (con el vector pCR®2.1) y células de *E. coli* químicamente competentes TOP One Shot® (Invitrogen Life Technologies, Carlsbad, CA) de acuerdo con el protocolo del fabricante.

C. Análisis de secuencia de nucleótidos de productos de genotipificado de productos PCR

En un tubo Falcon de 14 ml (Becton-Dickinson, Franklin Lakes, NJ) que contenía TB 2 ml complementado con kanamicina 50 μ l/ml, se inocularon las colonias individuales descritas en el Ejemplo 21B y se incubaron durante 16 horas a 37°C con agitación a 200 rpm. Después de la incubación, 1,5 ml de las células se transfirieron a un tubo de microcentrífuga Costar de 1,7 ml (Fisher Scientific, Pittsburgh, PA) y se sedimentaron a 16.000 x g durante 1 minuto. El sobrenadante se retiró y el ADN plasmídico se aisló como se ha descrito anteriormente usando el kit para plásmidos NucleoSpin® (BD Biosciences/Clontech/Macherey-Nagel, Palo Alto, CA). Tres μ g de plásmido aislado se digirieron con 10 unidades de Eco RI (New England Biolabs, Beverly, MA). Todas las digestiones de los plásmidos se incubaron durante 1 hora a 37 °C. El ADN restringido se sometió a electroforesis a 100 V durante 1 hora en un gel de agarosa TAE a 1,0 % complementado con bromuro de etidio a 0,5 %. Los fragmentos se visualizaron usando luz UV y el tamaño de los fragmentos se calculó mediante comparación con una escalera de ADN Plus de 1 Kpb (Invitrogen Life Technologies, Carlsbad, CA).

Los clones plasmídicos se diagnosticaron por la presencia de fragmentos de ADN insertados además del vector pCR®2.1 de 3,9 Kpb. Las reacciones de secuenciación bicatenarias de los clones plasmídicos se realizaron como describe el fabricante usando el kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). Las reacciones se purificaron usando cartuchos de filtración en gel Performa DTR (Edge BioSystems, Gaithersburg, MD) como se describe en el protocolo del fabricante. Las reacciones de secuencia se analizaron en un sistema de análisis de ADN de Beckman-Coulter CEQ™ 2000 XL y la caracterización de los nucleótidos se realizó usando la Versión 4.1.4 de Sequencher™ (Gene Codes Corporation, Ann Arbor, MI). Las alineaciones de los nucleótidos se realizaron usando la versión 10.1 del vector NTi (Invitrogen Life Technologies, Carlsbad, CA).

También se obtuvieron datos de las secuencias que abarcaban el límite entre la secuencia genómica de IPP2K aguas arriba (5'-) y el ADN donante procedente de acontecimientos múltiples de integración dirigida, incluyendo datos de secuencias que abarcaban el límite entre el ADN donante y la secuencia genómica de IPP2K aguas arriba (3'-) procedente de acontecimientos múltiples de integración dirigida, así como datos de secuencias incluyendo secuencias límite aguas arriba (5'-) procedentes de un solo acontecimiento de callo transformado (n.º 114). El acontecimiento (n.º 114) de integración dirigida transformado fue el resultado de la integración de un donante autónomo en la diana génica IPP2K.

En estos análisis, las reacciones de amplificación por PCR tanto primaria como secundaria se enfocaron a cualquiera de los límites 5'- o 3'- entre el genoma y el donante. El alineamiento de los fragmentos clonados que correspondían a estos productos de amplificación secundaria con la secuencia genómica IPP2K de tipo silvestre así como la secuencia esperada de un acontecimiento de integración dirigida reveló que se producía la integración del ADN donante en el sitio diana. La secuencia de nucleótidos del locus genómico de IPP2K, el límite del genoma/donante, la secuencia de nucleótidos de las regiones donantes correspondientes a los flancos de homología de IPP2K y la secuencia de nucleótidos del casete de tolerancia a herbicidas se preservaron todos en productos de PCR clonados múltiples procedentes de este acontecimiento.

Por lo tanto, este acontecimiento representa un genoma en el que aparece la reparación, dirigida por homología, de una rotura bicatenaria mediada por ZFN y la integración dirigida de un ADN donante en una diana génica específica. Se han obtenido acontecimientos transformados adicionales que representan apariciones exclusivas de integración dirigida, demostrando que los métodos enseñados en la presente memoria son reproducibles en callos de maíz. Un experto en la técnica podría aplicar estos métodos a cualquier diana génica en cualquier especie de planta en la que se considere que es deseable la integración dirigida.

15 **Ejemplo 22: Regeneración de plantas fértiles, intactas de tejido de callo de maíz**

Callos aislados de células de maíz tolerantes a herbicidas procedentes de cultivo de células Hill pueden regenerarse en plantas de maíz fértiles, intactas. Un experto en la técnica podría regenerar plantas de maíz fértiles, intactas a partir de una variedad de cultivos de células de maíz embriogénicas.

En este ejemplo, la regeneración de callos aislados de células Hill resistentes a bialafos, se inició transfiriendo tejido de callo aislado a un medio de inducción basado en citoquinina, 28 (1H), que contenía sales MS y vitaminas, sacarosa 30,0 g/l, bencilaminopurina 5 mg/l, 2,4-D 0,25 mg/l, bialafos 1 mg/l y Gelrite 2,5 g/l; pH 5,7. Se dejó que las células crecieran con poca luz (13 μ Em-2s-1) durante una semana, seguido de transferencia a condiciones de mayor luz (40 μ Em-2s-1) durante una semana. Después, las células se transfirieron a medio de regeneración, 36 (1H), que es idéntico al medio de inducción excepto que carece de reguladores de crecimiento de plantas. Se cortaron plántulas pequeñas (3-5 cm) con herramientas manuales y se colocaron en tubos de cultivo de vidrio estériles de 150 x 25 mm que contenían medio SHGA (Schenk yHildebrandt basal salts and vitamins, 1972, Can. J. Bot 50:199-204; mioinositol 1 g/l, sacarosa 10 g/l, Gelrite 2,0 g/l, pH 5,8).

Una vez que las plántulas desarrollaron un sistema radicular y de brotes suficientemente grande y diferenciado, se trasplantaron en macetas de 10,16 cm (4 pulgadas) que contenían medio de crecimiento Metro-Mix 360 (Sun Gro Horticulture Canadá Ltd.) se colocaron en un invernadero. Las plántulas se cubrieron total o parcialmente con copas de plástico transparente durante 2-7 días, después se trasplantaron a macetas de 0,02 m³ (5 galones) que contenían una mezcla que consistía en medio de crecimiento Metro-Mix 360 a 95 % y suelo de arcilla/marga a 5 % y crecieron hasta la madurez. Las plantas pueden autopolinizarse o polinizarse en cruzado con una línea endogámica para producir semillas T1 o F1, respectivamente. Un experto en la técnica podría autopolinizar plantas regeneradas o polinizar en cruzado plantas regeneradas con diversos germoplasmas para permitir el cultivo del maíz.

En las publicaciones de solicitud de patente de Estados Unidos US-2003-0232410; US-2005-0026157; US-2005-0064474; US-2005-0208489; y US-2006-0188987; y en la solicitud de patente de Estados Unidos n.º de serie 11/493,423, presentada el 26 de julio de 2006, puede encontrarse información adicional relacionada con escisión dirigida, recombinación dirigida e integración dirigida.

40 **Listado de secuencias**

<110> Cai, Qihua C
Miller, Jeffrey
Urnov, Fyodor
45 Shukla, Vipula K
Petolino, Joseph F
Baker, Lisa W
Garrison, Robbi J
Blue, Ryan C
50 Mitchell, Jon C
Arnold, Nicole L
Worden, Sarah E

<120> Proteínas de dedo de zinc no canónicas optimizadas

<130> 8325-4002.40

<140> PCT/US2007/025455

<141> 2007-12-13

<222> (2)..(2)
 <223> Xaa = cualquier aminoácido (A, K, o T son preferidos)

<220>
 5 <221> CARACT_MISCELÁNEA
 <222> (3)..(3)
 <223> Xaa = cualquier aminoácido (Q, E, o R son preferidos)

<220>
 10 <221> CARACT_MISCELÁNEA
 <222> (6)..(6)
 <223> Xaa = cualquier aminoácido (G es preferido)

<400> 2
 15 His Xaa Xaa Arg Cys Xaa
 1 5

<210> 3
 20 <211> 35
 <212> PRT
 <213> Artificial

<220>
 25 <223> motivo dedo de zinc

<220>
 <221> CARACT_MISCELÁNEA
 <222> (2) .. (3)
 30 <223> Xaa = cualquier aminoácido

<220>
 <221> CARACT_MISCELÁNEA
 <222> (4)..(5)
 35 <223> Xaa = cualquier aminoácido
 Xaa puede estar presente o ausente

<220>
 <221> CARACT_MISCELÁNEA
 <222> (7) .. (18)
 40 <223> Xaa = cualquier aminoácido

<220>
 <221> CARACT_MISCELÁNEA
 <222> (20)..(22)
 45 <223> Xaa = cualquier aminoácido

<220>
 <221> CARACT_MISCELÁNEA
 <222> (23)..(24)
 50 <223> Xaa = cualquier aminoácido
 Xaa puede estar presente o ausente

<220>
 55 <221> CARACT_MISCELÁNEA
 <222> (26)..(26)
 <223> Xaa = cualquier aminoácido

<220>
 60 <221> CARACT_MISCELÁNEA
 <222> (27) .. (35)
 <223> Xaa = cualquier aminoácido
 Xaa puede estar presente o ausente

<400> 3
 65

ES 2 586 210 T3

Cys Xaa Xaa Xaa Xaa Cys Xaa
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa
35

<210> 4
<211> 5
5 <212> PRT
<213> Artificial

<220>
10 <223> enlazador ZC
<400> 4

Gly Leu Arg Gly Ser
1 5

15 <210> 5
<211> 6
<212> PRT
<213> Artificial

20 <220>
<223> enlazador ZC
<400> 5

Gly Gly Leu Arg Gly Ser
25 1 5

<210> 6
<211> 1199
30 <212> DNA
<213> Zea mays

<400> 6

atggagatgg atggggttct gcaagccgcg gatgccaaagg actgggttta caagggggaa
60

ggcgccgcga atctcctcct cagctacacc ggctcgtcgc cctccatggt aagcgctgag
120

ES 2 586 210 T3

taggttctta ctgagcgtgc acgcatcgat cacttgactt taggggctca atgtgtgatt
180

cacgggtgcc gcggcgccat tcgagctcca gatccagtac cgctcgagca agtgataaaa
240

catggagcag ggacgatcac gtggtcactt gaaaattacg tgagggtccgg ggcgacgatg
300

tacggcgcg cgaactctca aacactcaca caacccaaaac cgcttcgtgt tcgtctttgt
360

tccaagcgac tgtgtgagtg tttgagagtt cgccagcgcg acatcgcccg atctgacaaa
420

ttaagctttc gttgcttttc catgattgtg cattttgtga gcatgcactg aatactatga
480

tggatatggt tggaggaagc attattccaa tttgatgata aggggtgttat ttacacttgt
540

tttcagcttg gcaaggtact gcggctcaag aagattctaa aaaacaagtc gcagcgggca
600

ccgagttgta ttgtattctc aagtcatgag caactcctgt ggggccatat ccgagaactg
660

gttgagtcgg tcaaacaaga ttgcttggct caagcctatg cagtgcattg tatgagccaa
720

cacctgggtg ccaatcatgt cgatgggtgg gtatggttca gattcagttc atttatgtcc
780

tgttattgtg attttgattg gtaacatatt gacaacctcg aacttggga tcagattcag
840

ttcacttatg gaagaaattg gagaattggt ataatttata tataatcacc cctactgaaa
900

tagaaataac atggcatcaa tgtgcatgct attggatttt gacacgaata tgctttatcc
960

tatcatatgt tggtaatcc agcaggcagc aggcactact ctttggatcc acgtgacttg
1020

acaaagaaat catgccatct ttccacaatg cagggtccglg tacgtgtttc tagggatttt
1080

ctggagcttg tcgaaaagaa tgttcttagc agccgtcctg ctgggagagt aatgcaagt
1140

tcaattgata aactgctga tgccgctctt ctaatagcag accactcttt atttctggt
1199

5 <210> 7
<211> 50
<212> DNA
<213> Zea mays

10 <400> 7

ES 2 586 210 T3

caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaaga
50

<210> 8
<211> 12
5 <212> DNA
<213> Zea mays

<400> 8

ctgtggggcc at
10 12

<210> 9
<211> 18
15 <212> DNA
<213> Zea mays

<400> 9

cttgaccaac tcagccag
20 18

<210> 10
<211> 59
25 <212> DNA
<213> Zea mays

<400> 10

aagtcatgag caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaag
59

30 <210> 11
<211> 53
<212> DNA
<213> Artificial

35 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 11

40 aagtcatgag caactcctgt ggggccaaga actggttgag tcgggtcaaac aag
53

<210> 12
<211> 12
45 <212> PRT
<213> Artificial

<220>
<223> motivo dedo de zinc

50 <400> 12

His Thr Lys Ile His Leu Arg Gly Ser Gln Leu Val
1 5 10

55 <210> 13
<211> 12
<212> PRT
<213> Artificial

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

 <400> 13
 5
 His Thr Lys Ile Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

 <210> 14
 <211> 12
 10 <212> PRT
 <213> Artificial

 <220>
 <223> motivo dedo de zinc
 15
 <400> 14

 His Thr Lys Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10
 20 <210> 15
 <211> 12
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo dedo de zinc

 <400> 15

 His Thr Lys Ala Cys Leu Arg Gly Ser Gln Leu Val
 30 1 5 10

 <210> 16
 <211> 12
 <212> PRT
 35 <213> Artificial

 <220>
 <223> motivo dedo de zinc
 40 <400> 16

 His Thr Lys Val Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10
 45 <210> 17
 <211> 12
 <212> PRT
 <213> Artificial

 <220>
 50 <223> motivo dedo de zinc

 <400> 17

 His Thr Lys Leu Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10
 55 <210> 18
 <211> 12
 <212> PRT
 <213> Artificial

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 18

His Thr Lys Ser Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 19
 <211> 12
 <212> PRT
 <213> Artificial

15 <220>
 <223> motivo dedo de zinc

<400> 19

His Thr Lys Asn Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

20 <210> 20
 <211> 12
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 20

30 His Thr Lys Lys Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 21
 <211> 12
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 21

His Thr Lys Arg Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

45 <210> 22
 <211> 14
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 22

His Thr Lys Ile Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

55 <210> 23
 <211> 13
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 23

His Thr Lys Ile Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 24
 <211> 14
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

15 <400> 24

His Thr Lys Ile Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

20 <210> 25
 <211> 14
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

25 <400> 25

30 His Thr Lys Ile Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 26
 <211> 15
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

35 <400> 26

His Thr Lys Ile Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

40 <210> 27
 <211> 13
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

45 <400> 27

His Leu Lys Gly Asn Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

50 <210> 28
 <211> 13
 <212> PRT
 <213> Artificial

55 <210> 28
 <211> 13
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 28

His Leu Lys Gly Asn Cys Pro Ala Gly Ser Gln Leu Val
 1 5 10

<210> 29
 <211> 13
 <212> PRT
 <213> Artificial

15 <220>
 <223> motivo dedo de zinc

<400> 29

His Ser Glu Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

20 <210> 30
 <211> 13
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 30

30 His Ser Glu Gly Gly Cys Pro Gly Gly Ser Gln Leu Val
 1 5 10

<210> 31
 <211> 13
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 31

His Ser Ser Ser Asn Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

45 <210> 32
 <211> 13
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 32

His Ser Ser Ser Asn Cys Thr Ile Gly Ser Gln Leu Val
 1 5 10

55 <210> 33
 <211> 15
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 33

His Thr Lys Ile Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 34
 <211> 16
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

15 <400> 34

His Thr Lys Ile Gly Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

20 <210> 35
 <211> 14
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 35

30 His Thr Lys Ile Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 36
 <211> 15
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 36

His Thr Lys Ile Gly Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

45 <210> 37
 <211> 16
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 37

His Thr Lys Ile Gly Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

55 <210> 38
 <211> 13
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 38

His Thr Lys Arg Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 39
 <211> 14
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

15 <400> 39

His Thr Lys Arg Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

20 <210> 40
 <211> 15
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 40

30 His Thr Lys Arg Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 41
 <211> 13
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 41

His Thr Lys Arg Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

45 <210> 42
 <211> 14
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 42

His Thr Lys Arg Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

55 <210> 43
 <211> 15
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 43

His Thr Lys Arg Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 44
 <211> 16
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

15 <400> 44

His Thr Lys Arg Gly Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

20 <210> 45
 <211> 14
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 45

30 His Thr Lys Arg Gly Gly Cys Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 46
 <211> 15
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 46

His Thr Lys Arg Gly Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

45 <210> 47
 <211> 16
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 47

His Thr Lys Arg Gly Gly Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

55 <210> 48
 <211> 14
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 48

His Leu Lys Gly Asn Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 49
 <211> 15
 <212> PRT
 <213> Artificial

15 <220>
 <223> motivo dedo de zinc

<400> 49

His Leu Lys Gly Asn Cys Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

20 <210> 50
 <211> 16
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 50

30 His Leu Lys Gly Asn Cys Gly Gly Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 51
 <211> 13
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 51

His Lys Glu Arg Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

45 <210> 52
 <211> 13
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 52

His Thr Arg Arg Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

55 <210> 53
 <211> 13
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
<223> motivo dedo de zinc

5 <400> 53

His Ala Gln Arg Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 54
10 <211> 14
<212> PRT
<213> Artificial

<220>
15 <223> motivo dedo de zinc

<400> 54

His Lys Lys Phe Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

20 <210> 55
<211> 14
<212> PRT
<213> Artificial

25 <220>
<223> motivo dedo de zinc

<400> 55

30 His Lys Lys His Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

<210> 56
35 <211> 14
<212> PRT
<213> Artificial

<220>
40 <223> motivo dedo de zinc

<400> 56

His Lys Lys Tyr Thr Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10

45 <210> 57
<211> 14
<212> PRT
<213> Artificial

50 <220>
<223> motivo dedo de zinc

<400> 57

His Lys Lys Tyr Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
55 1 5 10

<210> 58
60 <211> 14
<212> PRT
<213> Artificial

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 58

His Lys Gln Tyr Tyr Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 59
 <211> 14
 <212> PRT
 <213> Artificial

15 <220>
 <223> motivo dedo de zinc

<400> 59

His Leu Leu Lys Lys Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

20 <210> 60
 <211> 14
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 60

30 His Gln Lys Phe Pro Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 61
 <211> 14
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 61

His Gln Lys Lys Leu Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

45 <210> 62
 <211> 14
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 62

His Gln Ile Arg Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10

55 <210> 63
 <211> 15
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
 <223> motivo dedo de zinc

5 <400> 63

His Ile Lys Arg Gln Ser Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 64
 <211> 15
 <212> PRT
 <213> Artificial

<220>
 <223> motivo dedo de zinc

15 <400> 64

His Ile Arg Arg Tyr Thr Cys Gly Leu Arg Gly Ser Gln Leu Val
 20 1 5 10 15

<210> 65
 <211> 15
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

30 <400> 65

His Ile Ser Ser Lys Lys Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 66
 <211> 15
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

40 <400> 66

His Lys Ile Gln Lys Ala Cys Gly Leu Arg Gly Ser Gln Leu Val
 45 1 5 10 15

<210> 67
 <211> 15
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

55 <400> 67

His Lys Arg Ile Tyr Thr Cys Gly Leu Arg Gly Ser Gln Leu Val
 1 5 10 15

<210> 68
 <221> 15
 <212> PRT
 <213> Artificial

60

ES 2 586 210 T3

<220>
<223> motivo dedo de zinc

5 <400> 68

His Leu Lys Gly Gln Asn Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 69
<211> 15
<212> PRT
<213> Artificial

15 <220>
<223> motivo dedo de zinc

<400> 69

His Leu Lys Lys Asp Gly Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

20 <210> 70
<211> 15
<212> PRT
<213> Artificial

25 <220>
<223> motivo dedo de zinc

<400> 70

30 His Leu Lys Tyr Thr Pro Cys Gly Leu Arg Gly Ser Gln Leu Val
1 5 10 15

<210> 71
<211> 12
<212> PRT
<213> Artificial

35 <220>
<223> motivo dedo de zinc

40 <400> 71

His Thr Lys Arg Cys Gly Arg Gly Ser Gln Leu Val
1 5 10

45 <210> 72
<211> 14
<212> PRT
<213> Artificial

50 <220>
<223> motivo dedo de zinc

<400> 72

His Thr Lys Ile Gly Cys Gly Gly Arg Gly Ser Gln Leu Val
1 5 10

55 <210> 73
<211> 13
<212> PRT
<213> Artificial

60

ES 2 586 210 T3

<220>
<223> motivo dedo de zinc

5 <400> 73

His Leu Lys Gly Asn Cys Gly Arg Gly Ser Gln Leu Val
1 5 10

<210> 74
10 <211> 13
<212> PRT
<213> Artificial

<220>
15 <223> motivo dedo de zinc

<400> 74

His Leu Lys Gly Asn Cys Gly Gly Gly Ser Gln Leu Val
1 5 10

20 <210> 75
<211> 11
<212> PRT
<213> Artificial

25 <220>
<223> motivo dedo de zinc

<400> 75

30 His Ile Arg Thr Cys Thr Gly Ser Gln Lys Pro
1 5 10

<210> 76
35 <211> 12
<212> PRT
<213> Artificial

<220>
40 <223> motivo dedo de zinc

<400> 76

His Ile Arg Thr Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

45 <210> 77
<211> 12
<212> PRT
<213> Artificial

50 <220>
<223> motivo dedo de zinc

<400> 77

His Ile Arg Thr Gly Cys Thr Gly Ser Gln Lys Pro
55 1 5 10

<210> 78
<211> 13
<212> PRT
60 <213> Artificial

ES 2 586 210 T3

<220>
<223> motivo dedo de zinc

5 <400> 78

His Ile Arg Thr Gly Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

<210> 79
10 <211> 11
<212> PRT
<213> Artificial

<220>
15 <223> motivo dedo de zinc

<400> 79

His Ile Arg Arg Cys Thr Gly Ser Gln Lys Pro
1 5 10

20 <210> 80
<211> 12
<212> PRT
<213> Artificial

25 <220>
<223> motivo dedo de zinc

<400> 80

30 His Ile Arg Arg Gly Cys Thr Gly Ser Gln Lys Pro
1 5 10

<210> 81
35 <211> 11
<212> PRT
<213> Artificial

<220>
40 <223> motivo dedo de zinc

<400> 81

His Thr Lys Ile His Thr Gly Ser Gln Lys Pro
1 5 10

45 <210> 82
<211> 11
<212> PRT
<213> Artificial

50 <220>
<223> motivo dedo de zinc

<400> 82

His Thr Lys Ile Cys Thr Gly Ser Gln Lys Pro
55 1 5 10

<210> 83
60 <211> 11
<212> PRT
<213> Artificial

ES 2 586 210 T3

<220>
<223> motivo dedo de zinc

5 <400> 83

His Thr Lys Arg Cys Thr Gly Ser Gln Lys Pro
1 5 10

<210> 84
10 <211> 11
<212> PRT
<213> Artificial

<220>
15 <223> motivo dedo de zinc

<400> 84

His Ala Gln Arg Cys Thr Gly Ser Gln Lys Pro
1 5 10

20 <210> 85
<211> 12
<212> PRT
<213> Artificial

25 <220>
<223> motivo dedo de zinc

<400> 85

30 His Thr Lys Ile Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

<210> 86
35 <211> 12
<212> PRT
<213> Artificial

<220>
40 <223> motivo dedo de zinc

<400> 86

His Thr Lys Arg Cys Gly Thr Gly Ser Gln Lys Pro
1 5 10

45 <210> 87
<211> 12
<212> PRT
<213> Artificial

50 <220>
<223> motivo dedo de zinc

<400> 87

His Ala Gln Arg Cys Gly Thr Gly Ser Gln Lys Pro
55 1 5 10

<210> 88
60 <211> 12
<212> PRT
<213> Artificial

<220>
 <223> motivo dedo de zinc

5 <400> 88

His Thr Lys Ile His Leu Arg Gly Ser Gln Leu Val
 1 5 10

<210> 89
 <211> 7
 <212> PRT
 <213> Artificial

15 <220>
 <223> motivo dedo de zinc

<400> 89

His Ala Gln Arg Cys Gly Gly
 1 5

20 <210> 90
 <211> 8
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo dedo de zinc

<400> 90

30 His Ala Gln Arg Cys Gly Gly Gly
 1 5

<210> 91
 <211> 8
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo dedo de zinc

<400> 91

His Thr Lys Ile Cys Gly Gly Gly
 1 5

45 <210> 92
 <211> 8
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo dedo de zinc

<400> 92

His Thr Lys Arg Cys Gly Gly Gly
 1 5

55 <210> 93
 <211> 6
 <212> PRT
 <213> Artificial

60

<220>
 <223> motivo dedo de zinc
 5 <400> 93
 His Ala Gln Arg Cys Gly
 1 5
 <210> 94
 10 <211> 7
 <212> PRT
 <213> Artificial
 <220>
 15 <223> motivo dedo de zinc
 <400> 94
 Cys Cys His Thr Lys Ile His
 1 5
 20 <210> 95
 <211> 7
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo dedo de zinc
 <400> 95
 30 Cys Cys His Thr Lys Ile His
 1 5
 <210> 96
 <211> 7
 35 <212> PRT
 <213> Artificial
 <220>
 <223> motivo dedo de zinc
 40 <400> 96
 Cys Cys His Thr Lys Ile Cys
 1 5
 45 <210> 97
 <211> 10
 <212> PRT
 <213> Artificial
 50 <220>
 <223> motivo dedo de zinc
 <400> 97
 Cys Cys His Thr Lys Arg Cys Gly Gly Gly
 55 1 5 10
 <210> 98
 <211> 8
 <212> PRT
 60 <213> Artificial

<220>
 <223> motivo dedo de zinc
 5 <400> 98
 Cys Cys His Ala Gln Arg Cys Gly
 1 5
 <210> 99
 10 <211> 8
 <212> PRT
 <213> Artificial
 <220>
 15 <223> motivo dedo de zinc
 <400> 99
 Cys Cys His Ile Arg Thr Gly Cys
 1 5
 20 <210> 100
 <211> 7
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo dedo de zinc
 <400> 100
 30 Cys Cys His Thr Lys Ile His
 1 5
 <210> 101
 <211> 8
 35 <212> PRT
 <213> Artificial
 <220>
 <223> motivo dedo de zinc
 40 <400> 101
 Cys Cys His Ile Arg Thr Gly Cys
 1 5
 45 <210> 102
 <211> 7
 <212> PRT
 <213> Artificial
 50 <220>
 <223> motivo dedo de zinc
 <400> 102
 Cys Cys His Thr Lys Ile His
 55 1 5
 <210> 103
 <211> 9
 <212> PRT
 60 <213> Artificial

<220>
 <223> motivo dedo de zinc

5 <400> 103
 Cys Cys His Ala Gln Arg Cys Gly Gly
 1 5

<210> 104
 10 <211> 29
 <212> DNA
 <213> Artificial

<220>
 15 <223> cebador
 <400> 104
 atggagatgg atggggttct gcaagccgc
 20 29

<210> 105
 <211> 7
 <212> PRT
 <213> Artificial

25 <220>
 <223> motivo de unión dedo de zinc
 <400> 105

30 Asp Arg Ser Ala Leu Ser Arg
 1 5

<210> 106
 <211> 7
 35 <212> PRT
 <213> Artificial

<220>
 40 <223> motivo de unión dedo de zinc
 <400> 106
 Arg Asn Asp Asp Arg Lys Lys
 1 5

45 <210> 107
 <211> 7
 <212> PRT
 <213> Artificial

50 <220>
 <223> motivo de unión dedo de zinc
 <400> 107
 Arg Ser Asp Asn Leu Ser Thr
 55 1 5

<210> 108
 <211> 7
 <212> PRT
 60 <213> Artificial

<220>
 <223> motivo de unión dedo de zinc

5 <400> 108

His Ser His Ala Arg Ile Lys
 1 5

<210> 109
 <211> 7
 <212> PRT
 <213> Artificial

<220>
 <223> motivo de unión dedo de zinc

15 <400> 109

Arg Ser Asp Val Leu Ser Glu
 1 5

20 <210> 110
 <211> 7
 <212> PRT
 <213> Artificial

<220>
 <223> motivo de unión dedo de zinc

25 <400> 110

Gln Ser Gly Asn Leu Ala Arg
 1 5

<210> 111
 <211> 7
 <212> PRT
 <213> Artificial

35 <220>
 <223> motivo de unión dedo de zinc

40 <400> 111

Arg Ser Asp Asn Leu Ala Arg
 1 5

<210> 112
 <211> 7
 <212> PRT
 <213> Artificial

45 <220>
 <223> motivo de unión dedo de zinc

50 <400> 112

Thr Ser Gly Ser Leu Thr Arg
 1 5

<210> 113
 <211> 7
 <212> PRT
 <213> Artificial

60

<220>
 <223> motivo de unión dedo de zinc
 5 <400> 113
 Thr Ser Gly Asn Leu Thr Arg
 1 5
 <210> 114
 10 <211> 7
 <212> PRT
 <213> Artificial
 <220>
 15 <223> motivo de unión dedo de zinc
 <400> 114
 Arg Ser Asp His Leu Ser Glu
 1 5
 20 <210> 115
 <211> 7
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo de unión dedo de zinc
 <400> 115
 30 Gln Ser Ala Thr Arg Lys Lys
 1 5
 <210> 116
 <211> 7
 35 <212> PRT
 <213> Artificial
 <220>
 <223> motivo de unión dedo de zinc
 40 <400> 116
 Glu Arg Gly Thr Leu Ala Arg
 1 5
 45 <210> 117
 <211> 7
 <212> PRT
 <213> Artificial
 50 <220>
 <223> motivo de unión dedo de zinc
 <400> 117
 Arg Ser Asp Ala Leu Thr Gln
 55 1 5
 <210> 118
 <211> 7
 <212> PRT
 60 <213> Artificial

<220>
 <223> motivo de unión dedo de zinc
 5 <400> 118
 Arg Ser Asp Ser Leu Ser Ala
 1 5
 <210> 119
 10 <211> 7
 <212> PRT
 <213> Artificial
 <220>
 15 <223> motivo de unión dedo de zinc
 <400> 119
 Arg Ser Ala Ala Leu Ala Arg
 1 5
 20 <210> 120
 <211> 7
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo de unión dedo de zinc
 <400> 120
 30 Arg Ser Asp Asn Leu Ser Glu
 1 5
 <210> 121
 <211> 7
 35 <212> PRT
 <213> Artificial
 <220>
 <223> motivo de unión dedo de zinc
 40 <400> 121
 Ala Ser Lys Thr Arg Thr Asn
 1 5
 45 <210> 122
 <211> 7
 <212> PRT
 <213> Artificial
 50 <220>
 <223> motivo de unión dedo de zinc
 <400> 122
 Asp Arg Ser His Leu Ala Arg
 55 1 5
 <210> 123
 <211> 7
 <212> PRT
 60 <213> Artificial

<220>
 <223> motivo de unión dedo de zinc
 5 <400> 123
 Arg Ser Asp His Leu Ser Thr
 1 5
 <210> 124
 10 <211> 7
 <212> PRT
 <213> Artificial
 <220>
 15 <223> motivo de unión dedo de zinc
 <400> 124
 Gln Ser Gly Ser Leu Thr Arg
 1 5
 20 <210> 125
 <211> 7
 <212> PRT
 <213> Artificial
 25 <220>
 <223> motivo de unión dedo de zinc
 <400> 125
 30 Gln Asn His His Arg Ile Asn
 1 5
 <210> 126
 <211> 7
 35 <212> PRT
 <213> Artificial
 <220>
 <223> motivo de unión dedo de zinc
 40 <400> 126
 Thr Gly Ser Asn Leu Thr Arg
 1 5
 45 <210> 127
 <211> 7
 <212> PRT
 <213> Artificial
 50 <220>
 <223> motivo de unión dedo de zinc
 <400> 127
 Asp Arg Ser Ala Leu Ala Arg
 55 1 5
 <210> 128
 <211> 18
 <212> DNA
 60 <213> Artificial

ES 2 586 210 T3

<220>
<223> secuencia diana dedo de zinc IPP2K

5 <400> 128
gaactggttg agtcggtc
18

10 <210> 129
<211> 18
<212> DNA
<213> Artificial

15 <220>
<223> secuencia diana dedo de zinc IPP2K
<400> 129
gaactggttg agtcggtc
18

20 <210> 130
<211> 12
<212> DNA
<213> Artificial

25 <220>
<223> secuencia diana dedo de zinc IPP2K
<400> 130

30 atggccccac ag
12

35 <210> 131
<211> 15
<212> DNA
<213> Artificial

40 <220>
<223> secuencia diana dedo de zinc IPP2K
<400> 131
ggcaccagg tgttg
15

45 <210> 132
<211> 18
<212> DNA
<213> Artificial

50 <220>
<223> secuencia diana dedo de zinc IPP2K
<400> 132
gtcgaatggg gggatatg
18

55 <210> 133
<211> 18
<212> PRT
60 <213> Artificial

ES 2 586 210 T3

<220>

<223> NLS derivado de maíz op-2

<400> 133

5

Arg Lys Arg Lys Glu Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Tyr
1 5 10 15

Arg Lys

<210> 134

<211> 18

10

<212> PRT

<213> Artificial

<220>

<223> secuencia 2A de virus Thosea asigna

15

<400> 134

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15

Gly Pro

20

<210> 135

<211> 30

<212> DNA

<213> Artificial

25

<220>

<223> cebador

<400> 135

ggaagcatta ttccaatttg atgataatgg

30

30

<210> 136

<211> 29

<212> DNA

35

<213> Artificial

<220>

<223> cebador

40

<400> 136

cccaagtgtc gaggttgca atatgttac

29

45

<210> 137

<211> 26

<212> DNA

<213> Artificial

50

<220>

<223> cebador

<400> 137

ssscaccaag ttgtattgcc ttctca

26

ES 2 586 210 T3

<210> 138
<211> 26
<212> DNA
5 <213> Artificial

<220>
<223> cebador

10 <400> 138

sssataggct tgagccaagc aatcct
26

<210> 139
15 <211> 855
<212> DNA
<213> Artificial

<220>
20 <223> secuencia flanqueante homología 3' posición 1

<400> 139

gcgcccgcta gatagcagat gcagattgct tgcttctctg gtttgatctt tggagtcacc
60

atctctgctt ggttcgtgtg cctcagtgtc tgacagcagc agatcctcga tggagatgga
120

tggggttctg caagccgagg atgccaagga ctgggtttac aagggggaag gcgcccgaa
180

tctcactctc agctacaccg gcacgtcgcc ctccatggta agcgtgagt aggttcttac
240

tgagtgtgca cgcacgatc acttgacttt aggggctcaa tgtgtgattc acgggtgccg
300

25

ES 2 586 210 T3

ccattcgagc tccagatcca gtatcgctcg agcaagtgat aaaacatgga gcagggacga
360

tcacgtggtc acttgaaaat tatgtgaggt cggggcgac gatgtacggc gcggcgaact
420

ctcaaacact cacacagcca aaaccgcttc gtgttcgtct ttgttccaag cgaccgtgtg
480

gtgtgtttgt agtagttcgc cggcgccgca catcgctgcc ccggatctga caaattaagc
540

tttcgttgct tttccacgat tgtgcatttt ctgagcatgc actgaatact atgatggata
600

tgtttggagg aagcattatt ccaatttgat gataatgggtg ttatttacac ttgttttcag
660

cttggcaagg tactgcggtc caagaagatt ctaaaaaaca agttgcagcg ggcaccaagt
720

tgtattgcct tctcaagtca tgagcaacte ctgtggggcc atatccaga actggttgag
780

toggtcaaac aagattgctt ggctcaagcc tatgcagtgc atgttatgag ccaacacctg
840

ggtgccaata ctagt
855

<210> 140

<211> 845

5 <212> DNA

<213> Artificial

<220>

10 <223> secuencia flanqueante homología 3' posición 2

<400> 140

actagtcatg tcgatgggtg ggtatggttc agattcagtt catttatgtc ctgttattgt
60

gattttgatt ggtaacatat tgacaacctc gacacttggg atcagattca gttcacttat
120

ggaagaaatt ggagaattgt gataatttat ctataatcac ccctactgaa atagaaataa
180

catgacatca atgtgcatgc tattggattt tgacacgaat atgctttatt ctatcatatg
240

ES 2 586 210 T3

ttggtaattc cagcaggcag caggcactac tctttggatc cacgtgactt gacaaagaaa
300

tcatgccatc tttccacaat gcaggtcctg gtacgtggtt ctagggattt tctggagctt
360

gtcgaaaaga atgttcttag cagccgtcct gctgggagag taaatgcaag ttcaattgat
420

aacactgctg atgccgctct tctaatagca gaccactctt tattttctgg tacgtactct
480

atccctcttc ttaccataat ctgaatcttg ttaaggttta aaatatacga ttgattaagt
540

aaaatccaga gctctattca tatctcacgc actgatggtt tgatgaaacg cttgcagcaa
600

gacggttgcc tgttatttct atttgcatta gacaaacagt cacctttggt tataaaggtc
660

tttgaatttg cagttcttat aggtttaagt ttgcaactgt tacttacaac agcccaatgg
720

gtagcatcaa gattggtttt ttcagtgatt cataacttaa ctcttggtta aaccgctaga
780

acatggttgg tgtcttaaaa tgcaactggt cctgaggccg taacctgaaa tcattgtacg
840

tcgac
845

- <210> 141
- <211> 484
- <212> DNA
- <213> Artificial

5

- <220>
- <223> secuencia genómica curso arriba 5'-IPP2K de las regiones a las que apunta ZFN

10

- <400> 141

tggacggagc gagagccaga attcgacgct ggcggcggcg cgtcgccaat acgcagcgcg
60

gatgtggagc cacatgcaaa cgtgtgtccg cccgcgtggc gtccactctc cctccacggt
120

tcggcgtcct cgtcgccttc ctgggaaatc tccagctact gccactgcc ccttcccttc
180

ES 2 586 210 T3

agtccctttc cccgggctgt ggtaccagta ctagtaccag catctcttca ggctccacca
240

agcgcagaca ccgcagcagc ggcagcagca cgatccgggtg accccccgcc gcgtccagcc
300

tgctcctccg gtgatcgccg gactggcggg gtaggaacca gcggagcgca gcccgctcc
360

ttccgctggg aagagtgacg cccgcccgct cctcccttcg ctcgcttct tgctcttccg
420

attctggcgt accagtctca ccgcggcttg gggatttgat gcggagctag ttaaccagca
480

gagc
484

<210> 142

<211> 729

5 <212> DNA

<213> Artificial

<220>

10 <223> secuencia genómica curso abajo 3'-IPP2K de las regiones a las que apunta ZFN

<400> 142

attgtttttt tcagtgattc ataacttaac tcttggttaa accgctagaa catgggtggg
60

gtcttaaaat gcaactggtc ctgaggccgt aacctgaaat cattgtactt ttctctcatt
120

tctttagata tttccaaaac tctacattag atgatttatg ttgcttact tagtctttct
180

taatctcagg caatcctaag ggtagcagct gcatagctgt agagataaag gtactttgca
240

agcttctctt tttattctta tttttcattt cttatgtata tttctcctca accatttgac
300

ttcttttcgg catgctctac cttgcaggcc aatgtgggt ttctgccatc atcagaatat
360

atatcagaag ataatactat caagaaacta gtaacgagat ataagatgca tcagcacctc
420

aaattttatc agggtgaggt gtgtagattg gaatgcttga tgccttgatc caagataaaa
480

ES 2 586 210 T3

ttccactctc ttttgcgcac ttaaaaaaca tccatcgatg atacaaactt gatcaaaata
540

ccttaaggct tggtatttac ggcactggtg taatattata ccgtctcttg ctttttgaca
600

tcaggttgat tcccaatata ttcttgcaca catttcagat atcgaagact agtgagtaca
660

atcctcttga tctatcttct gggtcaaaag agagaatatg catggccatc aagtcctttt
720

tctcaactc
729

<210> 143
<211> 42
5 <212> DNA
<213> Artificial

<220>
<223> cebador

10 <400> 143

gcggccgcgt ctcaccgcgg cttggggatt ggatacggag ct
42

15 <210> 144
<211> 38
<212> DNA
<213> Artificial

20 <220>
<223> cebador

<400> 144

actagtgata tggceccaca ggagttgctc atgacttg
25 38

<210> 145
<211> 40
30 <212> DNA
<213> Artificial

<220>
<223> cebador

35 <400> 145

actagtccag aactggttga gtcggtcaaa caagattgct
40

40 <210> 146
<211> 32
<212> DNA
<213> Artificial

45 <220>
<223> cebador

<400> 146

gtcgaccttg atgctaccca ttgggctggt gt
 32
 <210> 147
 <211> 30
 5 <212> DNA
 <213> Artificial
 <220>
 <223> cebador
 10 <400> 147
 ggggcccgcta gatagcagat gcagattgct
 30
 15 <210> 148
 <211> 29
 <212> DNA
 <213> Artificial
 20 <220>
 <223> cebador
 <400> 148
 actagtattg gcacccaggt gttggctca
 25 29
 <210> 149
 <211> 38
 <212> DNA
 30 <213> Artificial
 <220>
 <223> cebador
 35 <400> 149
 actagtcatg tcgatggtgg ggtatggttc agattcag
 38
 40 <210> 150
 <211> 37
 <212> DNA
 <213> Artificial
 <220>
 45 <223> cebador
 <400> 150
 gtcgacgtac aatgatttca ggttacggcc tcaggac
 37
 50 <210> 151
 <211> 41
 <212> DNA
 <213> Artificial
 55 <220>
 <223> cebador
 <400> 151
 60

ES 2 586 210 T3

actagttaac tgacctcact cgaggtcatt catatgcttg a
41

5 <210> 152
<211> 29
<212> DNA
<213> Artificial

10 <220>
<223> cebador

<400> 152

actagtgtga attcagcact taaagatct
29

15 <210> 153
<211> 101
<212> DNA
<213> Artificial

20 <220>
<223> cebador

<400> 153

actagtggcg gcggagaggg cagaggaagt cttctaacat gcggtgacgt ggaggagaat
60

25 cccggcccta ggatggcttc tccggagagg agaccagttg a
101

30 <210> 154
<211> 35
<212> DNA
<213> Artificial

35 <220>
<223> cebador

<400> 154

actagtatgc atgtgaattc agcacttaaa gatct
35

40 <210> 155
<211> 28
<212> DNA
<213> Artificial

45 <220>
<223> cebador

<400> 155

ctgtgggtacc agtactagta ccagcatc
50 28

55 <210> 156
<211> 31
<212> DNA
<213> Artificial

<220>
<223> cebador

<400> 156

5 tcttgatca aggcataag cattccaatc t
31

<210> 157
<211> 20
10 <212> DNA
<213> Artificial

<220>
<223> cebador

15 <400> 157

tgggtaactg gcctaactgg
20

20 <210> 158
<211> 20
<212> DNA
<213> Artificial

25 <220>
<223> cebador

<400> 158

tggaaggcta ggaacgctta
30 20

<210> 159
<211> 20
35 <212> DNA
<213> Artificial

<220>
<223> cebador

40 <400> 159

ccagttaggc cagttacca
20

<210> 160
<211> 20
45 <212> DNA
<213> Artificial

<220>
50 <223> cebador

<400> 160

taagcgttcc tagccttcca
55 20

<210> 161
<211> 31
60 <212> DNA
<213> Artificial

<220>
 <223> cebador

 <400> 161
 5 cttggcaagg tactgoggct caagaagatt c
 31

 <210> 162
 <211> 26
 10 <212> DNA
 <213> Artificial

 <220>
 <223> cebador
 15 <400> 162

 atgaagaaag acagggaatg aaggac
 26

 20 <210> 163
 <211> 32
 <212> DNA
 <213> Artificial

 25 <220>
 <223> cebador

 <400> 163

 atgaagaaag acagggaatg aaggaccgcc ac
 30 32

 <210> 164
 <211> 28
 <212> DNA
 35 <213> Artificial

 <220>
 <223> cebador
 40 <400> 164

 catggagggc gacgagccgg tgtagctg
 28

 <210> 165
 <211> 27
 <212> DNA
 <213> Artificial

 <220>
 <223> cebador
 50 <400> 165

 atcgacatga ttggcaccca ggtgttg
 27
 55 <210> 166
 <211> 31
 <212> DNA
 <213> Artificial
 60

<220>
<223> cebador

<400> 166

5 tttcgacaag ctccagaaaa tccctagaaa c
 31

<210> 167
<211> 28
10 <212> DNA
 <213> Artificial

<220>
<223> cebador

15 <400> 167

 acaagctcca gaaaatccct agaaacac
 28

20 <210> 168
 <211> 32
 <212> DNA
 <213> Artificial

25 <220>
 <223> cebador

 <400> 168

 ttcgacaagc tccagaaaaat ccctagaaac ac
 32

30 <210> 169
 <211> 29
 <212> DNA
35 <213> Artificial

<220>
<223> cebador

40 <400> 169

 tgctaagaac attcttttcg acaagctcc
 29

45 <210> 170
 <211> 32
 <212> DNA
 <213> Artificial

<220>

50 <223> cebador

 <400> 170

 gaacattctt ttcgacaagc tccagaaaaat cc
 32

55 <210> 171
 <211> 30
 <212> DNA
 <213> Artificial

60 <220>

ES 2 586 210 T3

<223> cebador
<400> 171
tggacggagc gagagccaga attcgacgct
5 30
<210> 172
<211> 31
<212> DNA
10 <213> Artificial
<220>
<223> cebador
15 <400> 172
gtgcaagaat gtattgggaa tcaacctgat g
 31
<210> 173
20 <211> 59
<212> DNA
<213> Zea mays
<400> 173
25 aagtcatgag caactcctgt ggggccatat cccagaactg gttgagtcgg tcaaacaag
 59
<210> 174
<211> 53
30 <212> DNA
<213> Artificial
<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
35 <400> 174
aagtcatgag caactcctgt ggggccaaaga actgggtgag tcggtcaaac aag
 53
40 <210> 175
<211> 55
<212> DNA
<213> Artificial
45 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
<400> 175
aagtcatgag caactcctgt ggggccataa gaactgggtg agtcggtcaa acaag
50 55
<210> 176
<211> 54
<212> DNA
55 <213> Artificial
<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
60 <400> 176

ES 2 586 210 T3

aagtcattgag caactcctgt ggggccatag aactgggtga gtcgggtcaa caag
54

<210> 177
5 <211> 56
<212> DNA
<213> Artificial

<220>
10 <223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 177

aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag
56

15 <210> 178
<211> 57
<212> DNA
<213> Artificial

20 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 178

25 aagtcattgag caactcctgt ggggccatac cagaactgggt tgagtcgggtc aaacaag
57

<210> 179
30 <211> 52
<212> DNA
<213> Artificial

<220>
35 <223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 179

aagtcattgag caactcctgt ggggccagaa ctgggtgagt cgggtcaaaca ag
52

40 <210> 180
<211> 57
<212> DNA
<213> Artificial

45 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 180

aagtcattgag caactcctgt ggggccatat cagaactgggt tgagtcgggtc aaacaag
50 57

<210> 181
55 <211> 54
<212> DNA
<213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

60 <400> 181

ES 2 586 210 T3

aagtcacgag caactcctgt ggggcatag aactgggtga gtcgggtcaaa caag
54

<210> 182
<211> 54
5 <212> DNA
<213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

10 <400> 182

aagtcacgag caactcctgt ggggcatag aactgggtga gtcgggtcaaa caag
54

15 <210> 183
<211> 57
<212> DNA
<213> Artificial

20 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 183

aagtcacgag caactcctgg tggggcata cagaactggt tgagtcgggtc aaacaag
25 57

<210> 184
<211> 54
<212> DNA
30 <213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

35 <400> 184

aagtcacgag caactcctgt ggggcatag aactgggtga gtcgggtcaaa caag
54

40 <210> 185
<211> 53
<212> DNA
<213> Artificial

<220>
45 <223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 185

aagcatgagc aactcctgtg gggcataga actgggtgag tcgggtcaaac aag
50 53

<210> 186
<211> 56
<212> DNA
55 <213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 186
60

ES 2 586 210 T3

aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag
56

<210> 187
<211> 57
5 <212> DNA
<213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

10 <400> 187

aagtcattgag caactcctgt ggggccatat cagaactgggt tgagtcgggtc aaacaag
57

15 <210> 188
<211> 54
<212> DNA
<213> Artificial

20 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 188

aagtcattgag caactcctgt ggggccatag aactgggttga gtcgggtcaaa caag
25 54

<210> 189
<211> 54
<212> DNA
30 <213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

35 <400> 189

aagtcattgag caactcctgt ggggccacag aactgggttga gtcgggtcaaa caag
54

40 <210> 190
<211> 54
<212> DNA
<213> Artificial

<220>
45 <223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 190

aagtcattgag caactcctgt ggggccatag aactgggttga gtcgggtcaaa caag
50 54

<210> 191
<211> 57
<212> DNA
<213> Artificial

55 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 191
60

ES 2 586 210 T3

aagtcattgag caactcctgt ggggccatac cagaactggt tgagtcggtc aaacaag
57

<210> 192
<211> 54
5 <212> DNA
<213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
10
<400> 192

aagtcattgag caactcctgt ggggccatag aactgggtga gtcgggtcaaa caag
54

15 <210> 193
<211> 55
<212> DNA
<213> Artificial

20 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
<400> 193

aagtcattgag caactcctgt ggggccataa gaactgggtg agtcgggtcaa acaag
25 55

<210> 194
<211> 56
<212> DNA
30 <213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
35 <400> 194

aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag
56

40 <210> 195
<211> 56
<212> DNA
<213> Artificial

45 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
<400> 195

aagtcattgag caactcctgt ggggccatat agaactgggt gagtcgggtca aacaag
50 56

<210> 196
<211> 56
<212> DNA
55 <213> Artificial

<220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN
60 <400> 196

ES 2 586 210 T3

aagtcattgag caactcctgt ggggccatac agaactgggt gagtcgggtca aacaag
56

5 <210> 197
<211> 57
<212> DNA
<213> Artificial

10 <220>
<223> Secuencia de gen IPP2K de maíz con mutación mediada por ZFN

<400> 197

aagtcattgag caactcctgt ggggccatac cagaactgggt tgagtcgggtc aacaag
57

15 <210> 198
<211> 821
<212> DNA
<213> Artificial

20 <220>
<223> secuencia flanqueante homología 5' posición 1

<400> 198

25 gcggccgct ctcaccgagg cttggggatt ggatacggag ctagttaacc agcagagcta
60

ES 2 586 210 T3

gatagcagac gcagategct tgcttctctg gtttgatttt tggagtcacc atttctgttt
120

ggttcgtgtg cctcagtgtc tgacagcagc agatcctcga tggagatgga tggggttctg
180

caagccgcgg atgccaagga ctgggtttac aagggggaag gcgccgcgaa tctcatcctc
240

agctacaccg gcacgtcgcc ctccatggta agcgcctgagt aggttcttac tgagtgtgca
300

cgcatcgatc acttgacttt aggggctcaa tgtgtgattc acgggtgccg ccattcgagc
360

tccagatcca gtatcgctcg agcaagtgat aaaacatgga gcagggacga tcacgtggtc
420

acttgaaaat tatgtgaggt ccggggcgac gatgtacggc gcggcgaact ctcaaacact
480

cacacagcca aaaccgcttc gtgttcgtct ttgttccaag cgaccgtgtg gtgtgtttgt
540

agtagttcgc cggcgccgca catcgtcgcc ccggatctga caaattaagc tttcgttgct
600

tttccacgat tgtgcatttt ctgagcatgc actgaatact atgatggata tgtttgagg
660

aagcattatt ccaatttgat gataatgggt ttatttacac ttgttttcag cttggcaagg
720

tactgcccgt caagaagatt ctaaaaaaca agttgcagcg ggcaccaagt tgtattgcct
780

tctcaagtca tgagcaactc ctgtggggcc atatactag t
821

<210> 199

<211> 821

5 <212> DNA

<213> Artificial

<220>

10 <223> secuencia flanqueante homología 3' posición 1

<400> 199

actagtccag aactgggtga gtcgggtcaaa caagattgct tggctcaagc ctatgcagtg
60

ES 2 586 210 T3

catggtatga gccaacacct gggtgccaat catgtcgatg gtggggtatg gttcagatc
120

agttcattta tgcctgtta ttgtgatttt gattggtaac atattgaaa cctcgacct
180

tgggatcaga ttcagttcac ttatggaaga aattggagaa ttgtgataat ttatctataa
240

tcacccctac tgaaatagaa ataacatgac atcaatgtgc atgctattgg attttgacac
300

gaatatgctt tattctatca tatgttgga attccagcag gcagcaggca ctactctttg
360

gatccacgtg acttgacaaa gaaatcatgc catctttcca caatgcaggt ccgtgtacgt
420

gtttctaggg attttctgga gcttgtcgaa aagaatgttc ttagcagccg tctgtctggg
480

agagtaaatg caagttcaat tgataacact gctgatgccg ctcttctaag agcagaccac
540

tctttatctt ctggtacgta ctctatccct cttcttacca taatctgaat cttggttaagg
600

tttaaaatat atgattgatt aagtaaaatc cagagctcta ttcatatctc acgcaactgat
660

gttttgatga aacgcttgca gcaagacggg tgccctgttat ttctatttgc attagacaaa
720

cagtcacctt tgtttataaa ggtctttgaa tttgcagttc ttataggttt aagtttgcaa
780

ctgttactta caacagccca atgggtagca tcaagggtcga c
821

REIVINDICACIONES

1. Una proteína de dedo de zinc que comprende un dedo de zinc no canónico (no-C2H2), en donde el dedo de zinc no canónico tiene una parte helicoidal implicada en la unión al ADN y en donde al menos un dedo de zinc comprende la secuencia Cys-(X^A)₂₋₄-Cys-(X^B)₁₂-His-(X^C)₃₋₅-Cys-(X^D)₁₋₁₀ (SEQ ID NO:3), donde X^A, X^B, X^C y X^D puede ser cualquier aminoácido e His-(X^C)₃₋₅-Cys(X^D)₁₋₁₀ comprende una de las secuencias como se muestra en cualquiera de las SEQ ID NO:14-80, 83-87 y 89-92 y adicionalmente en donde la proteína de dedo de zinc se modifica genéticamente para unirse a una secuencia diana.
2. A Una proteína de dedo de zinc que comprende una pluralidad de dedos de zinc, en donde al menos uno de los dedos de zinc comprende un dedo de zinc CCHC de acuerdo con la reivindicación 1.
3. La proteína de dedo de zinc de cualquiera de las reivindicaciones 1 a 2, en donde la proteína de dedo de zinc comprende cualquiera de las secuencias mostradas en la Tabla 8 y se modifica genéticamente para unirse a una secuencia diana en un gen IPP2-K.
4. Una proteína de fusión que comprende una proteína de dedo de zinc de cualquiera de las reivindicaciones 1 a 3 y uno o más dominios funcionales.
5. La proteína de fusión de la reivindicación 4, en donde el dominio funcional comprende un semidominio de escisión.
6. La proteína de fusión de la reivindicación 5, que adicionalmente comprende un enlazador ZC entre el semidominio de escisión y la proteína de dedo de zinc.
7. Un polinucleótido que codifica al menos una proteína de dedo de zinc o proteína de fusión de acuerdo con cualquiera de las reivindicaciones 1 a 6.
8. Una célula de planta que comprende al menos una proteína de dedo de zinc, al menos una proteína de fusión o al menos un polinucleótido de acuerdo con cualquiera de las reivindicaciones 1 a 7.
9. La célula de planta de la reivindicación 8, en donde la célula es una semilla.
10. La célula de planta de la reivindicación 8 a 9, en donde IPP2-K está parcial o completamente inactivado y los niveles de ácido fítico en la semilla están reducidos.
11. Un método para la escisión dirigida de la cromatina celular en una célula de planta, comprendiendo el método expresar, en la célula, un par de proteínas de fusión de acuerdo con la reivindicación 5 o al menos un polinucleótido que codifica el par de proteínas de fusión; en donde:
- (a) las secuencias diana de las proteínas de fusión están a una distancia de 10 nucleótidos entre sí; y
- (b) las proteínas de fusión dimerizan y escinden ADN localizado entre las secuencias diana.
12. Un método de recombinación genética dirigida en una célula de planta, comprendiendo el método:
- (a) expresar, en la célula hospedadora, un par de proteínas de fusión de acuerdo con la reivindicación 5 o al menos un polinucleótido que codifica el par de proteínas de fusión, en donde las secuencias diana de las proteínas de fusión están presentes en un locus diana del hospedador seleccionado; y
- (b) identificar una célula hospedadora recombinante que exhibe una alteración de secuencia en el locus diana del hospedador y
- (c) opcionalmente introducir un polinucleótido exógeno en la célula hospedadora, en donde, cuando está presente el polinucleótido exógeno está integrado en el genoma de la célula de la planta hospedadora.
13. El método de la reivindicación 12, en donde la alteración de secuencia es una mutación seleccionada del grupo que consiste en una delección de material genético, una inserción de material genético, una sustitución de material genético y cualquiera de sus combinaciones.
14. Un método para reducir el nivel de ácido fítico en una semilla de planta o para fabricar fósforo más disponible desde el punto de vista metabólico en una semilla de planta, comprendiendo el método inactivar o alterar un gen IPP2-K de acuerdo con la reivindicación 12.

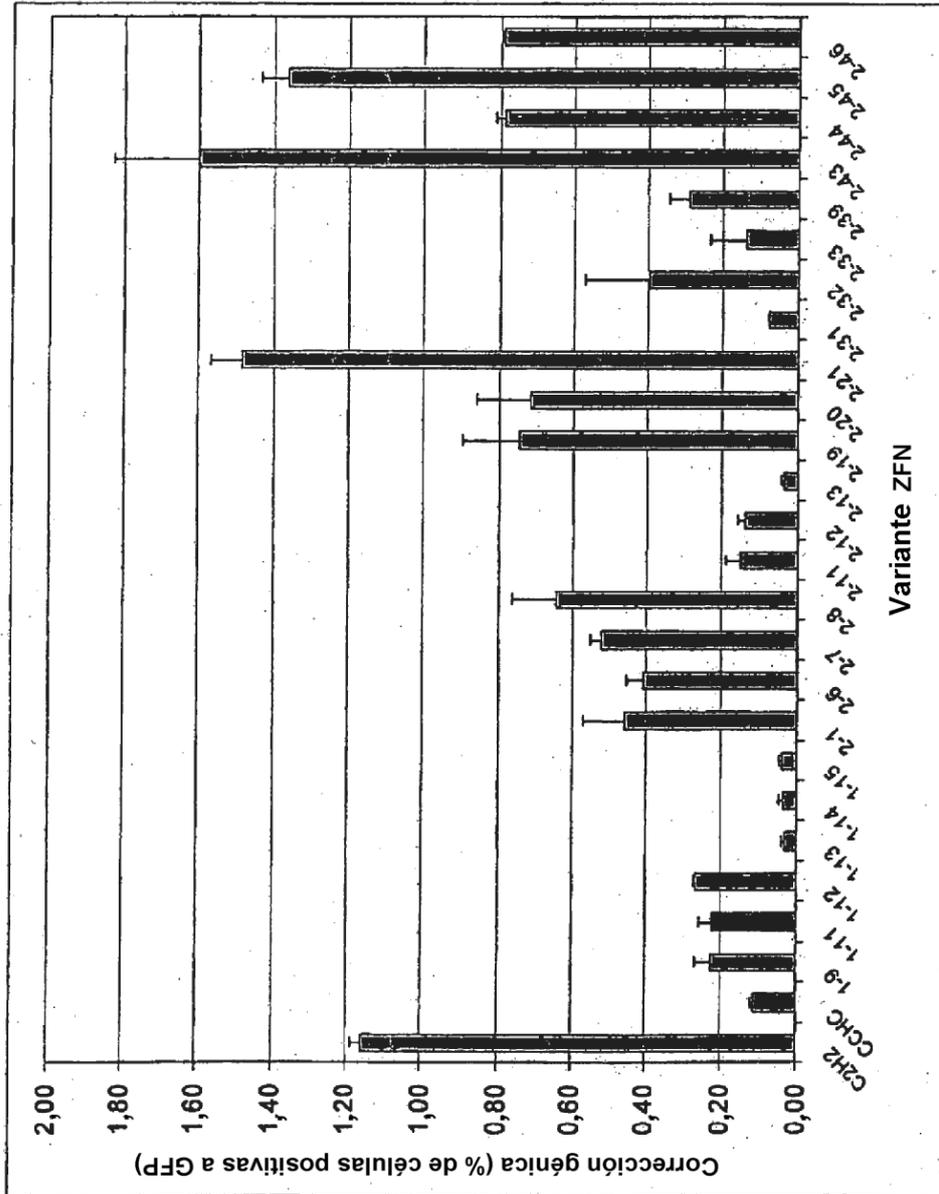


FIG. 1

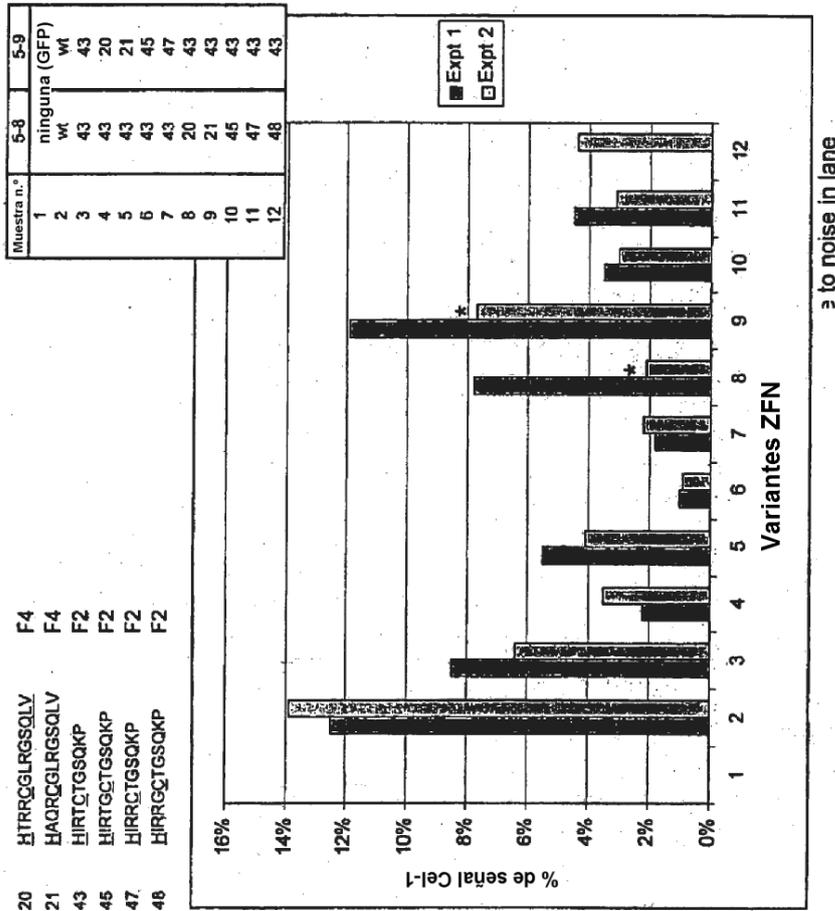


FIG. 2

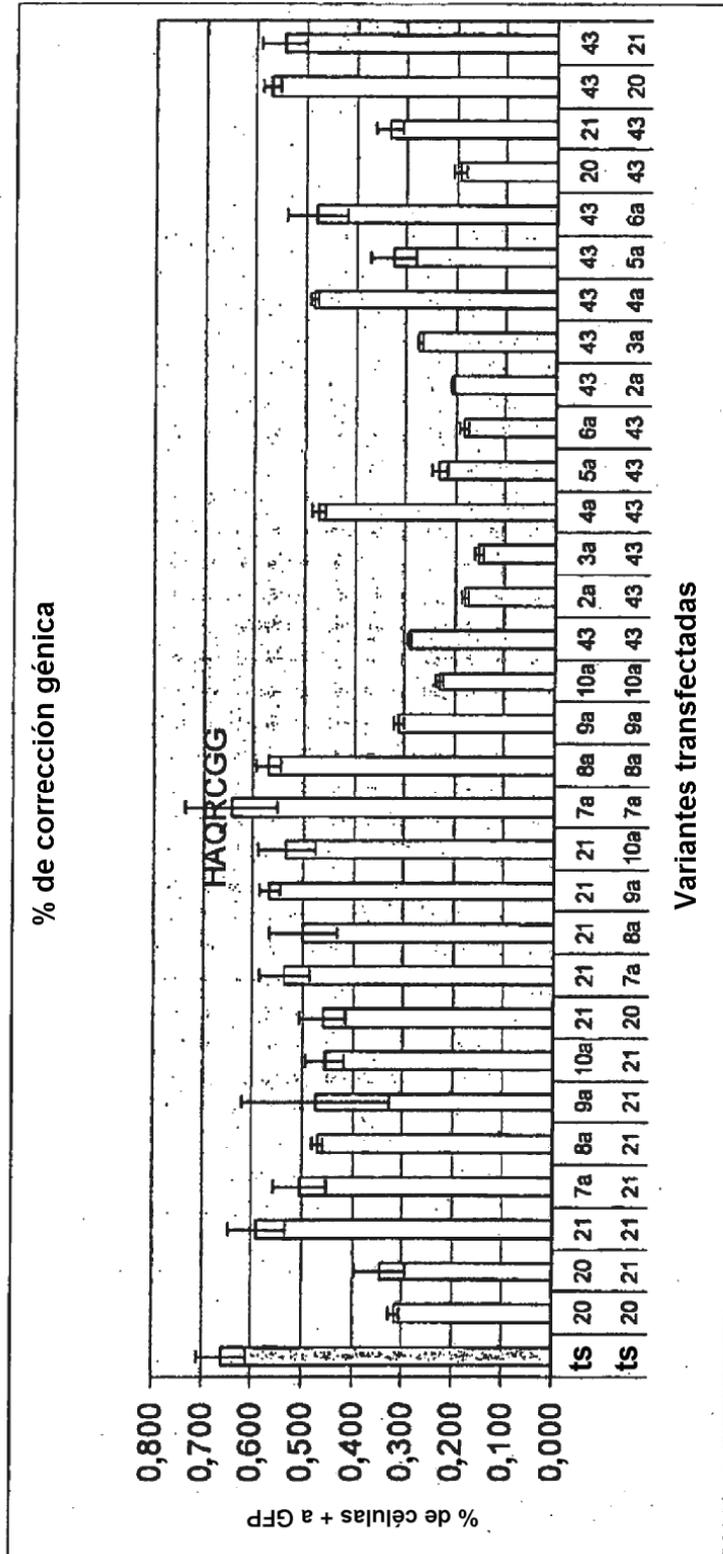


FIG. 3

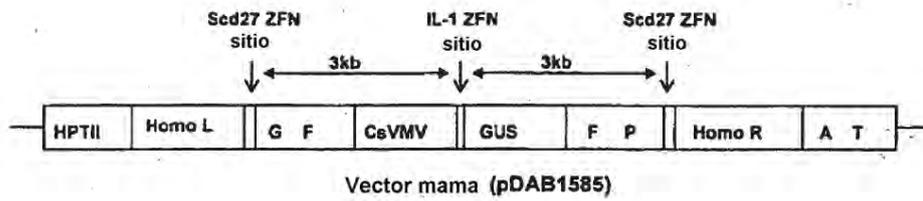


Figura 4

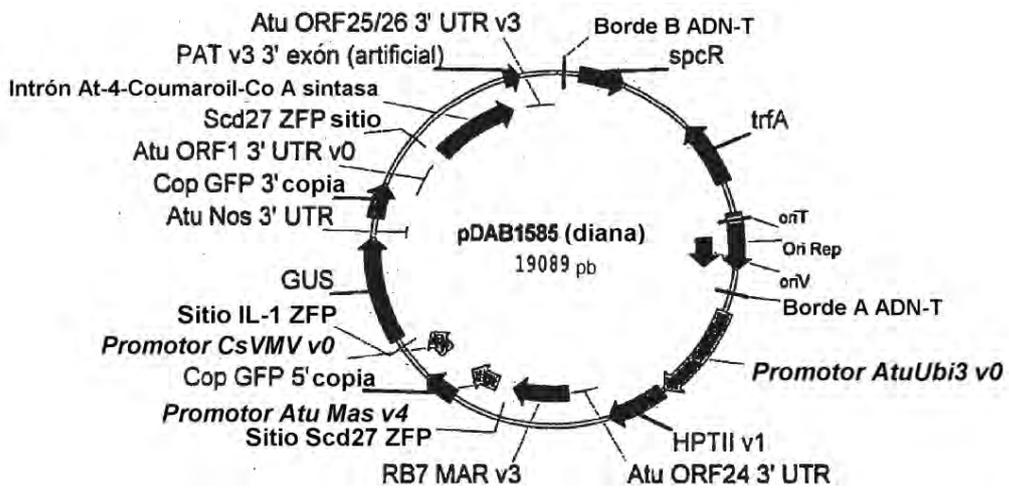


Figura 5: Representación esquemática del plásmido pDAB1585-Vector diana para el tabaco

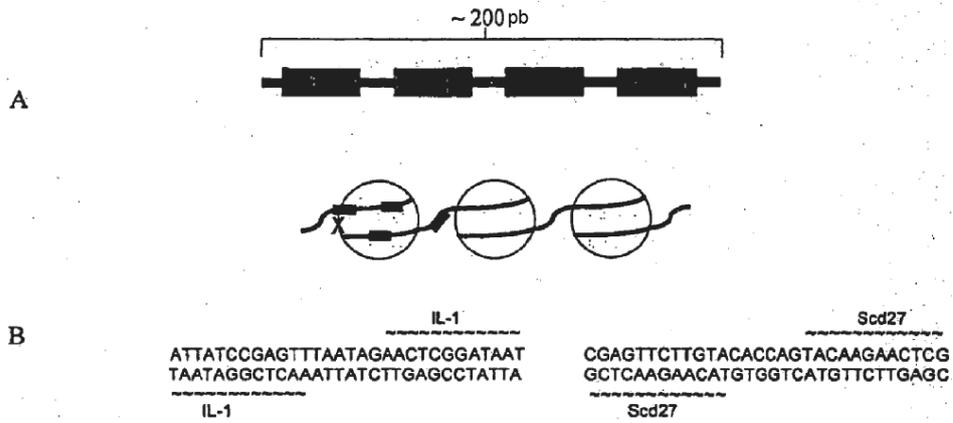


Figura 6: Secuencias de unión a nucleasas de dedo de zinc y el sitio diana diseñado para el tabaco

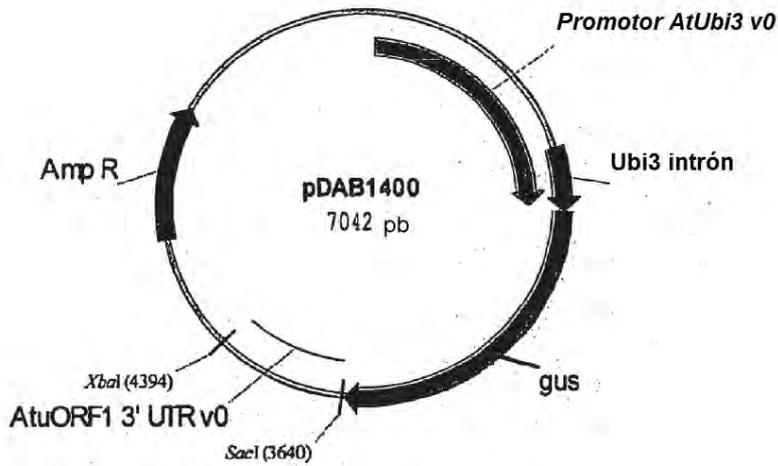


Figura 7: Representación esquemática del plásmido pDAB1400

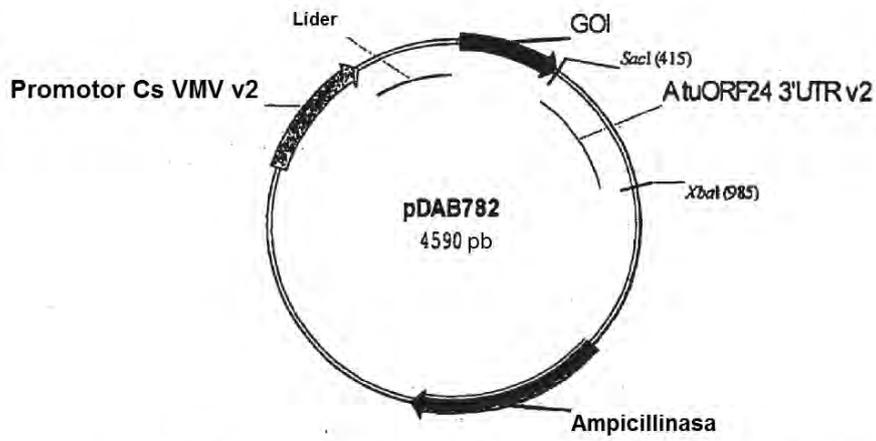


Figura 8: Representación esquemática del plásmido pDAB782

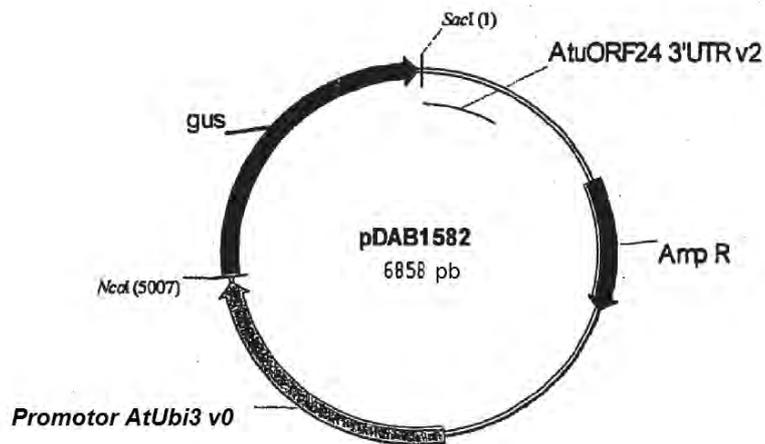


Figura 9: Representación esquemática del plásmido pDAB1582

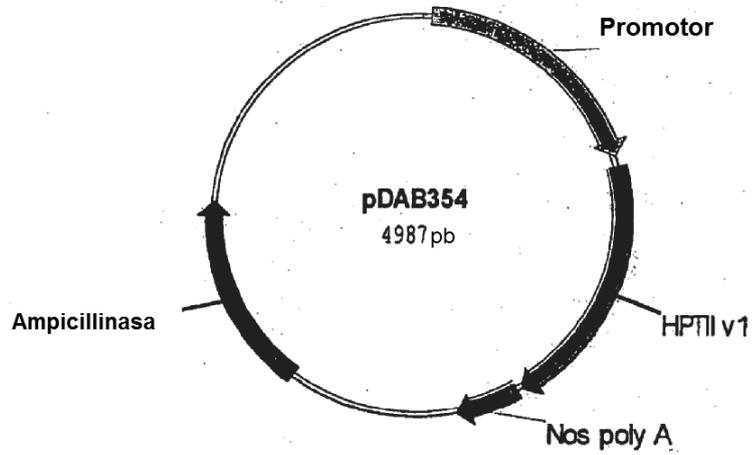


Figura 10: Representación esquemática del plásmido pDAB354

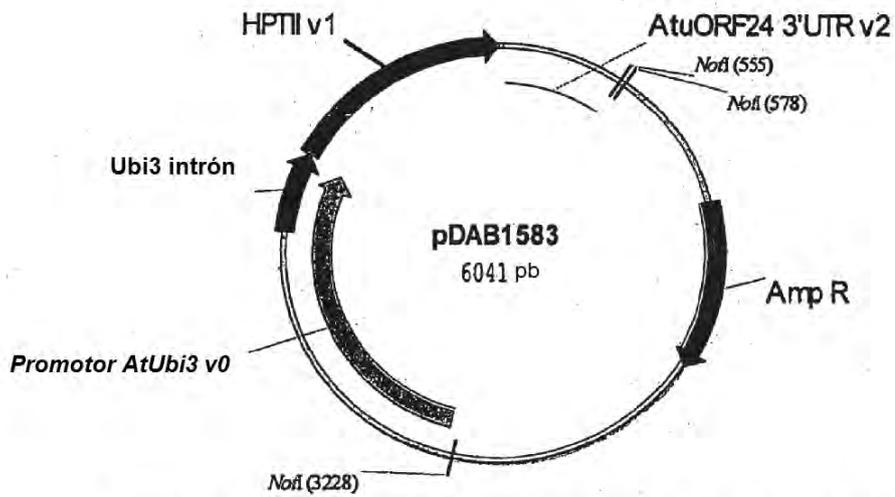


Figura 11: Representación esquemática del plásmido pDAB1583

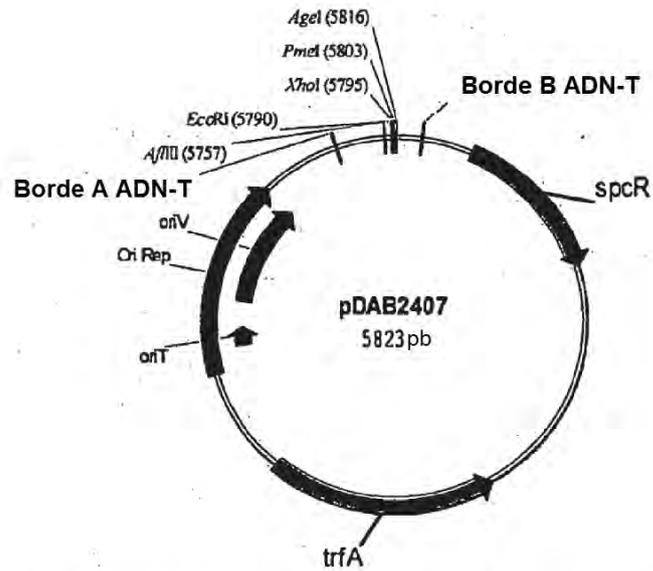


Figura 12: Representación esquemática del plásmido pDAB2407

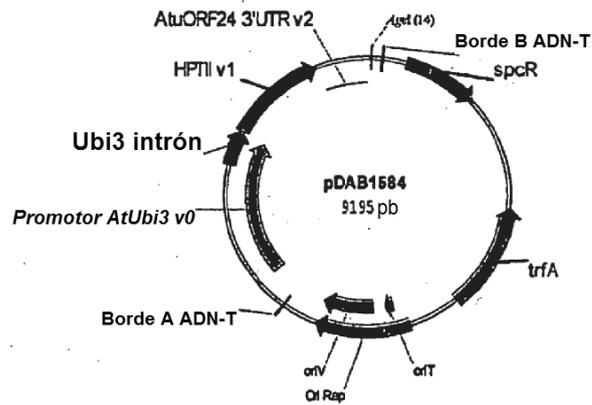


Figura 13: Representación esquemática del plásmido pDAB1584

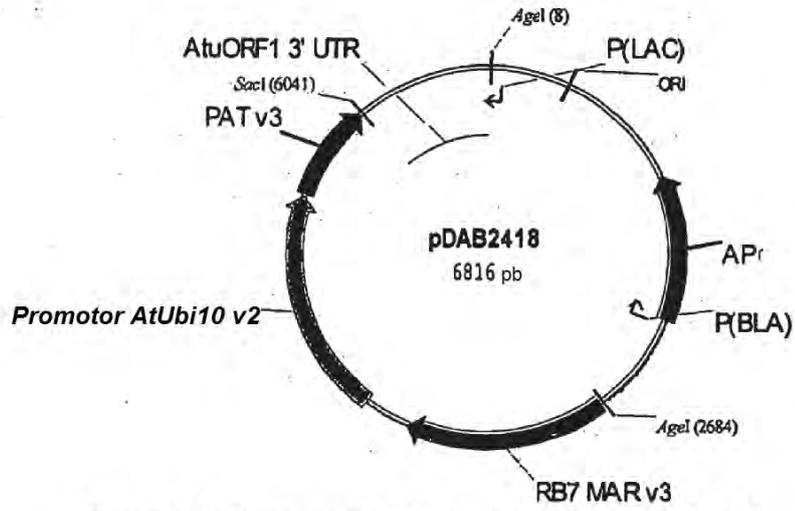


Figura 14: Representación esquemática del plásmido pDAB2418

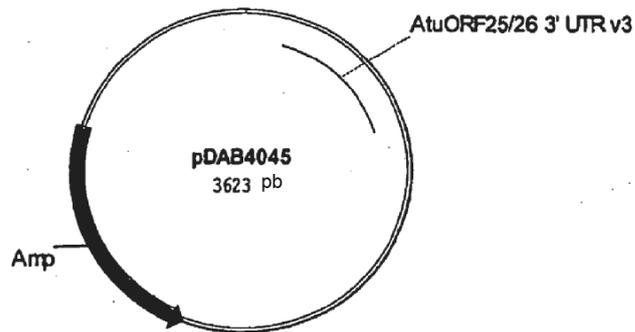


Figura 15: Representación esquemática del plásmido pDAB4045

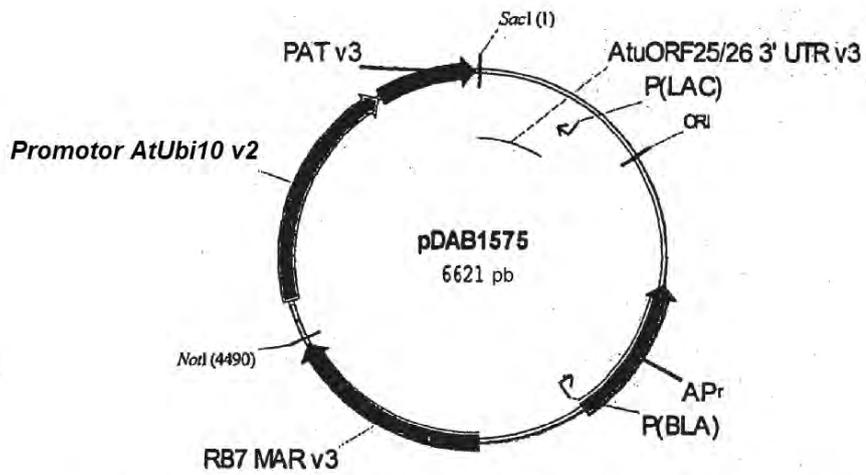


Figura 16: Representación esquemática del plásmido pDAB1575

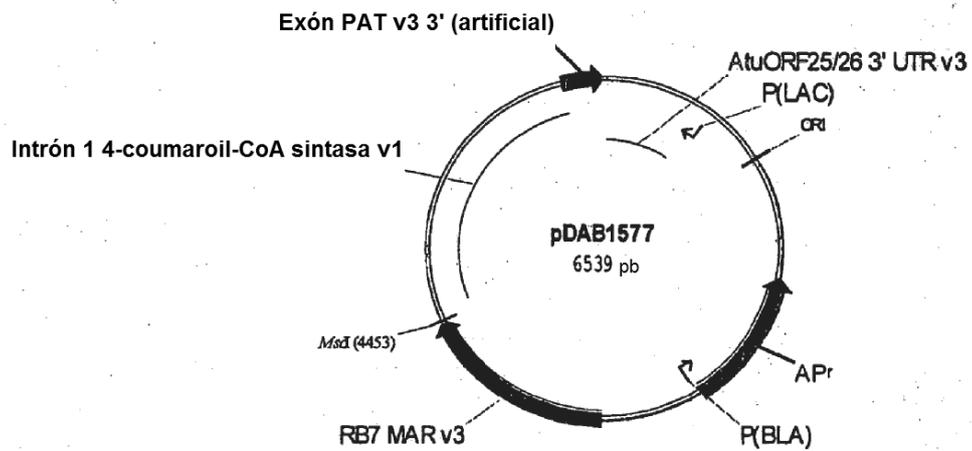


Figura 17: Representación esquemática del plásmido pDAB1577

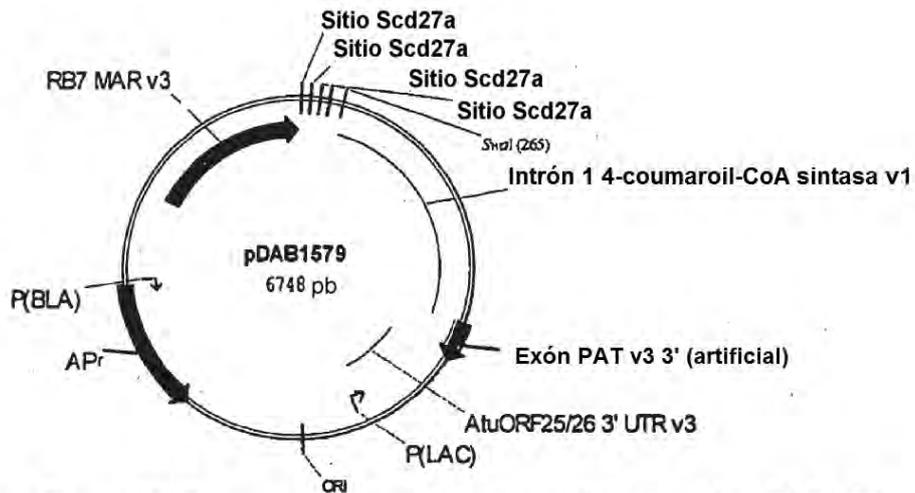


Figura 18: Representación esquemática del plásmido pDAB1579

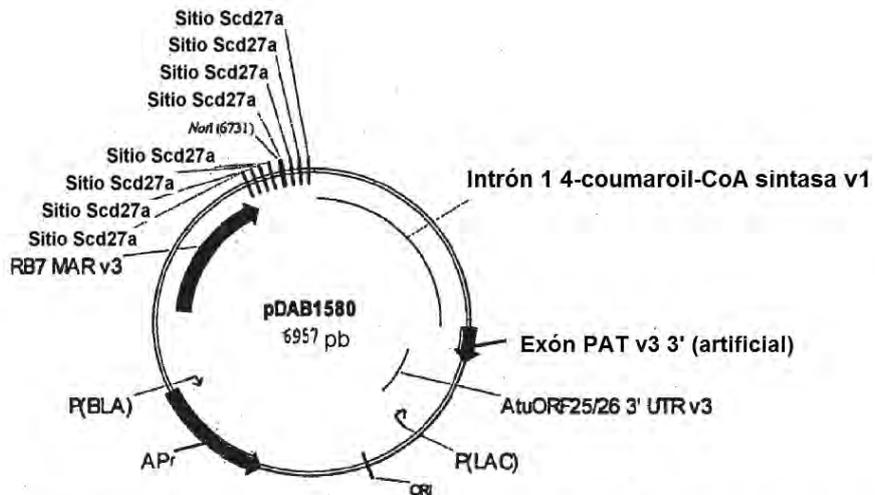


Figura 19: Representación esquemática del plásmido pDAB1580

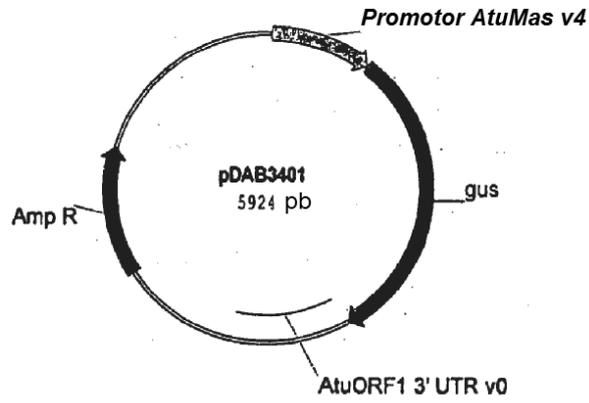


Figura 20: Representación esquemática del plásmido pDAB3401

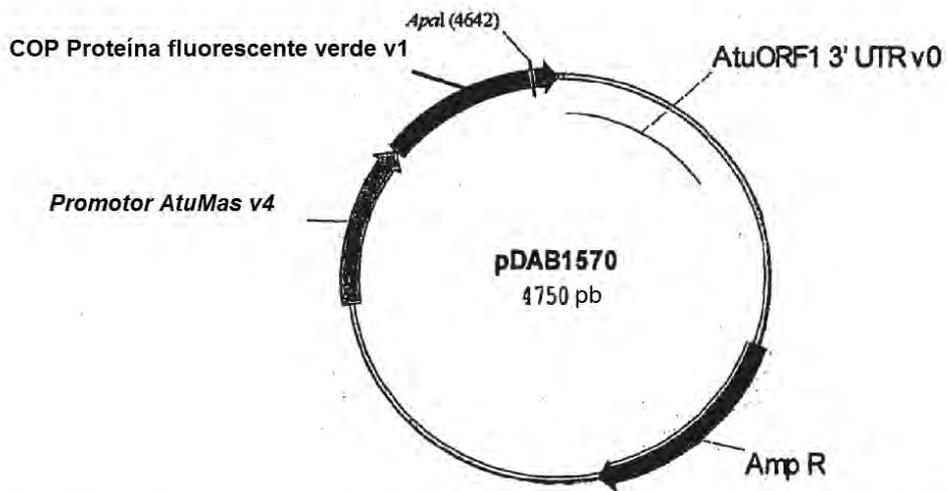


Figura 21: Representación esquemática del plásmido pDAB1570

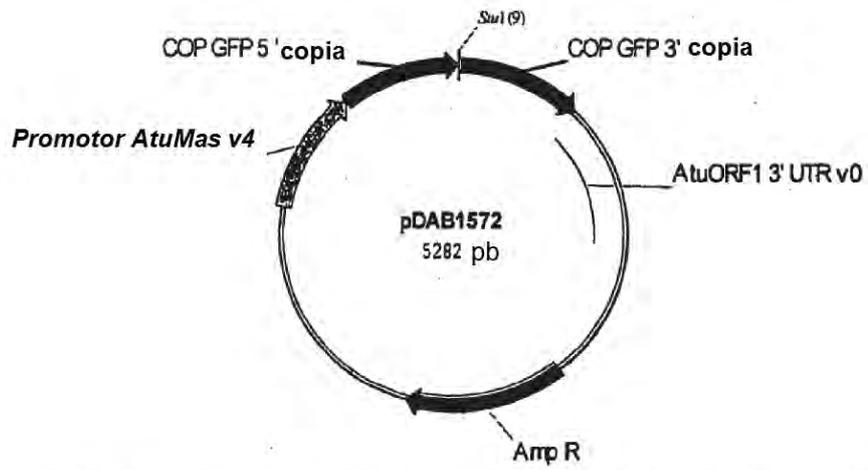


Figura 22: Representación esquemática del plásmido pDAB1572

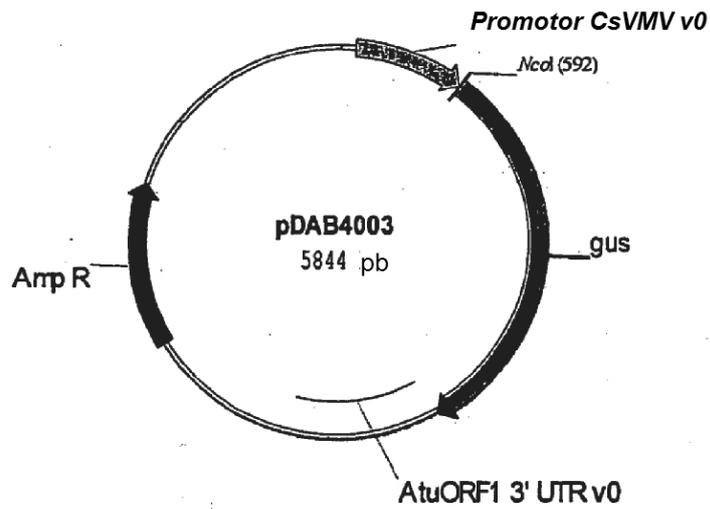


Figura 23: Representación esquemática del plásmido pDAB4003

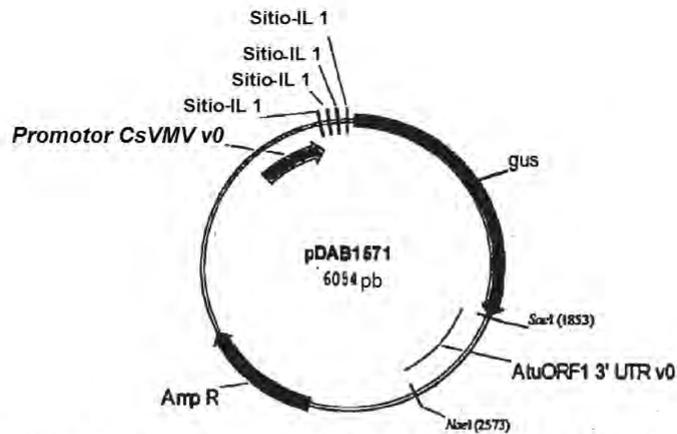


Figura 24: Representación esquemática del plásmido pDAB1571

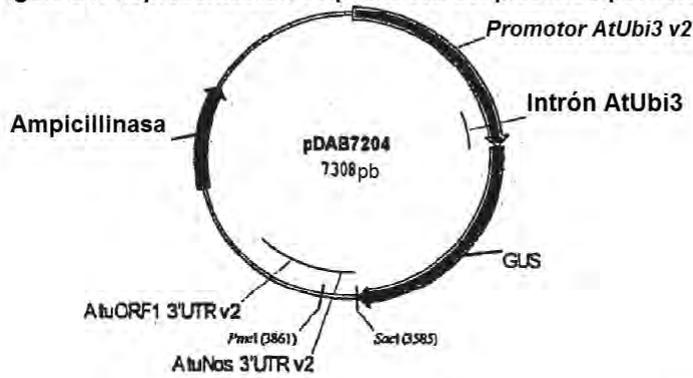


Figura 25: Representación esquemática del plásmido pDAB7204

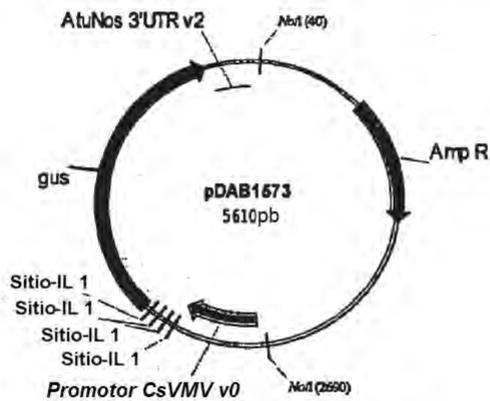


Figura 26: Representación esquemática del plásmido pDAB1573

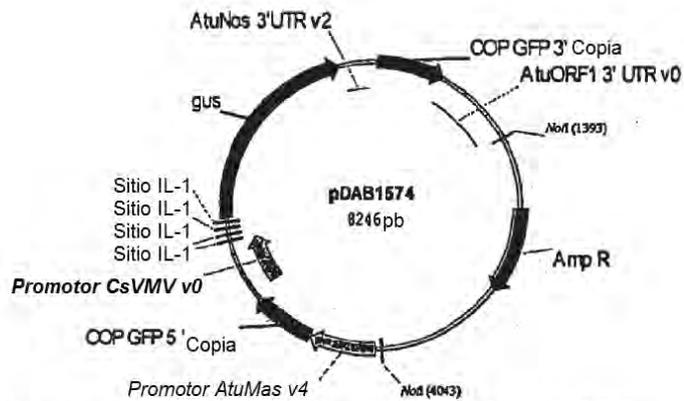


Figura 27: Representación esquemática del plásmido pDAB1574

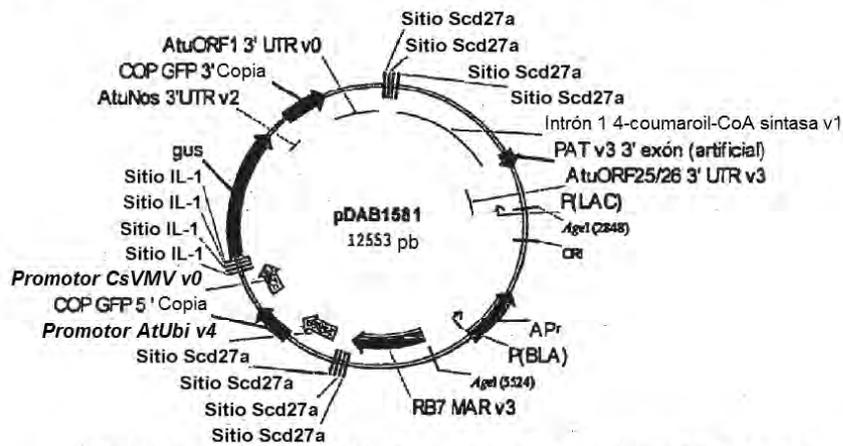


Figura 28: Representación esquemática del plásmido pDAB1581

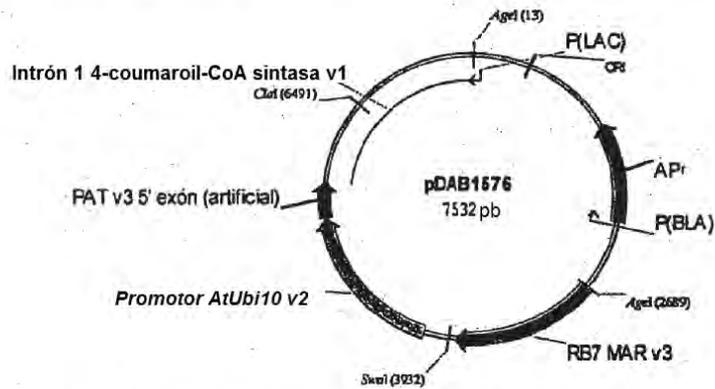


Figura 29: Representación esquemática del plásmido pDAB1576

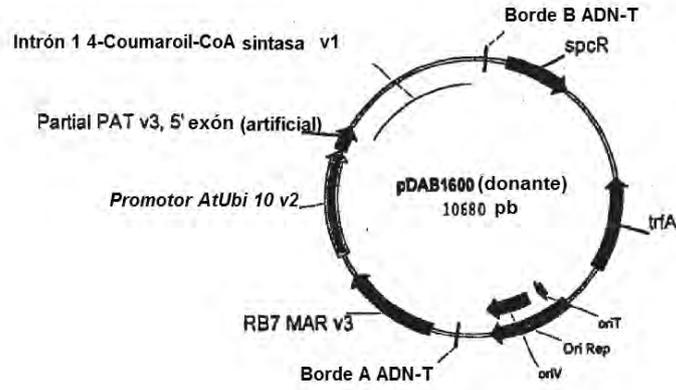
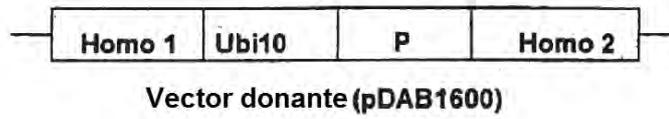


Figura 30: Representación esquemática del plásmido pDAB1600. Donante ADN Vector para el tabaco.

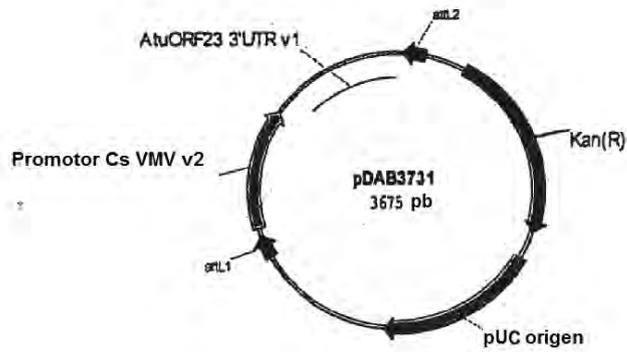


Figura 31: Representación esquemática del plásmido pDAB3731

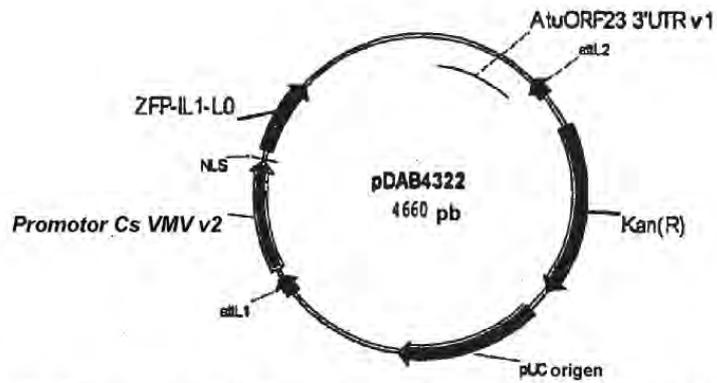


Figura 32: Representación esquemática del plásmido pDAB4322

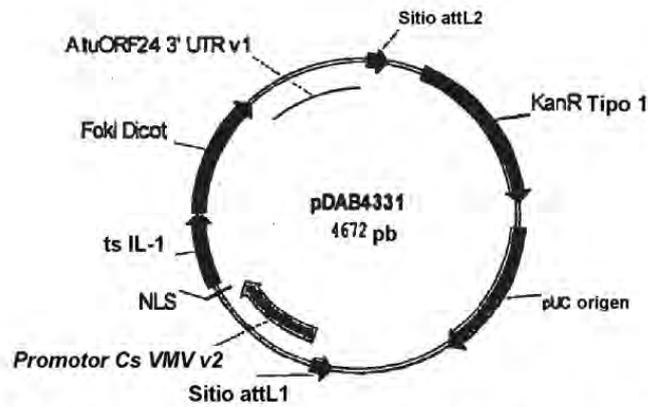


Figura 33: Representación esquemática del plásmido pDAB4331

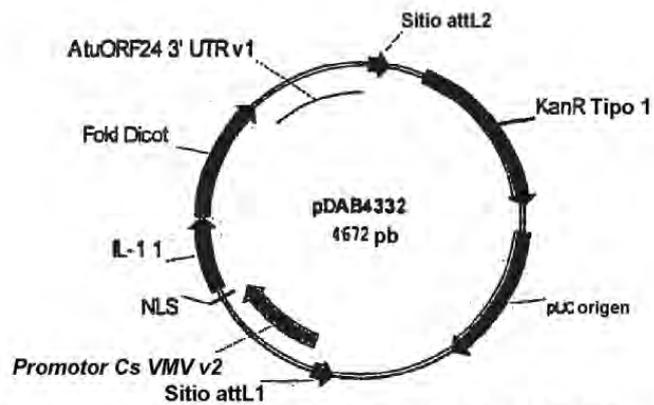


Figura 34: Representación esquemática del plásmido pDAB4332

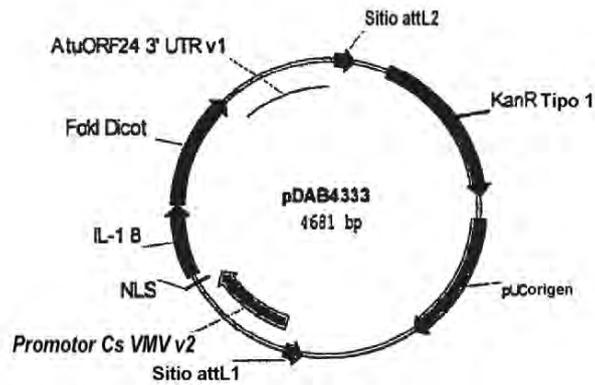


Figura 35: Representación esquemática del plásmido pDAB4333

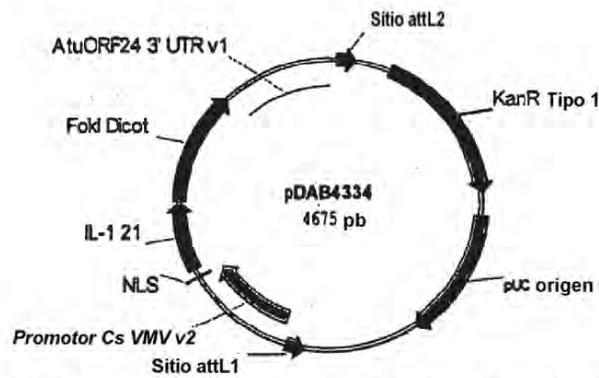


Figura 36: Representación esquemática del plásmido pDAB4334

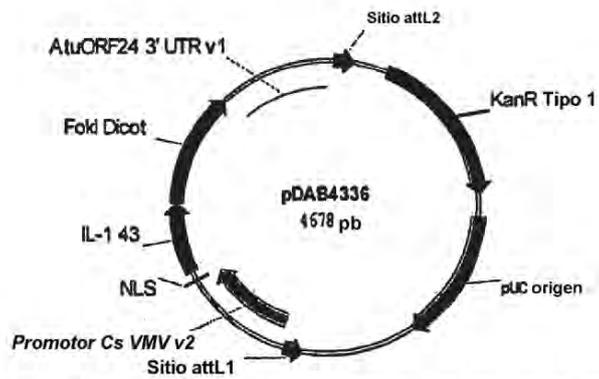


Figura 37: Representación esquemática del plásmido pDAB4336

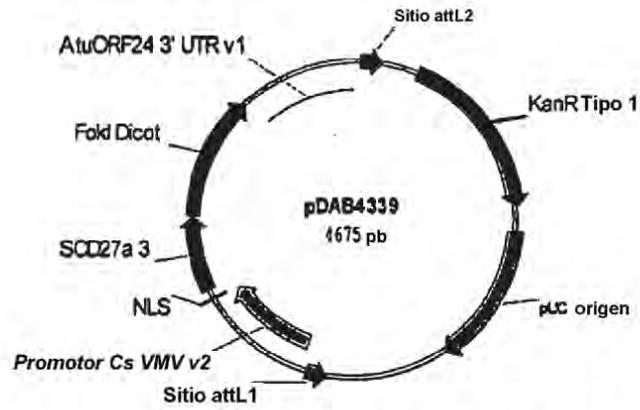


Figura 38: Representación esquemática del plásmido pDAB4339

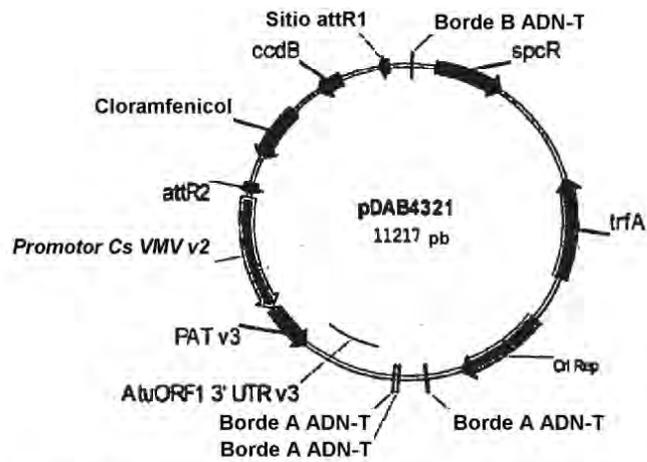


Figura 39: Representación esquemática del plásmido pDAB4321

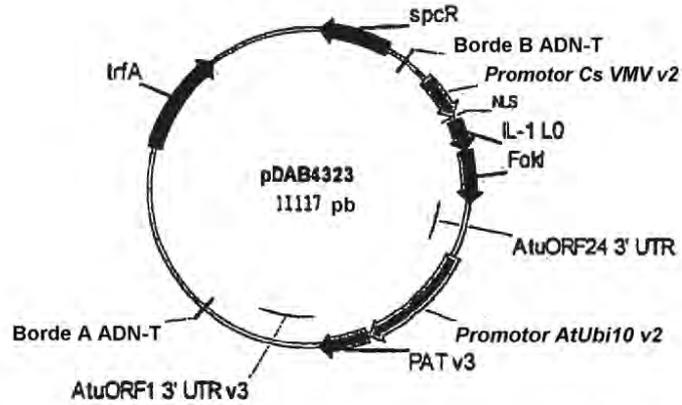


Figura 40: Representación esquemática del plásmido pDAB4323

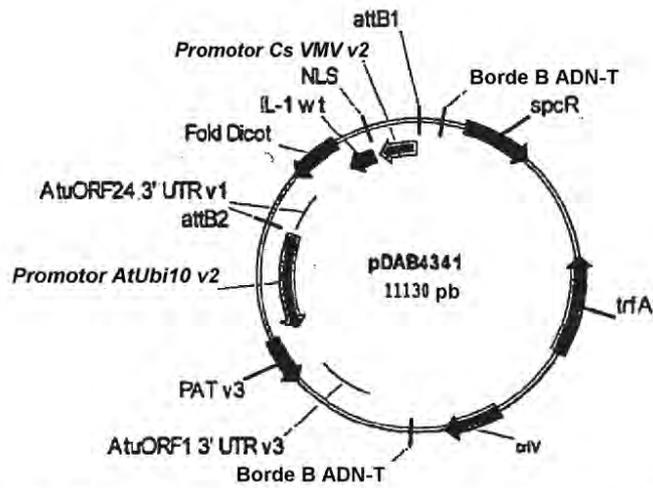


Figura 41: Representación esquemática del plásmido pDAB4341

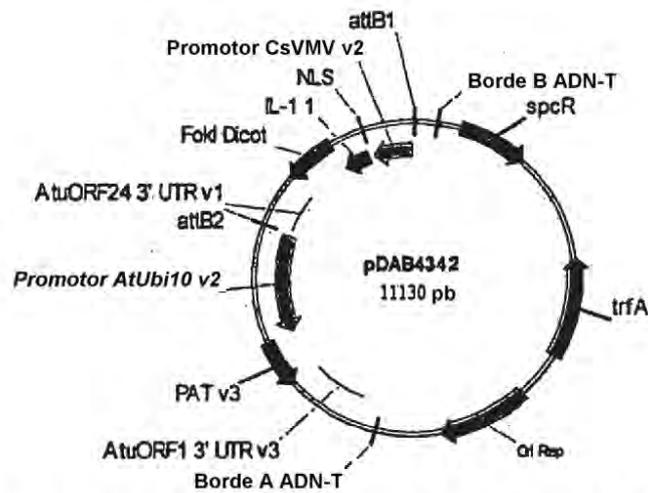


Figura 42: Representación esquemática del plásmido pDAB4342

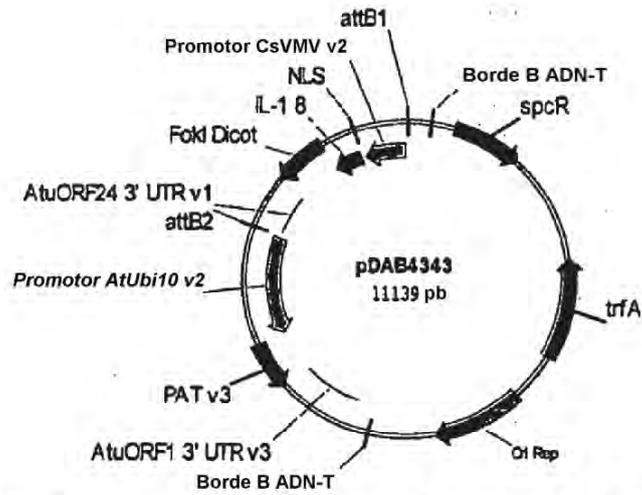


Figura 43: Representación esquemática del plásmido pDAB4343

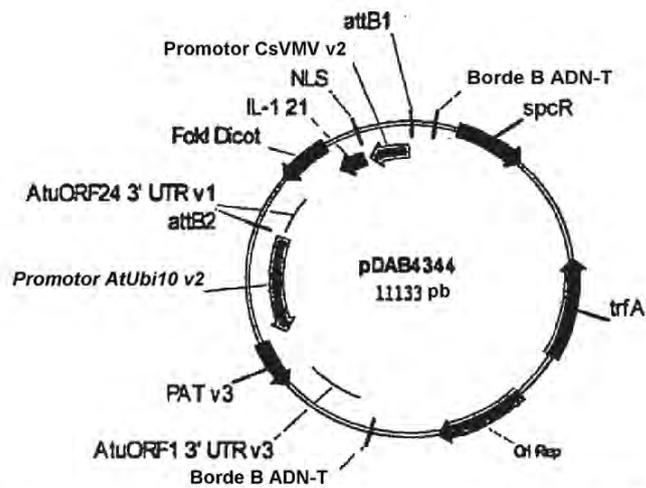


Figura 44: Representación esquemática del plásmido pDAB4344

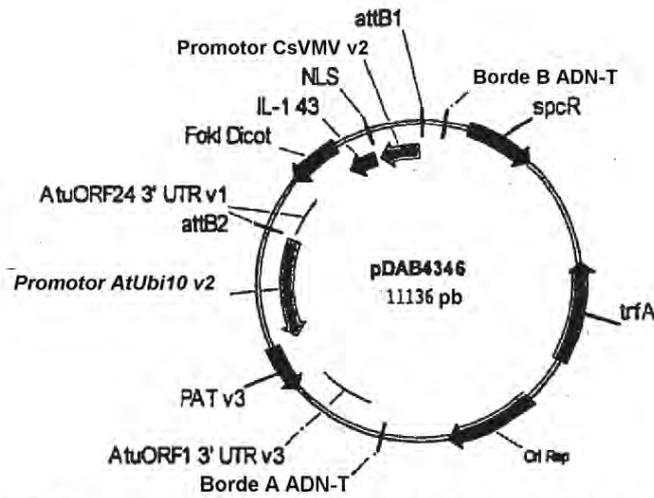


Figura 45: Representación esquemática del plásmido pDAB4346

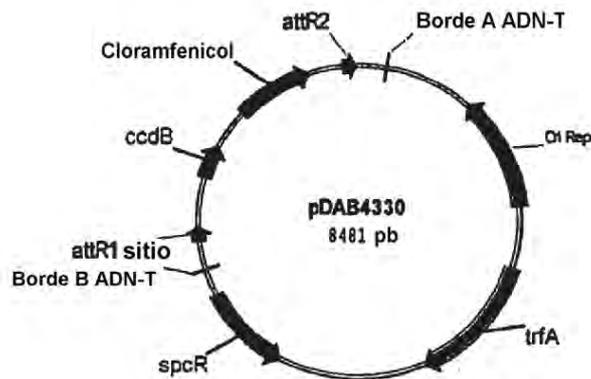


Figura 46: Representación esquemática del plásmido pDAB4330

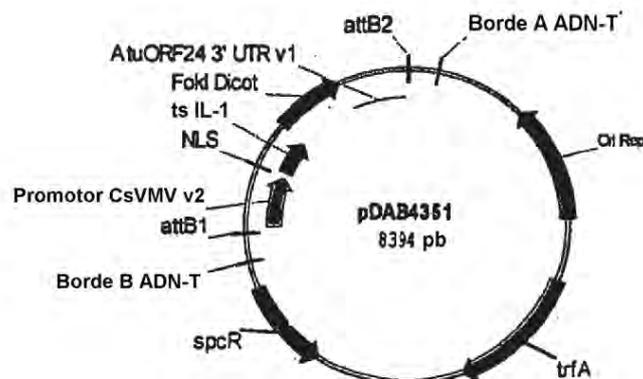


Figura 47: Representación esquemática del plásmido pDAB4351

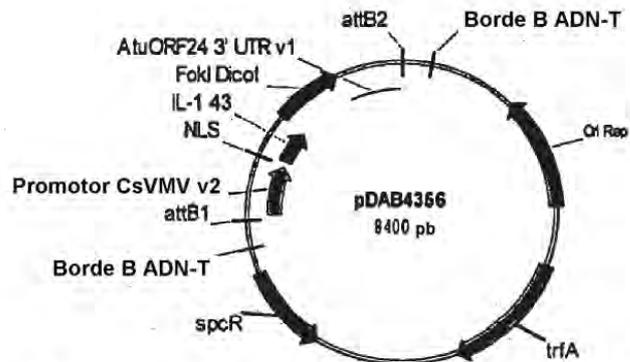


Figura 48: Representación esquemática del plásmido pDAB4356

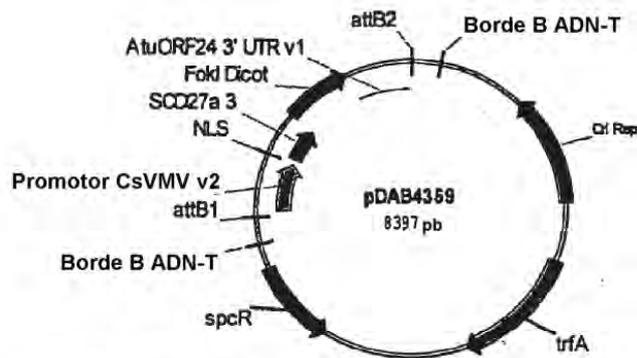


Figura 49: Representación esquemática del plásmido pDAB4359

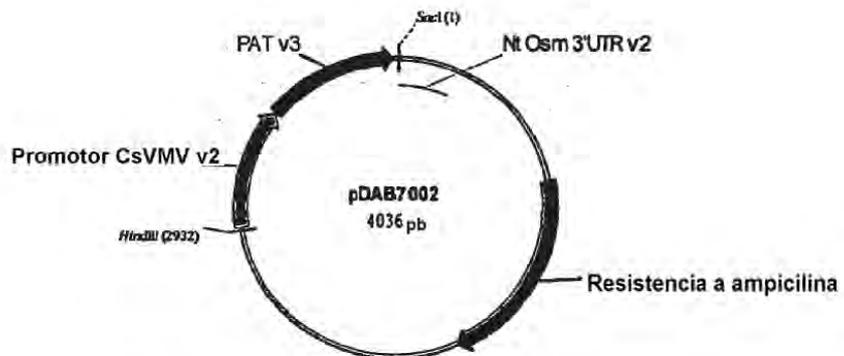


Figura 50: Representación esquemática del plásmido pDAB7002

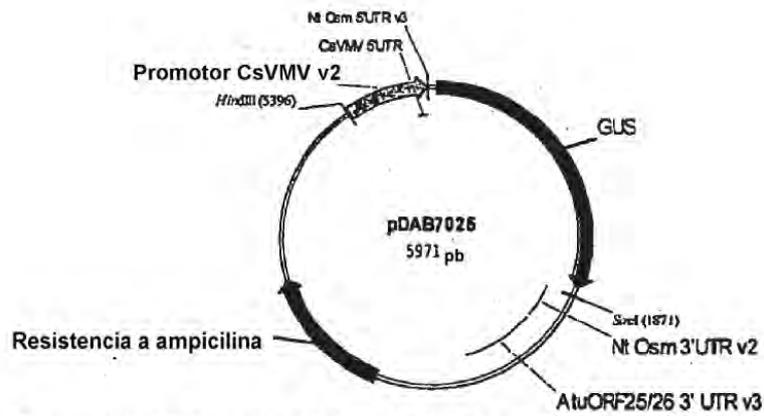


Figura 51: Representación esquemática del plásmido pDAB7026

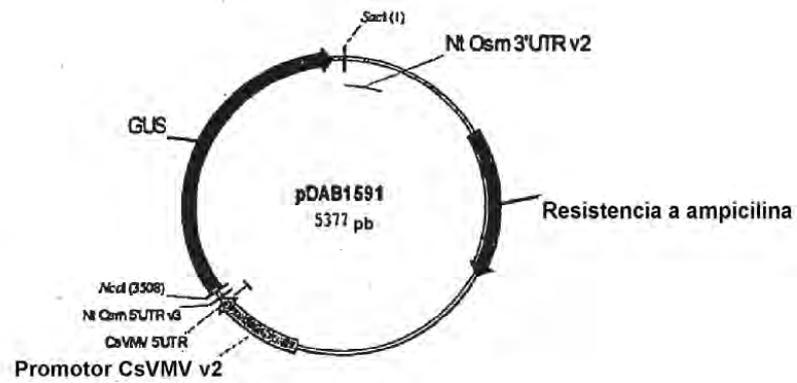


Figura 52: Representación esquemática del plásmido pDAB1591

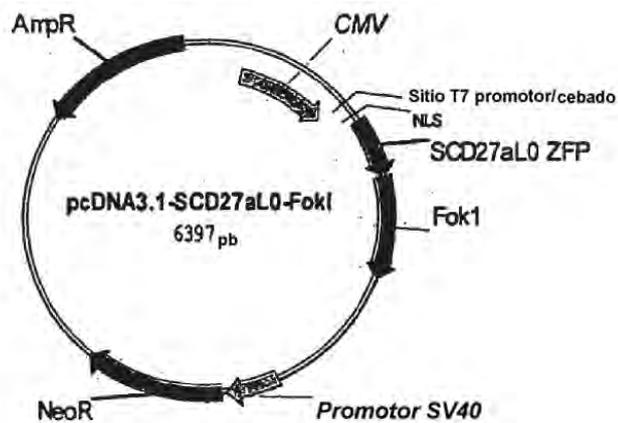


Figura 53: Representación esquemática del plásmido pcDNA3.1-SCD27a-L0-FokI

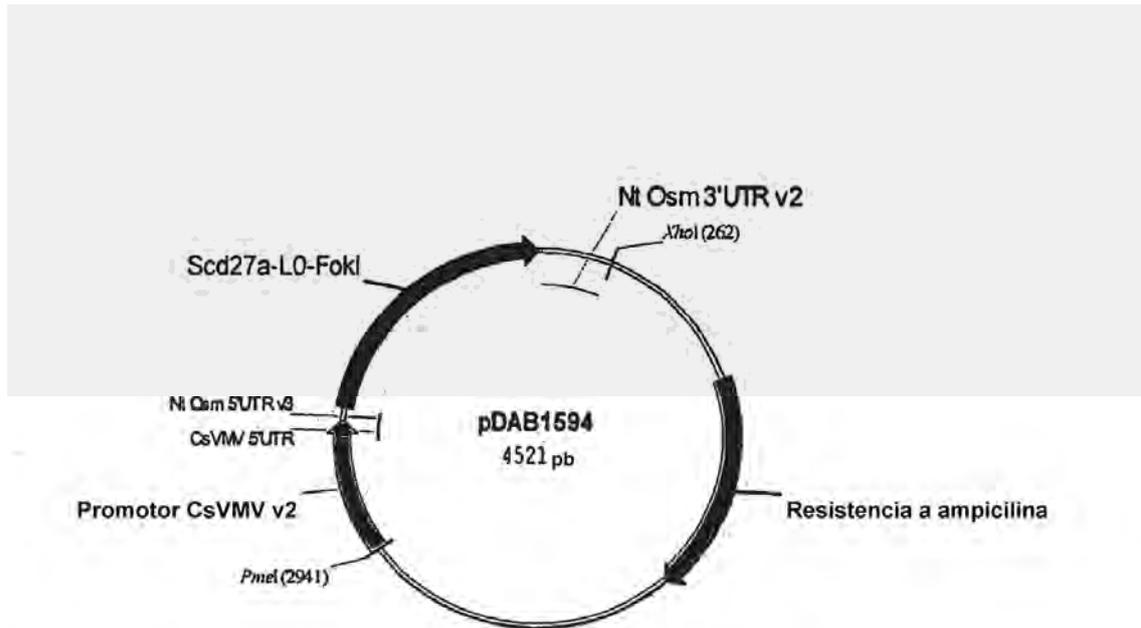


Figura 54: Representación esquemática del plásmido pDAB1594

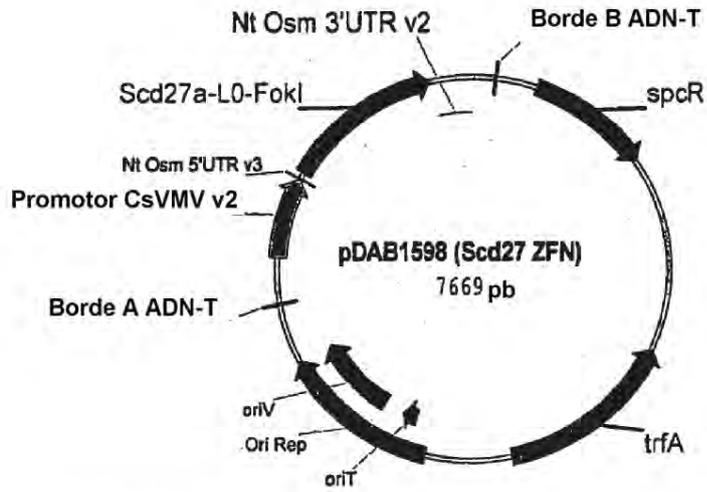


Figura 55: Representación esquemática del plásmido pDAB1598-Vectores de expansión del gen de la proteína de fusión dedo de cinc-Fok 1

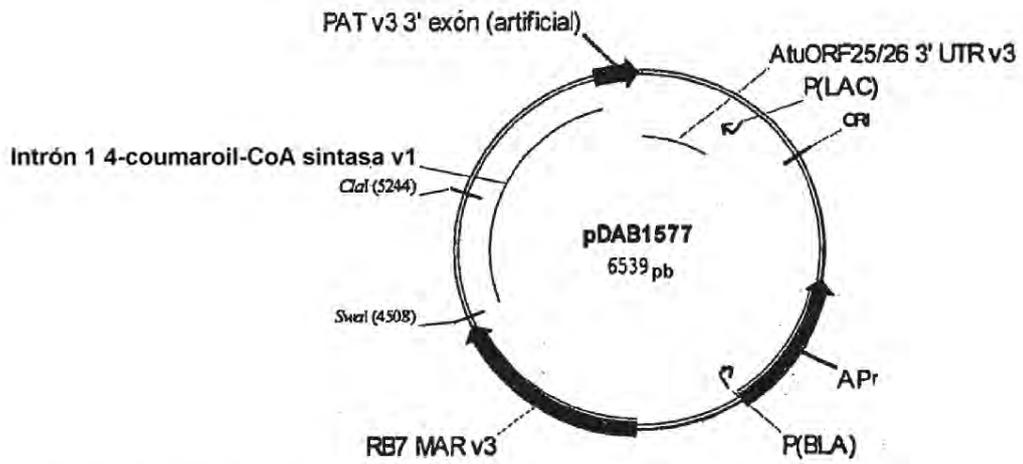


Figura 56: Representación esquemática del plásmido pDAB1577

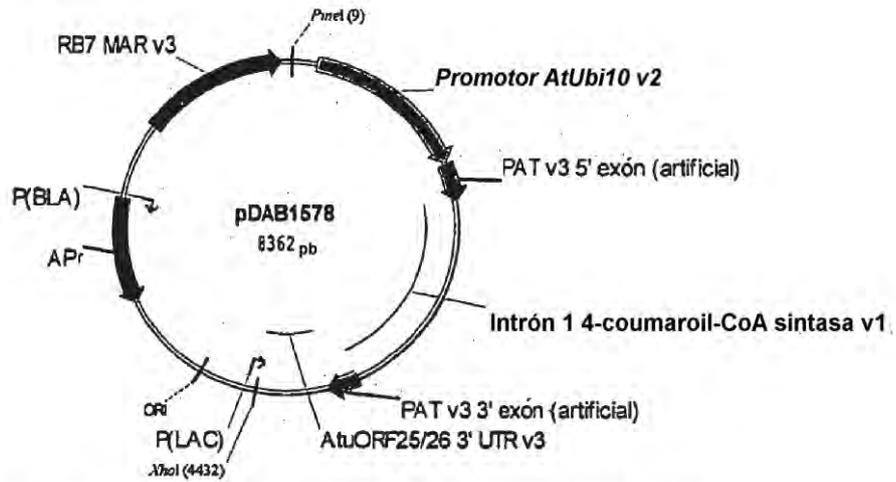
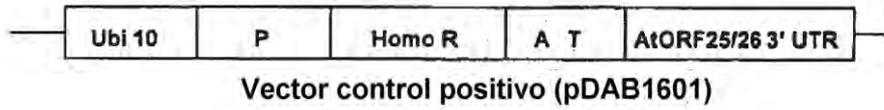


Figura 57: Representación esquemática del plásmido pDAB1578



Vector control positivo (pDAB1601)

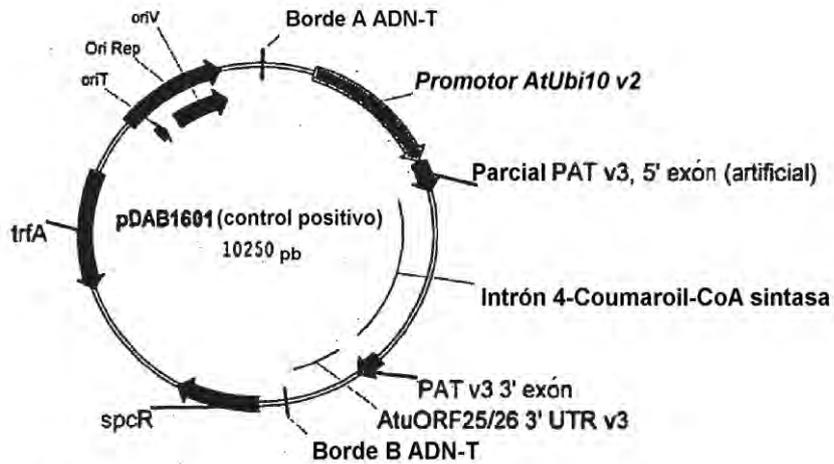


Figura 58: Representación esquemática del plásmido pDAB1601-vector control gen PAT

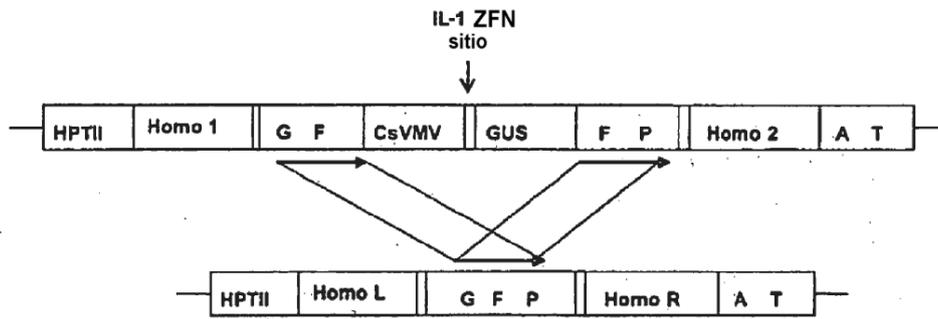
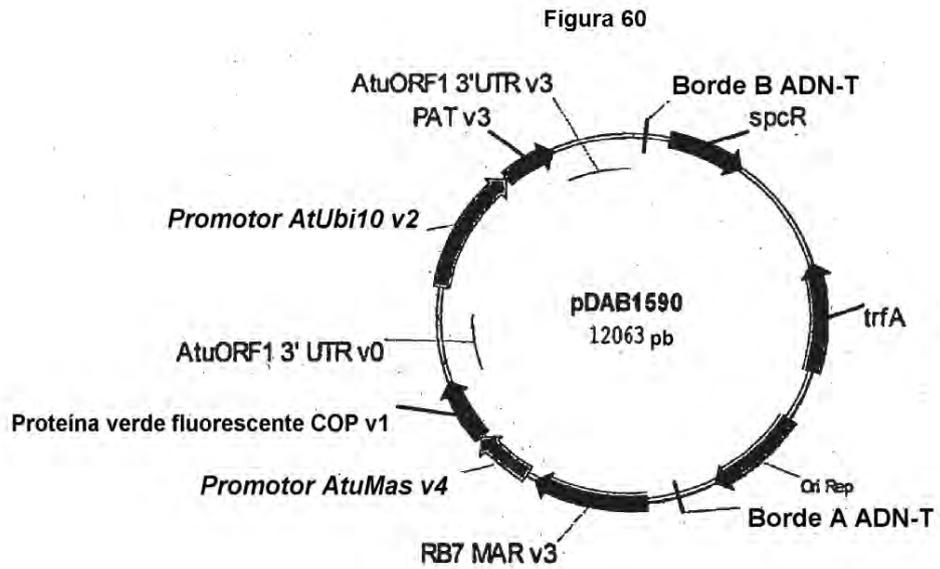


Figura 59: Recombinación homóloga intracromosómica prevista estimulada por la proteína de fusión IL-1-Fok1



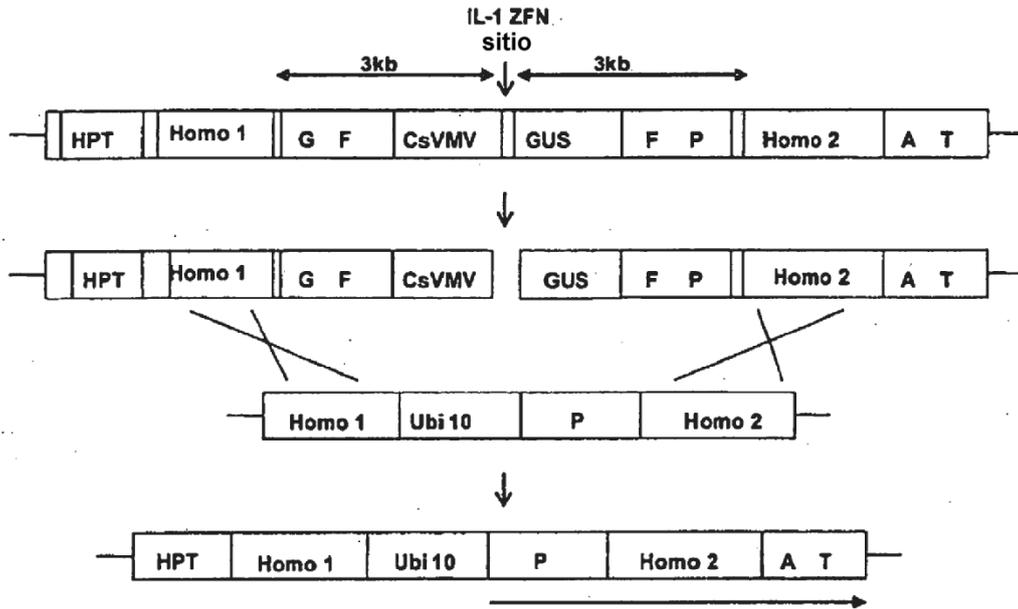


Figura 61: Recombinación homóloga intracromosómica prevista estimulada por la proteína de fusión dedo de cinc-Fok1

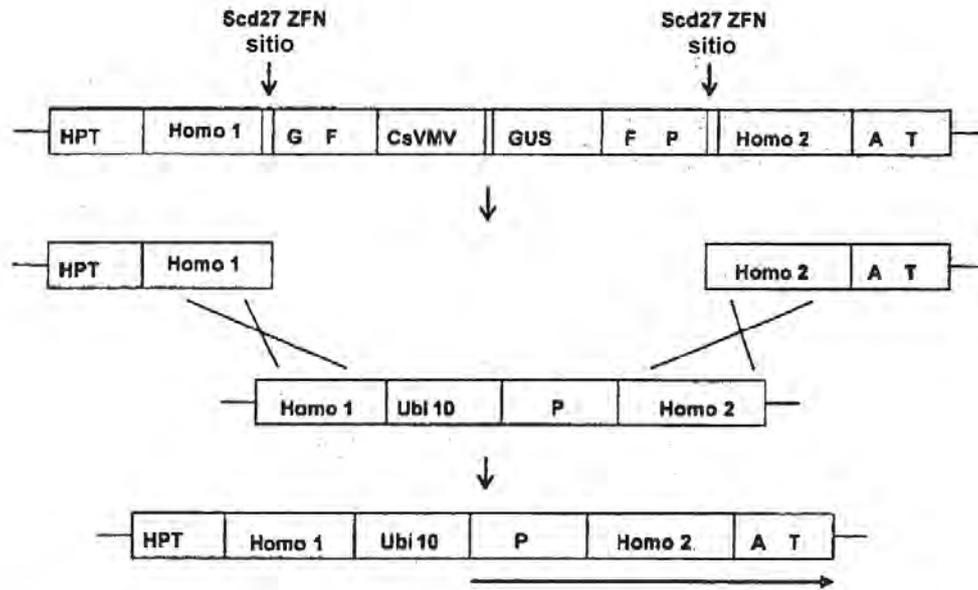


Figura 62: Recombinación homóloga intracromosómica prevista estimulada por la proteína de fusión Scd27 dedo de cinc-Fok1

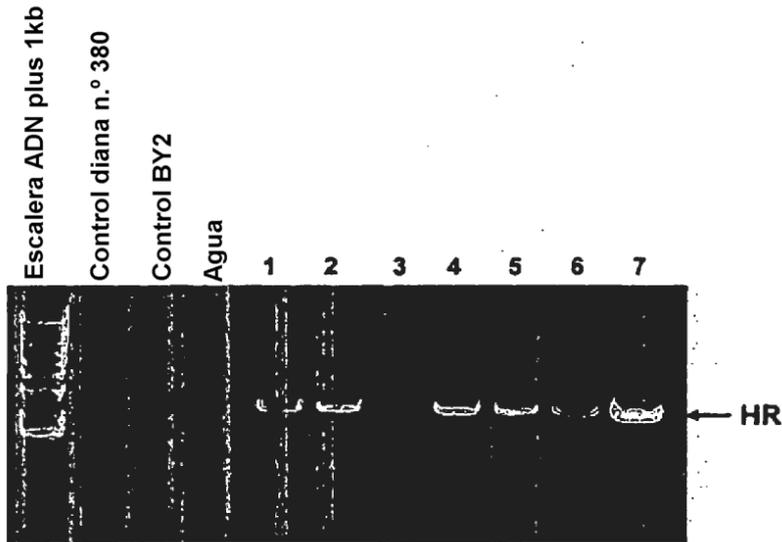


Figura 63: Análisis PCR de los recombinantes. 1-5: Eventos HR de la transformación de BY2-380 con el gen de la proteína de fusión C3H IL-1-Fok1. 6-7: Eventos HR de la transformación de BY2-380 con el gen de la proteína de fusión C3H SCD27-Fok1

Figura 64: Secuencia génica de IPP2K de

```

1   ATGGAGATGG ATGGGGTTCT GCAAGCCGCG GATGCCAAGG ACTGGGTTTA
51  CAAGGGGGAA GGCGCCGCGA ATCTCATCCT CAGCTACACC GGCTCGTCGC
101 CCTCCATGGT AAGCGCTGAG TAGGTTCTTA CTGAGCGTGC ACGCATCGAT
151 CACTTGACTT TAGGGGCTCA ATGTGTGATT CACGGGTGCC GCGGCGCCAT
201 TCGAGCTCCA GATCCAGTAC CGCTCGAGCA AGTGATAAAA CATGGAGCAG
251 GGACGATCAC GTGGTCACTT GAAAATTACG TGAGGTCCGG GCAGACGATG
301 TACGGCGCGG CGAACTCTCA AACACTCACA CAACCAAAC  CGCTTCGTGT
351 TCGTCTTTGT TCCAAGCGAC TGTGTGAGTG TTTGAGAGTT CGCCAGCGCG
401 ACATCGCCCG ATCTGACAAA TTAAGCTTTC GTTGCTTTTC CATGATTGTG
451 CATTTTGTGA GCATGCACTG AATACTATGA TGGATATGTT TGGAGGAAGC
501 ATTATTCCAA TTTGATGATA AGGGTGTTAT TTACACTTGT TTCAGCTTG
551 GCAAGGTACT GCGGCTCAAG AAGATTCTAA AAAACAAGTC GCAGCGGGCA
601 CCGAGTTGTA TTGTATTCTC AAGTCATGAG CAACTCCTGT GGGGCCATAT
651 CCCAGAACTG GTTGAGTCGG TCAAACAAGA TTGCTTGGCT CAAGCCTATG
701 CAGTGCAATG TATGAGCCAA CACCTGGGTG CCAATCATGT CGATGGTGGG
751 GTATGGTTCA GATTCAGTTC ATTTATGTCC TGTTATTGTG ATTTTGATTG
801 GTAACATATT GACAACCTCG ACACTTGGGA TCAGATTGAG TTCACTTATG
851 GAAGAAATG  GAGAATTGTT ATAATTTATC TATAATCACC CCTACTGAAA
901 TAGAAATAAC ATGGCATCAA TGTGCATGCT ATTGGATTTT GACACGAATA
951 TGCTTTATTC TATCATATGT TGGTAATTCC AGCAGGCAGC AGGCACTACT
1001 CTTTGGATCC ACGTGACTTG ACAAAGAAAT CATGCCATCT TTCCACAATG
1051 CAGGTCCTGT TACGTGTTTC TAGGGATTTT CTGGAGCTTG TCGAAAAGAA
1101 TGTTCTTAGC AGCCGTCCCTG CTGGGAGAGT AAATGCAAGT TCAATTGATA
1151 ACACTGCTGA TGCCGCTCTT CTAATAGCAG ACCACTCTTT ATTTTCTGG
    
```

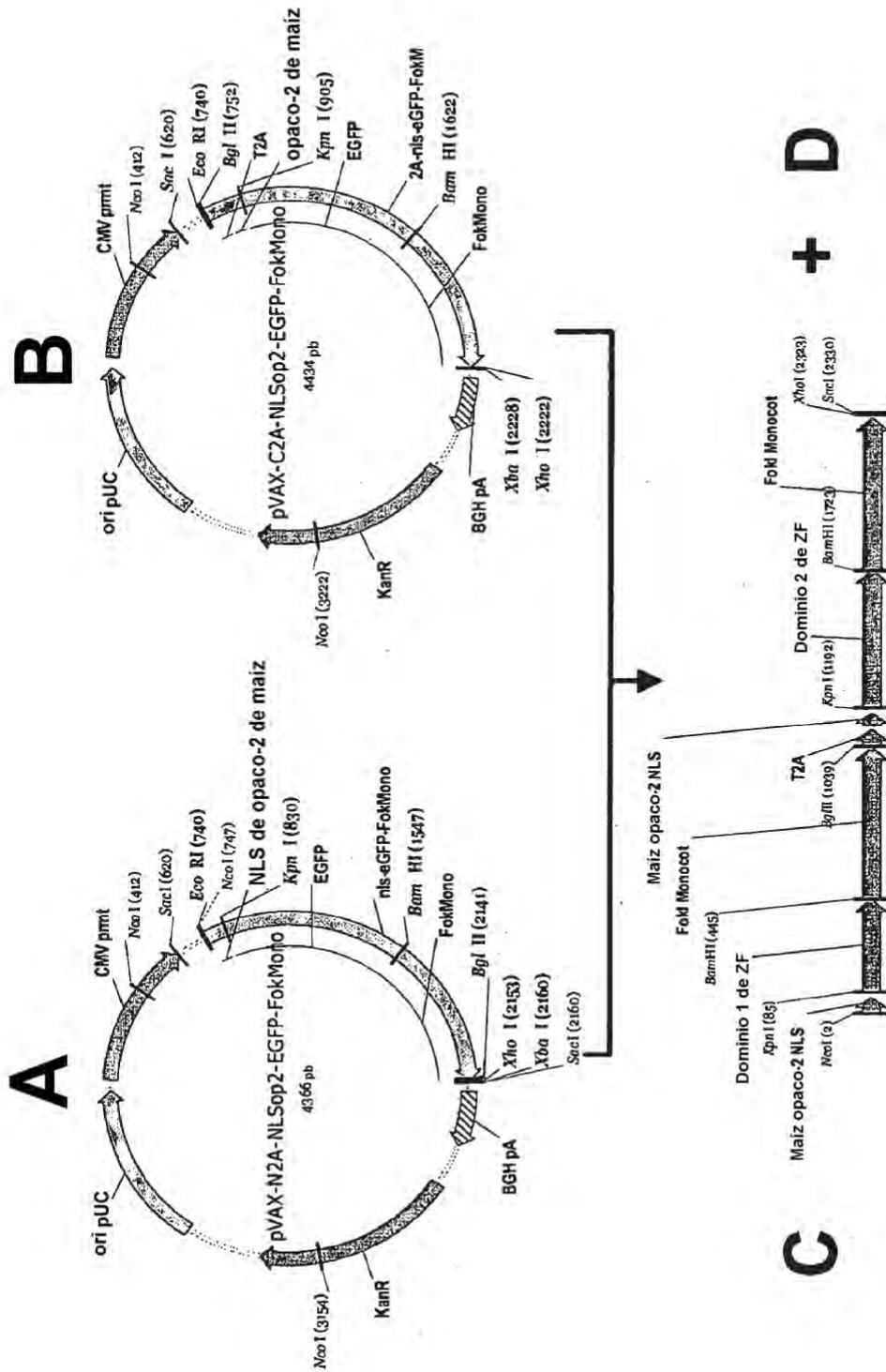


Figura 65.1: Esquema de clonación del vector de expresión de ZFN

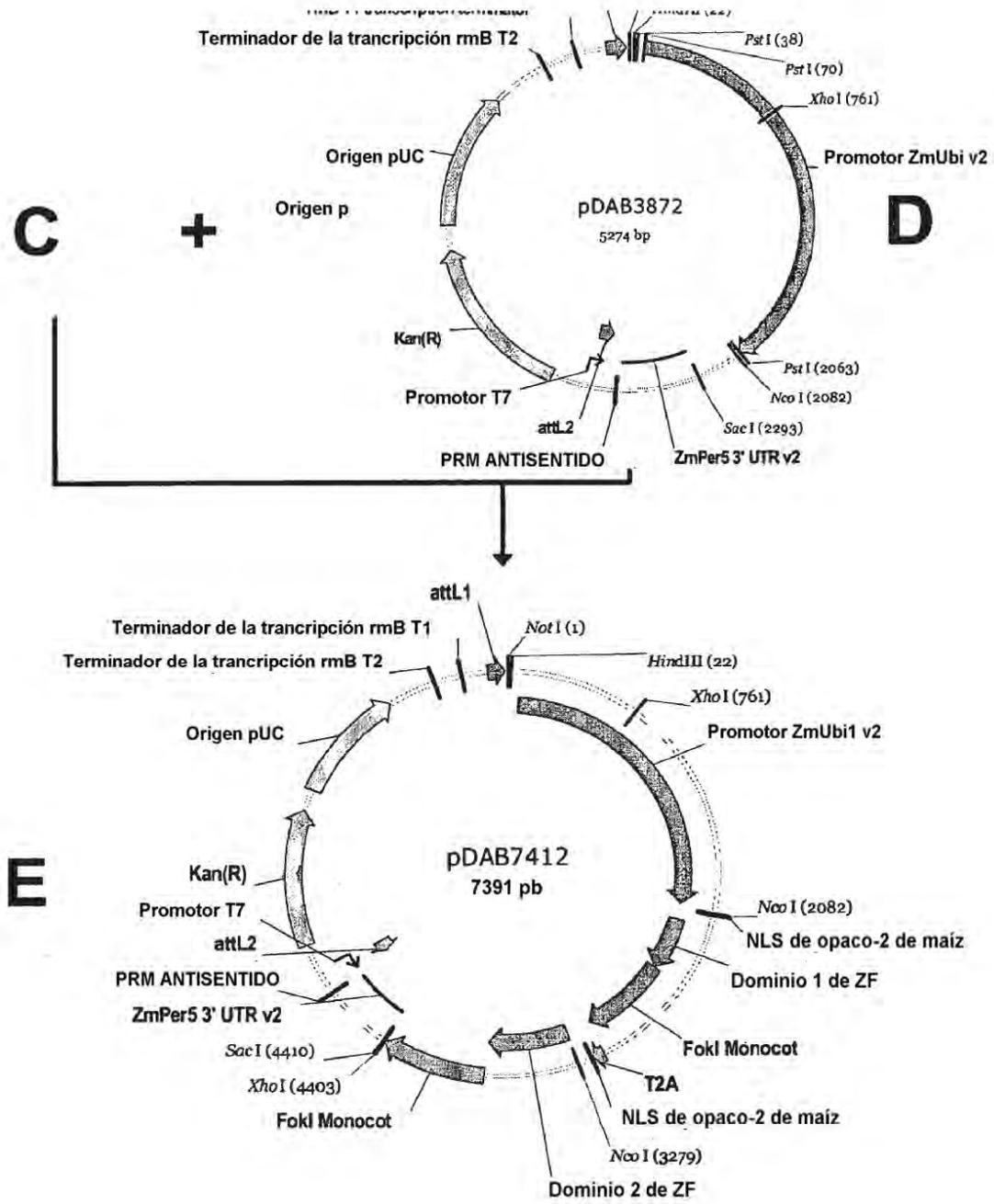


Figura 65.2: Esquema de clonación del vector de expresión de ZFN

Figura 66: Unión y escisión de los ZFN para el gen IPP2K del maíz

```

      sitio de unión 8705
caactcctgtggggccatataccagaactggttgagtoggtcaacaaga
gttgaggacacccoggtatagggtcttgaccaactcaaccagtttgttct
      8684 sitio de unión
      8685
      8686
    
```

Figura 67: Delección mediada por ZFN

```

tipo silvestre AAGTCATGAGCAACTCCTGTGGGGCCATATCCAGAACTGGTTGAGTCGGTCAAACAAG
clon 127       AAGTCATGAGCAACTCCTGTGGGGCCA-----AGAACTGGTTGAGTCGGTCAAACAAG
    
```

Figura 68: Delecciones mediadas por ZFN

```

tipo silvestre AAGTCATGAGCAACTCCTGTGGGGCCATATCCAGAACTGGTTGAGTCGGTCAAACAAG
experim.      AAGTCATGAGCAACTCCTGTGGGGCC: :A: : :AGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : :AGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : : :CCAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATAT: : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATAT: : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCC: :A: : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATAT: : :AGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :GAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :AGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATAT: : :AGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CAGAACTGGTTGAGTCGGTCAAACAAG
              AAGTCATGAGCAACTCCTGTGGGGCCATA: : : :CCAGAACTGGTTGAGTCGGTCAAACAAG
    
```

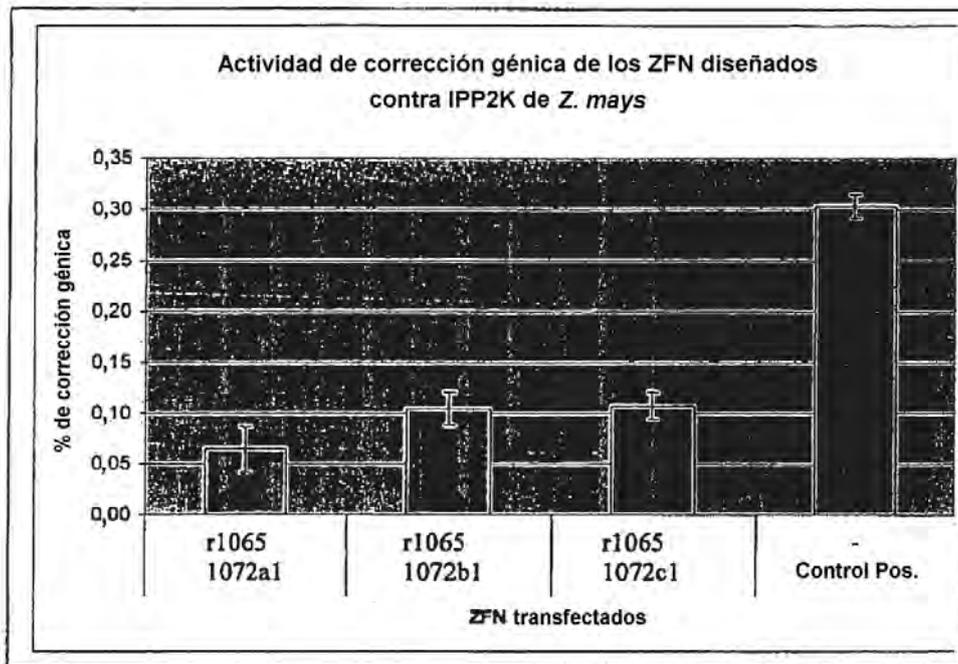


FIG. 69

Figura 70: Contrucción del plásmido pDAB7471

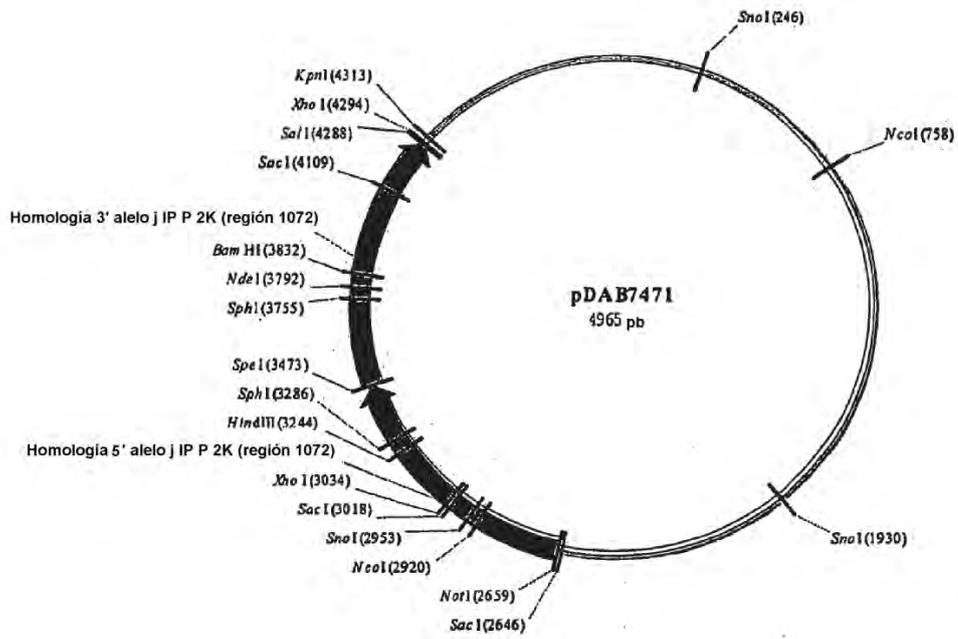


Figura 71: Construcción del plásmido pDAB7451

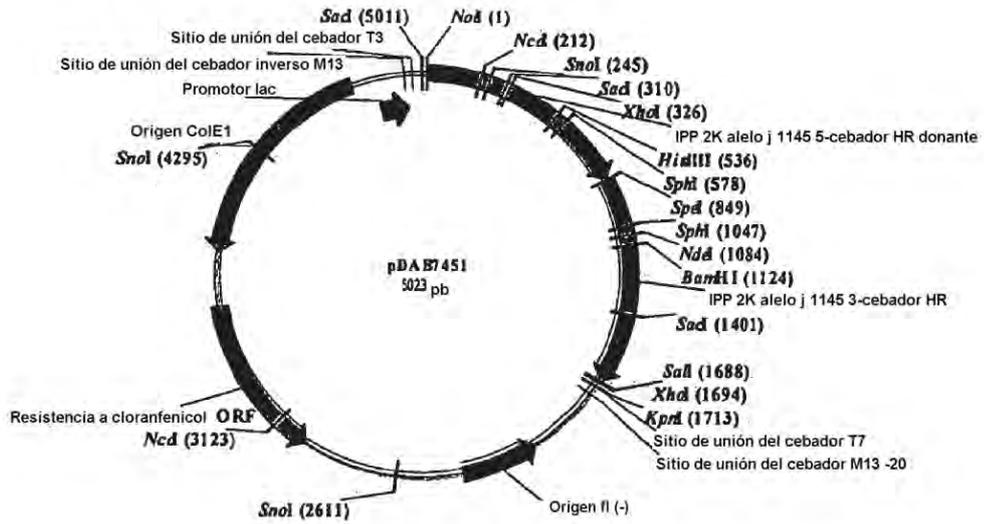


Figura 72: Casete autónomo de expresión génica de tolerancia a herbicida

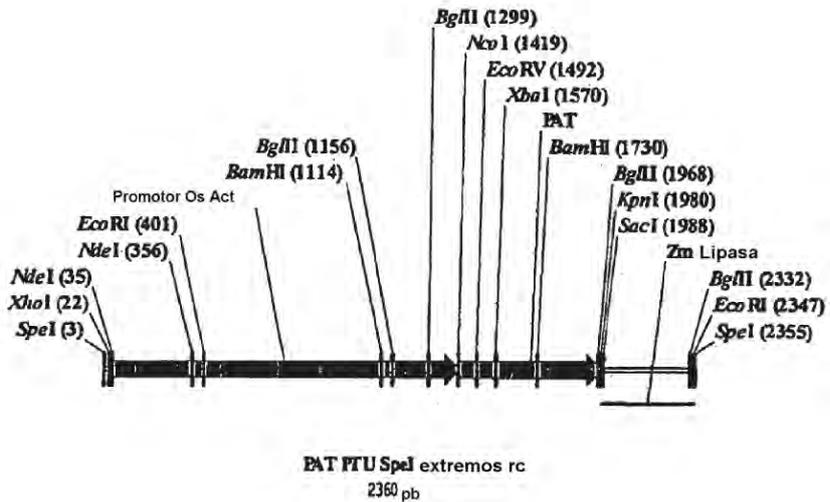


Figura 73

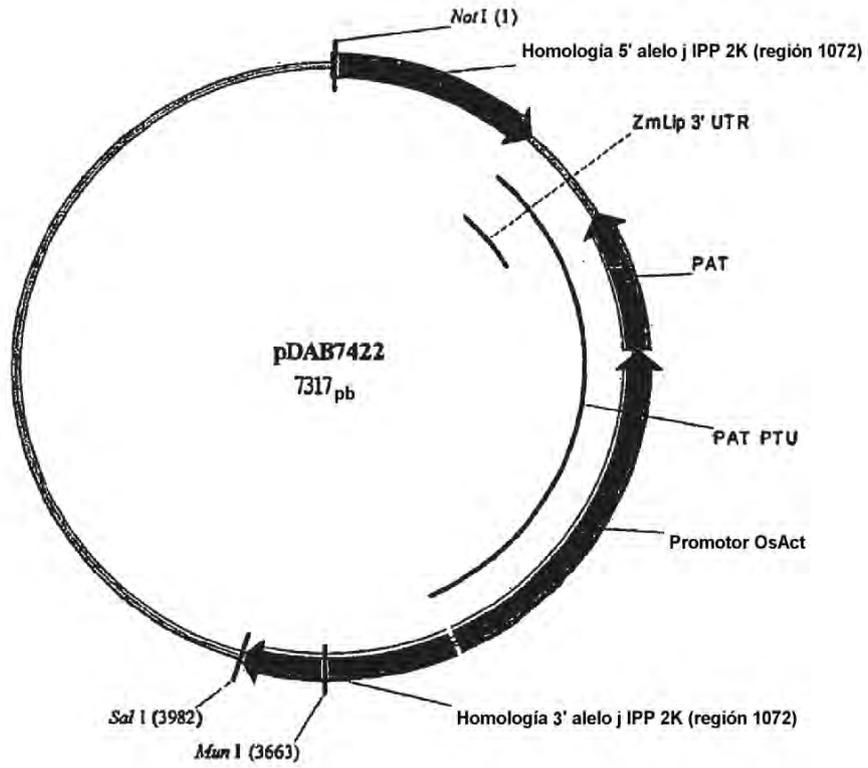


Figura 74: Construcción del plásmido pDAB7452

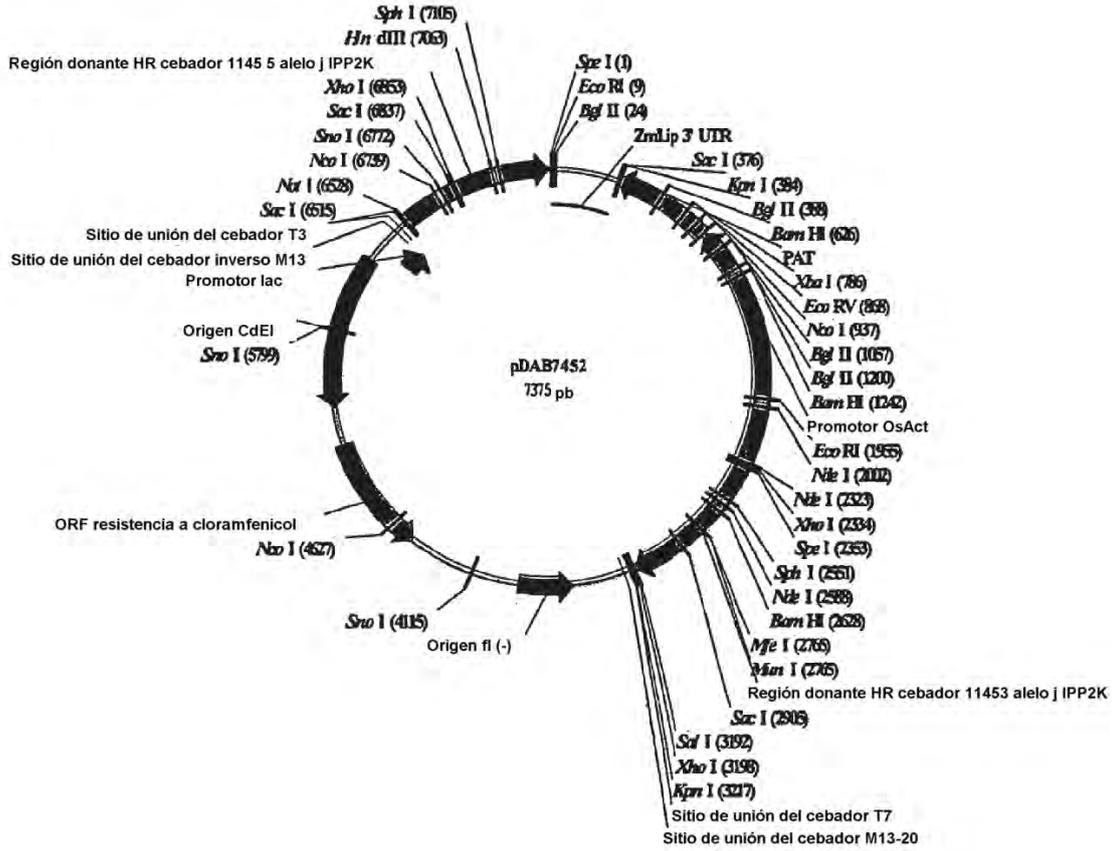


Figura 75: Casete de expresión génica de tolerancia a herbicida no autónomo.

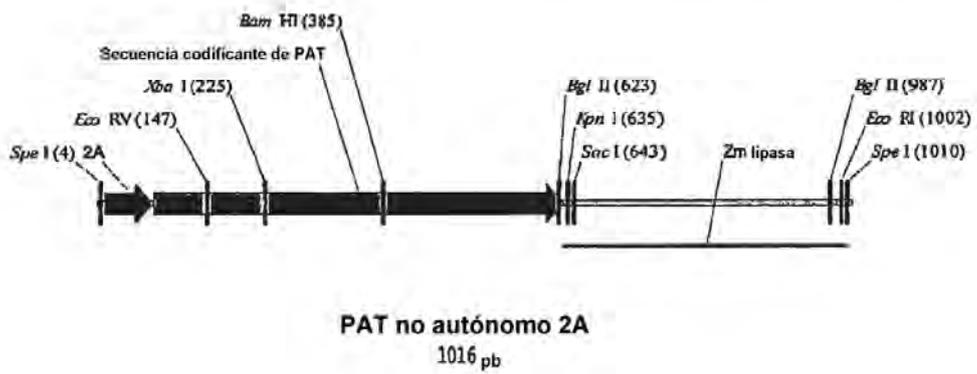


Figura 76: Construcción del plásmido pDAB7423

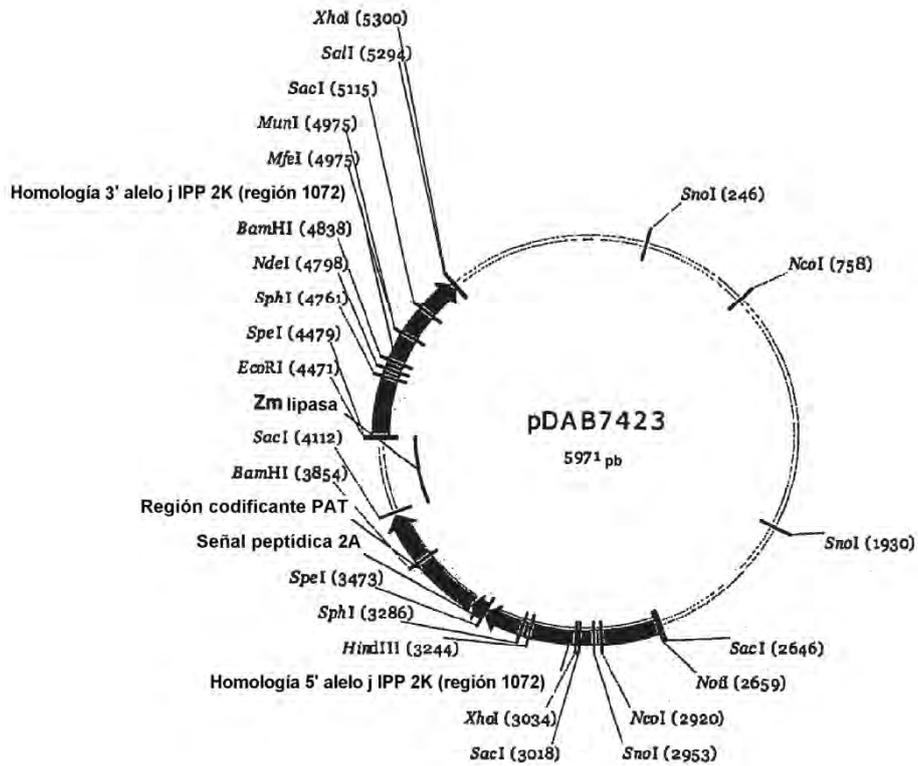


Figura 77: Construcción del plásmido pDAB7454

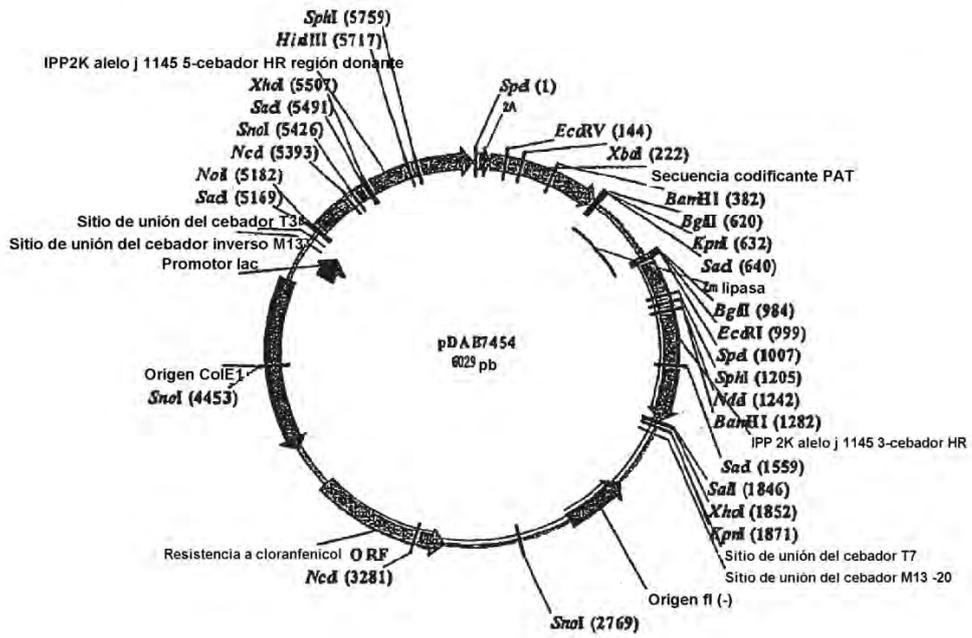


Figura 78: Construcción del plásmido pDAB7424

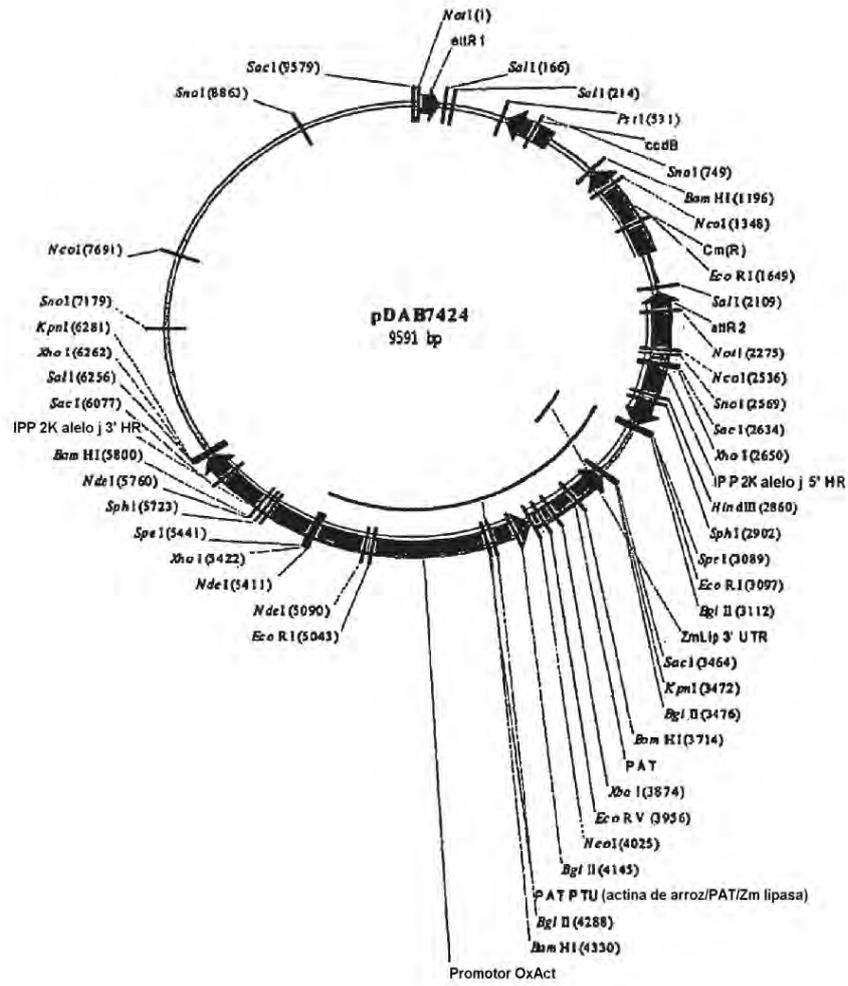


Figura 79: Construcción del plásmido pDAB7425

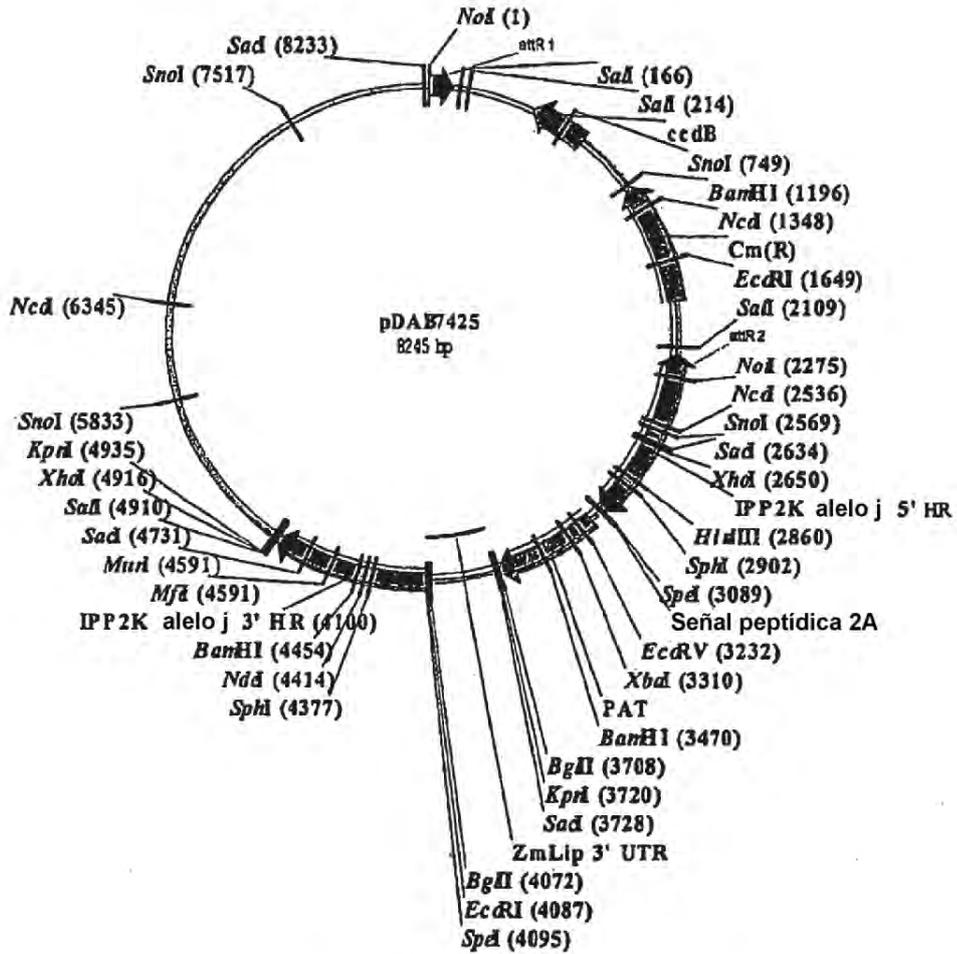


Figura 80: Construcción del plásmido pDAB7426

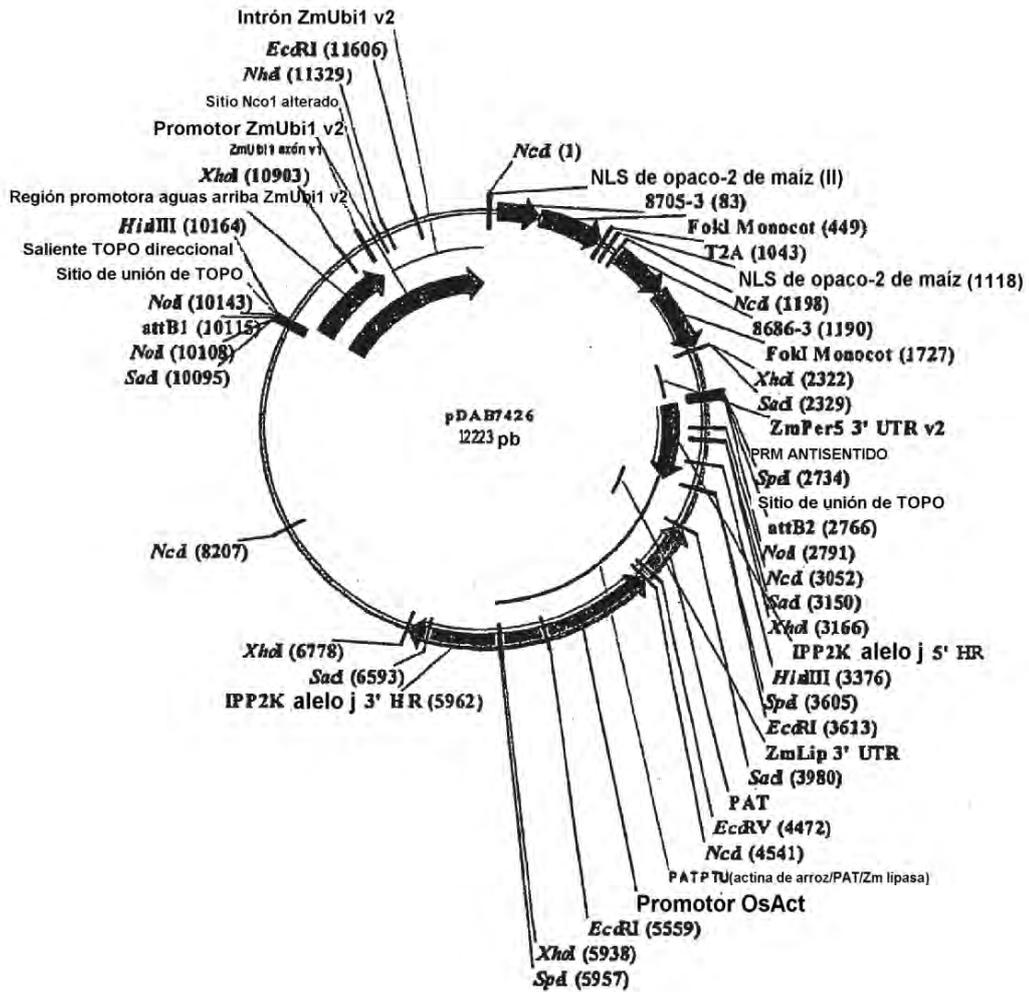


Figura 81: Construcción del plásmido pDAB7427

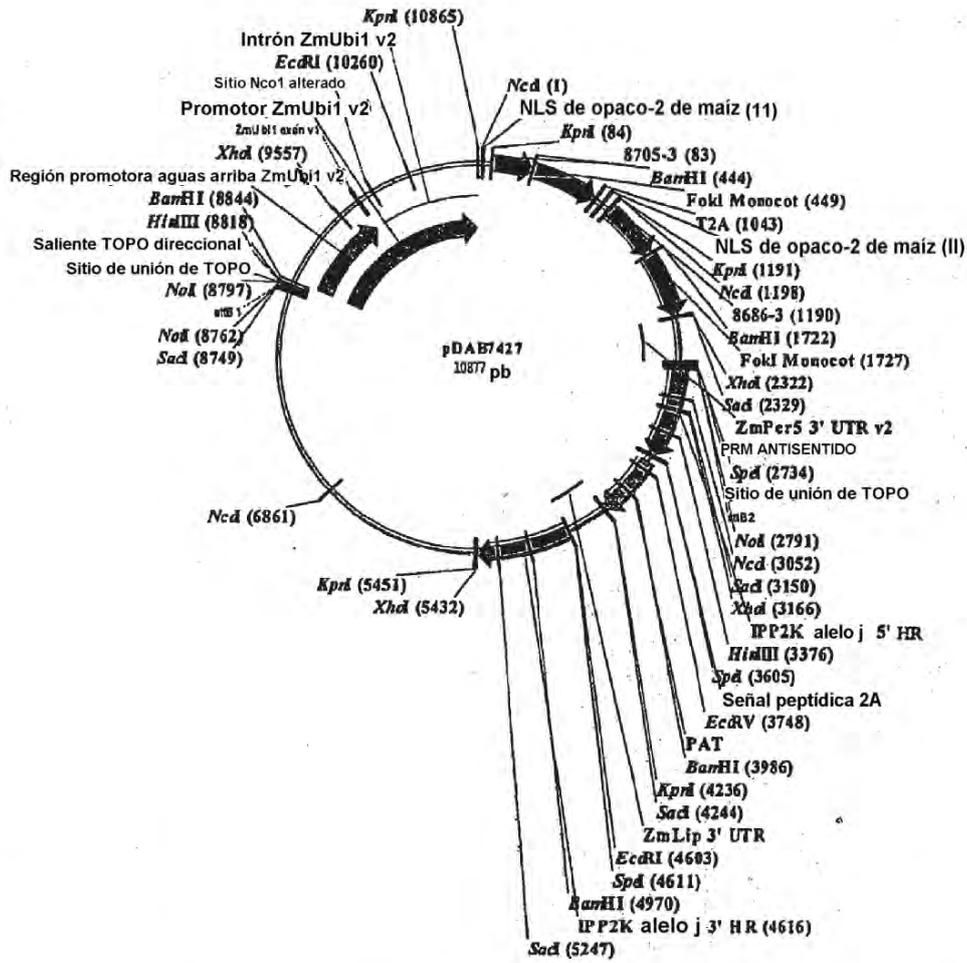


Figura 82

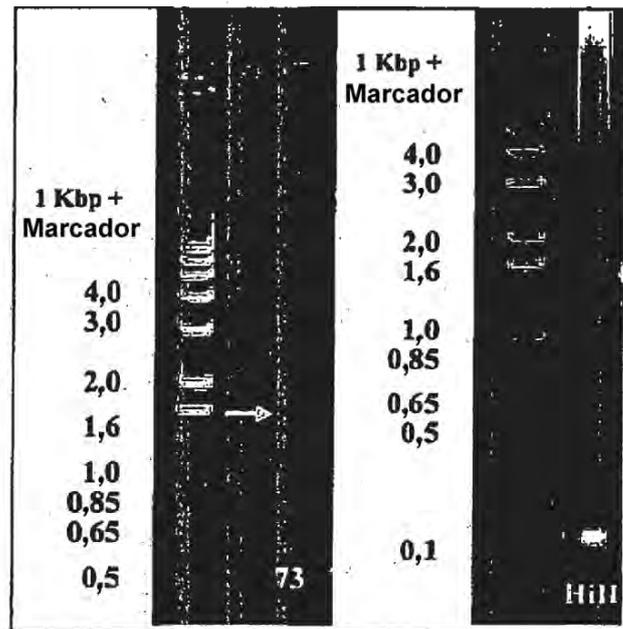
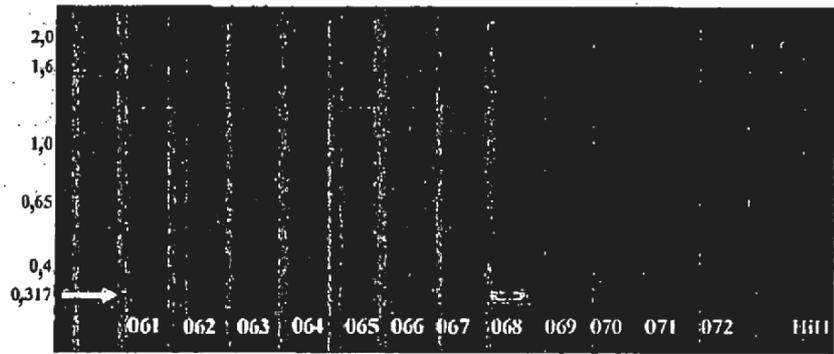


Figura 83

Figura 84

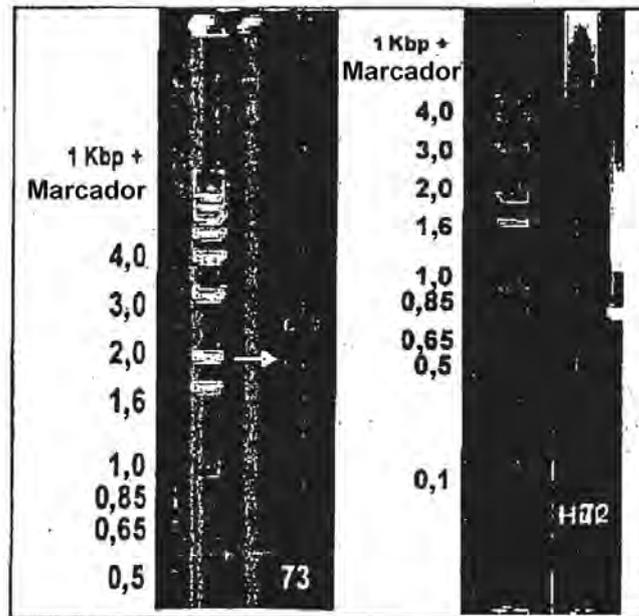


Figura 85

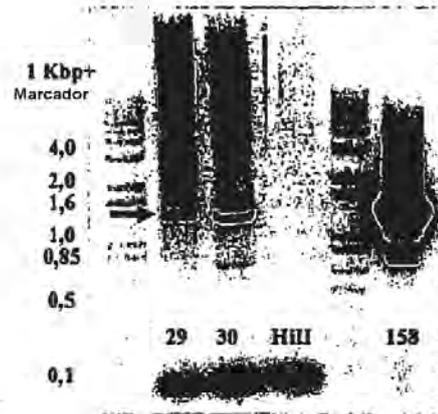


Figura 86

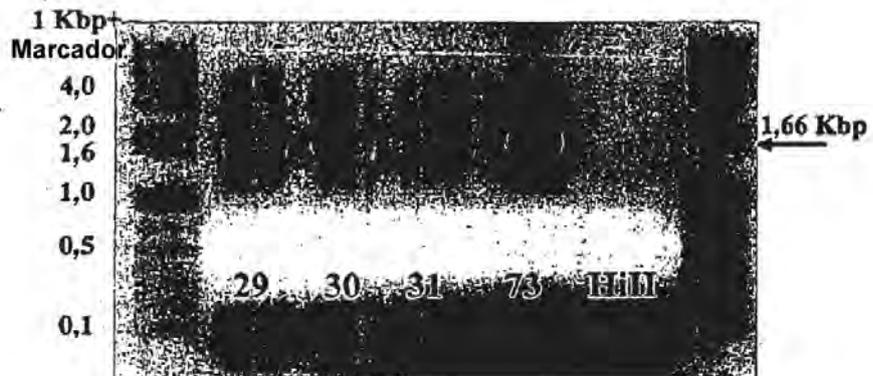


Figura 87: Secuencia flanqueante de homología 5' en posición 1 (SEQ ID NO: 1)

```

GCGGCCGCGT CTCACCGCGG CTTGGGGATT GGATACGGAG CTAGTTAACC
AGCAGAGCTA GATAGCAGAC GCAGATCGCT TGCTTCTCTG GTTTGATTTT
TGGAGTCACC ATTTCTGTTT GGTTTCGTGTG CCTCAGTGTC TGACAGCAGC
AGATCCTCGA TGGAGATGGA TGGGGTTCTG CAAGCCGCGG ATGCCAAGGA
CTGGGTTTAC AAGGGGGAAG GCGCCGCGAA TCTCATCCTC AGCTACACCG
GCACGTCGCC CTCCATGGTA AGCGCTGAGT AGGTTCTTAC TGAGTGTGCA
CGCATCGATC ACTTGACTTT AGGGGCTCAA TGTGTGATTC ACGGGTGCCG
CCATTCGAGC TCCAGATCCA GTATCGCTCG AGCAAGTGAT AAAACATGGA
GCAGGGACGA TCACGTGGTC ACTTGAAAAT TATGTGAGGT CCGGGCGCAC
GATGTACGGC GCGGCGAACT CTCAAACACT CACACAGCCA AAACCGCTTC
GTGTTTCGTCT TTGTTCCAAG CGACCGTGTG GTGTGTTTGT AGTAGTTCGC
CGGCGCCGCA CATCGTCGCC CCGGATCTGA CAAATTAAGC TTTCGTTGCT
TTCCACGAT TGTGCATTTT CTGAGCATGC ACTGAATACT ATGATGGATA
TGTTTGGAGG AAGCATTATT CCAATTTGAT GATAATGGTG TTATTTACAC
TTGTTTTAGC CTTGGCAAGG TACTGCGGCT CAAGAAGATT CTAATAAACA
AGTTGCAGCG GGCACCAAGT TGTATTGCCT TCTCAAGTCA TGAGCAACTC
CTGTGGGGCC ATATCACTAG T
    
```

Figura 88: Secuencia flanqueante de homología 3' en posición 1 (SEQ ID NO: 4)

```

ACTAGTCCAG AACTGGTTGA GTCGGTCAAA CAAGATTGCT TGGCTCAAGC
CTATGCAGTG CATGTTATGA GCCAACACCT GGGTGCCAAT CATGTCGATG
GTGGGGTATG GTTCAGATTC AGTTCATTTA TGTCCTGTTA TTGTGATTTT
GATTGGTAAC ATATTGACAA CCTCGACACT TGGGATCAGA TTCAGTTCAC
TTATGGAAGA AATTGGAGAA TTGTGATAAT TTATCTATAA TCACCCCTAC
TGAATAGAA ATAACATGAC ATCAATGTGC ATGCTATTGG ATTTTGACAC
GAATATGCTT TATTCTATCA TATGTTGGTA ATTCCAGCAG GCAGCAGGCA
CTACTCTTTG GATCCACGTG ACTTGACAAA GAAATCATGC CATCTTTCCA
CAATGCAGGT CCGTGTACGT GTTTCTAGGG ATTTTCTGGA GCTTGTGCGAA
AAGAATGTTT TTAGCAGCCG TCCTGCTGGG AGAGTAAATG CAAGTTCAAT
TGATAACACT GCTGATGCCG CTCTTCTAAT AGCAGACCAC TCTTTATTTT
CTGGTACGTA CTCTATCCCT CTTCTTACCA TAATCTGAAT CTTGTTAAGG
TTTTAAATAT ATGATTGATT AAGTAAAATC CAGAGCTCTA TTCATATCTC
ACGCACTGAT GTTTTGATGA AACGCTTGCA GCAAGACGGT TGCCTGTTAT
TTCTATTTGC ATTAGACAAA CAGTCACTT TGTTTATAAA GGTCTTTGAA
TTTGCACTT TTATAGGTTT AAGTTTGCAA CTGTTACTTA CAACAGCCCA
ATGGGTAGCA TCAAGGTCGA C
    
```

Figura 89: Secuencia flanqueante de homología 5' en posición 2 (SEQ ID NO: 139)

```

GCGGCCGCTA GATAGCAGAT GCAGATTGCT TGCTTCTCTG GTTTGATTTT
TGGAGTCACC ATTTCTGTTT GGTTCGTGTG CCTCAGTGTG TGACAGCAGC
AGATCCTCGA TGGAGATGGA TGGGGTTCTG CAAGCCGCGG ATGCCAAGGA
CTGGGTTTAC AAGGGGAAG GCGCCGCGAA TCTCATCCTC AGCTACACCG
GCACGTCGCC CTCCATGGTA AGCGCTGAGT AGGTTCTTAC TGAGTGTGCA
CGCATCGATC ACTTGACTTT AGGGGCTCAA TGTGTGATTG ACGGGTGCCG
CCATTGCGAG TCCAGATCCA GTATCGCTCG AGCAAGTGAT AAAACATGGA
GCAGGGACGA TCACGTGGTC ACTTGAAAAT TATGTGAGGT CCGGGGCGAC
GATGTACGGC GCGGCCAACT CTCAAACACT CACACAGCCA AAACCGCTTC
GTGTTTCTCT TTGTTCCAAG CGACCGTGTG GTGTGTTTGT AGTAGTTCGC
CGGCGCCGCA CATCGTCGCC CCGGATCTGA CAAATTAAGC TTTGTTGCT
TTTCCACGAT TGTGCATTTT CTGAGCATGC ACTGAATACT ATGATGGATA
TGTTTGGAGG AAGCATTATT CCAATTTGAT GATAATGGTG TTATTTACAC
TTGTTTTTCT CTTGGCAAGG TACTGCGGCT CAAGAAGATT CTAATAAACA
AGTTGCAGCG GGCACCAAGT TGTATTGCCT TCTCAAGTCA TGAGCAACTC
CTGTGGGGCC ATATCCAGA ACTGGTTGAG TCGGTCAAAC AAGATTGCTT
GGCTCAAGCC TATGCAGTGC ATGTTATGAG CCAACACCTG GGTGCCAATA
CTAGT
    
```

Figura 90: Secuencia flanqueante de homología 3' en posición 2 (SEQ ID NO: 140)

```

ACTAGTCATG TCGATGGTGG GGTATGGTTC AGATTCAGTT CATTATATGTC
CTGTTATTGT GATTTTGATT GGTAACATAT TGACAACCTC GACACTTGGG
ATCAGATTCA GTTCACTTAT GGAAGAAATT GGAGAATTGT GATAATTTAT
CTATAATCAC CCCTACTGAA ATAGAAATAA CATGACATCA ATGTGCATGC
TATTGGATTT TGACACGAAT ATGCITTTAT CTATCATATG TTGGTAATTC
CAGCAGGCAG CAGGCACTAC TCTTTGGATC CACGTGACTT GACAAAGAAA
TCATGCCATC TTTCCACAAT GCAGGTCCGT GTACGTGTTT CTAGGGATTT
TCTGGAGCTT GTCGAAAAGA ATGTTCTTAG CAGCCGTCTT GCTGGGAGAG
TAAATGCAAG TTCAATTGAT AACACTGCTG ATGCCGCTCT TCTAATAGCA
GACCACTCTT TATTTTCTGG TACGTA CTCT ATCCCTCTTC TTACCATAAT
CTGAATCTTG TTAAGGTTTA AAATATACGA TTGATTAAGT AAAATCCAGA
GCTCTATTCA TATCTCACGC ACTGATGTTT TGATGAAACG CTTGCAGCAA
GACGGTTGCC TGTTATTTCT ATTTGCATTA GACAAACAGT CACCTTTGTT
TATAAAGGTC FTTGAATTTG CAGTTCTTAT AGGTTTAAGT TTGCAACTGT
TACTTACAAC AGCCCAATGG GTAGCATCAA GATTGTTTTT TTCAGTGATT
CATAACTTAA CTCTTGTTA AACCGCTAGA ACATGGTTGG TGTCTTAAAA
TGCAACTGTT CCTGAGGCCG TAACCTGAAA TCATTGTACG TCGAC
    
```

Figura 91: Secuencia genómica aguas arriba (5') IPP2K de las regiones diana ZFN (SEQ ID NO: 141)

```
TGGACGGAGC GAGAGCCAGA ATTCGACGCT GCGGCGGCG CGTCGCCAAT
ACGCAGCGCG GATGTGGAGC CACATGCAAA CGTGTGTCCG CCCGCGTGGC
GTCCACTCTC CCTCCACGTT TCGGCGTCCT CGTCGCCTTC CTGGGAAATC
TCCAGCTACT GCCCACTGCC CCTTCCCTTC AGTCCCTTTC CCCGGGCTGT
GGTACCAGTA CTAGTACCAG CATCTCTTCA GGCTCCACCA AGCGCAGACA
CCGCAGCAGC GGCAGCAGCA CGATCCGGTG ACCCCCCGCC GCGTCCAGCC
TGCTCCTCCG GTGATCGCCG GACTGGCGGG GTAGGAACCA GCGGAGCGCA
GCCCCCCTCC TTCCGCTGGT AAGAGTGACG CCCGCCCGCT CCTCCCTTCG
CTCGCTTCCT TGCTCTCCG ATTCTGGCGT ACCAGTCTCA CCGCGGCTTG
GGGATTTGAT GCGGAGCTAG TTAACCAGCA GAGC
```

Figura 92: Secuencia genómica aguas arriba (3') IPP2K de las regiones diana ZFN (SEQ ID NO: 142)

```
ATTGTTTTTT TCAGTGATTC ATAACTTAAC TCTTGGTTAA ACCGCTAGAA
CATGGTTGGT GTCTTAAAT GCAACTGGTC CTGAGGCCGT AACCTGAAAT
CATTGTACTT TTCTCTCATT TCTTTAGATA TTCCAAAAC TCTACATTAG
ATGATTTATG TTTGCTTACT TAGTCTTCT TAATCTCAGG CAATCCTAAG
GGTAGCAGCT GCATAGCTGT AGAGATAAAG GTACTTTGCA AGCTTCCTCT
TTTATTCTTA TTTTTCATTT CTTATGTATA TTTCTCCTCA ACCATTTGAC
TTCTTTTCGG CATGCTCTAC CTTGCAGGCC AAATGTGGGT TTCTGCCATC
ATCAGAATAT ATATCAGAAG ATAATACTAT CAAGAAACTA GTAACGAGAT
ATAAGATGCA TCAGCACCTC AAATTTTATC AGGGTGAGGT GTGTAGATTG
GAATGCTTGA TGCCTTGATC CAAGATAAAA TTCCACTCTC TTTTGCGCAC
TTAAAAACA TCCATCGATG ATACAACTT GATCAAAATA CCTTAAGGCT
TGTTATTTAC GGCCTGTTG TAATATTATA CCGTCTCTTG CTTTTTGACA
TCAGGTTGAT TCCCAATACA TTCTTGACCA CATTTCAGAT ATCGAAGACT
AGTGAGTACA ATCCTCTTGA TCTATTTTCT GGGTCAAAAG AGAGAATATG
CATGGCCATC AAGTCCCTTT TCTCAACTC
```