

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 657 067**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**C12N 5/04** (2006.01)

**A01H 5/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.10.2010 E 14151570 (0)**

97 Fecha y número de publicación de la concesión europea: **22.11.2017 EP 2722392**

54 Título: **Proteínas con dedos de zinc modificadas genéticamente que se dirigen a genes vegetales implicados en la biosíntesis de ácidos grasos**

30 Prioridad:

**22.10.2009 US 279528 P**

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

**01.03.2018**

73 Titular/es:

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

72 Inventor/es:

**DEKELVER, RUSSELL;  
GUPTA, MANJU;  
MILLER, JEFFREY C.;  
NOVAK, STEPHEN y  
PETOLINO, JOSEPH F.**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 657 067 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 con dedos de zinc modificadas genéticamente que se dirigen a genes vegetales implicados en la biosíntesis de ácidos grasos

**Campo técnico**

- 5 La descripción se refiere en general a los campos de la modificación genética del genoma y expresión proteica en vegetales. En particular, la presente descripción se refiere a dominios que se unen a ADN modificados genéticamente, por ejemplo, proteínas con dedos de zinc, que se dirigen a genes implicados en la síntesis de ácidos grasos y a métodos de empleo de tales proteínas con dedos de zinc en la modulación de la expresión génica, la inactivación de genes y la modificación génica dirigida, para generar plantas con perfiles de ácidos grasos alterados.

**10 Antecedentes**

- 15 Las dietas ricas en grasas saturadas aumentan las lipoproteínas de baja densidad (LDL), que a su vez provocan el depósito de colesterol en los vasos sanguíneos, una afección previa, estrechamente correlacionada con la aterosclerosis y la enfermedad cardíaca coronaria (Conner et al., *Coronary Heart Disease: Prevention, Complications, and Treatment*, J.B. Lippincott, Philadelphia, 1984 p. 43-64). Por el contrario, se ha mostrado que las dietas ricas en grasas monoinsaturadas reducen las enfermedades cardíacas. El ácido oleico, la única grasa monoinsaturada en la mayoría de los aceites vegetales comestibles, disminuye las LDL tan eficazmente como el ácido linoleico, pero no afecta a los niveles de lipoproteínas de alta densidad (HDL) (Mensink et al. (1989) *New England J. Med.*, 321:436-441). Adicionalmente, las dietas ricas en grasas monoinsaturadas también están correlacionadas con una presión sanguínea sistólica reducida (Williams et al. (1987) *J. Am. Med. Assoc.*, 257:3251-3256, 1987).

- 20 De cara al efecto de los ácidos grasos sobre la dieta y la salud, se han realizado intentos de alterar el perfil de ácidos grasos de plantas utilizadas con fines comestibles e industriales. Sin embargo, los métodos convencionales de alteración de plantas para mejorar el perfil de ácidos grasos se basan en la mutagénesis (por ejemplo, la radiación química, etc.) y/o en el cultivo y requieren mucho tiempo, son laboriosos y no se dirigen específicamente a genes seleccionados. Véase, por ejemplo, el documento de Patente de EE.UU. nº 5.861.187.

- 25 Por ejemplo las tres referencias siguientes se refieren más particularmente a la tioesterasa de ácidos grasos (Fat):

Salas Joaquin et al. ((2002) *ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS* 403(1):25-34) describe la especificidad de sustrato de las tioesterasas de acil-ACP FatA y FatB.

El documento WO 00/09721 A1 se refiere a métodos para incrementar el contenido en estearato en aceite de soja mediante la expresión de la tioesterasa de acil-ACP.

- 30 El documento US 2002/178467 describe secuencias de péptidos de tránsito de plástidos de una tioesterasa de acil-ACP de la *Cuphea* para la orientación eficaz hacia plástidos.

- Recientemente, se han empleado ventajosamente dominios que se unen a ADN modificados genéticamente, tales como dominios de meganucleasa que se unen a ADN y proteínas con dedos de zinc (ZFPs) para modular selectivamente la expresión génica y para alterar de forma dirigida secuencias de genes en plantas (véanse, por ejemplo, los documentos de Patentes de EE.UU. nº 7.262.054, 7.235.354, 7.220.719, 7.001.768 y 6.534.261; los documentos de publicación de Patentes de EE.UU. nº 2008/0182332 y nº de serie 12/284.888). Las proteínas con dedos de zinc (ZFPs) son proteínas que se unen a ADN, ARN y/o proteína, de una manera específica de la secuencia, en virtud de un dominio estabilizado con metal, conocido como dedos de zinc. Véase, por ejemplo, Miller et al. (1985) *EMBO J.* 4:1609-1614; Rhodes et al. (1993) *Sci. Amer.* 268(2):56-65; y Klug (1999) *J. Mol. Biol.* 293:215-218. Las ZFPs se encuentran comúnmente en los factores de transcripción y hasta la fecha, se han identificado más de 10.000 secuencias de dedos de zinc en varios miles de factores de transcripción conocidos o putativos.

- Los dominios que se unen a ADN también se pueden usar con dominios de nucleasas para preparar nucleasas modificadas genéticamente. Por ejemplo, el dominio que se une a ADN de una endonucleasa de migración dirigida se puede alterar para generar nuevas endonucleasas de migración dirigida. Del mismo modo, los dominios de dedos de zinc también se han combinado con dominios de escisión de nucleasas para producir nucleasas de dedos de zinc (ZFNs) para dirigir de forma específica la rotura de una cadena doble en la región de un genoma en la que se desea una modificación (por ejemplo, delección, mutación, recombinación homóloga o inserción de una secuencia exógena) (véase, por ejemplo, el documento de solicitud de Patente de EE.UU. nº 2007/0134796; 2005/0064474; 2008/0182332). Las ZFPs modificadas genéticamente facilitan enormemente la inserción de secuencias exógenas o la modificación de secuencias endógenas en sitios diana específicos en plantas y proporcionan una alteración dirigida de los genomas de plantas con mayor eficacia que los métodos convencionales (véanse, por ejemplo, los documentos de Patentes de EE.UU. nº 7.262.054, 7.235.354, 7.220.719, 7.001.768 y 6.534.261).

- No obstante, sigue habiendo una necesidad de composiciones y métodos para la alteración dirigida de genes implicados en la síntesis de ácidos grasos, con el fin de producir plantas y productos vegetales (por ejemplo, aceites vegetales) que tengan ácidos grasos seleccionados. Mediante la producción de variedades de plantas con niveles

reducidos de grasas saturadas individuales y totales en el aceite de la semilla, se pueden producir productos alimenticios a base de aceite, que contienen menos grasas saturadas. Tales productos beneficiarán la salud pública mediante una reducción de la incidencia de la aterosclerosis y la enfermedad coronaria cardiaca.

### Compendio

5 La presente descripción describe composiciones y métodos para modular la expresión y para una alteración de forma dirigida, en plantas enteras o en células vegetales, uno o varios genes vegetales implicados en la biosíntesis de ácidos grasos, alterando de este modo la composición de los ácidos grasos en las plantas enteras o en las células vegetales. Las plantas enteras o las células vegetales pueden proceder de especies de plantas monocotiledóneas (monocots) o dicotiledóneas (dicots), incluyendo en algunas realizaciones particulares, plantas productoras de aceite y también incluyendo células cultivadas, células en una planta en cualquier etapa de desarrollo y células vegetales que se han retirado de una planta entera y cuyas células (o sus descendientes) se regenerarán formando plantas. Las células vegetales pueden contener una o varias secuencias de genes homeólogos o parálogos, en donde la modificación se puede dirigir a cualquier cantidad de los mismos o a todos ellos por los métodos descritos en esta memoria.

15 En un aspecto, en esta memoria se describe un dominio que se une a ADN (por ejemplo, una proteína con dedos de zinc (ZFP)) que se une específicamente a un gen implicado en una ruta de biosíntesis de ácidos grasos vegetales. En algunas realizaciones, el gen es un gen de *Brassica napus*. En algunos aspectos particulares, el gen de *Brassica napus* puede codificar carboxilasa de acetil-CoA (ACCasa), sintetasas de  $\beta$ -cetoacil-ACP (KAS, por ejemplo, KAS I - KAS IV), tioesterasa B de ácidos grasos (FATB, por ejemplo, FATB1 - FATB5, u otras tioesterasas plastidiales), sintasa de ácidos grasos (FAS), elongasa de ácidos grasos (FAE, por ejemplo, FAE1), tioesterasa A de ácidos grasos (FatA), desaturasa de ácidos grasos (Fad2, Fad3), deshidrogenasa plastidial de G-3-P (GPDH), glicerocinasa (GK), desaturasa de la proteína portadora de estearoil-acilo (S-ACP-DES) e hidrolasa de oleoil-ACP. En algunos aspectos particulares, el gen puede ser un ortólogo o un homólogo de estos genes en otras especies de plantas productoras de aceite. En una realización preferida, la presente invención proporciona una proteína de fusión que comprende un dominio funcional, que es un dominio regulador transcripcional o un dominio de escisión y una proteína con dedos de zinc no presente en la naturaleza que modula la expresión de al menos un gen FatB endógeno vegetal, en donde la proteína con dedos de zinc comprende las regiones de reconocimiento de hélices, ordenadas como se muestra en una fila simple de la Tabla 10B. Más preferentemente, el gen FatB está contenido dentro de una planta Brassica. En todavía una realización preferida, la proteína con dedos de zinc se une a un sitio diana como se muestra en la Tabla 2.

En todavía un aspecto adicional, también se proporcionan proteínas de fusión que comprenden cualquiera de los dominios que se unen a ADN (por ejemplo, ZFPs) descritos en este documento. En ciertas realizaciones, la proteína de fusión comprende una proteína con dedos de zinc y un dominio regulador de la transcripción (por ejemplo, dominio de activación o represión), también conocido como un ZFP TF. En ciertos aspectos, la proteína de fusión comprende una ZFP y un dominio de nucleasa para formar una nucleasa con dedos de zinc (ZFN) que escinde en una región genómica de interés con el gen diana. En ciertas realizaciones, la ZFN comprende un polipéptido de fusión que comprende un dominio de unión de dedos de zinc modificado genéticamente que tiene especificidad para un gen implicado en una vía de biosíntesis de ácidos grasos vegetales (por ejemplo, un gen que codifica ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, FATB5, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK o S-ACP-DES) y un dominio de escisión, y/o uno o varios polipéptidos de fusión que comprenden un dominio de unión de dedos de zinc modificado genéticamente y un medio dominio de escisión. En ciertos aspectos, los dominios de unión de dedos de zinc se unen a un sitio diana, como se muestra en la Tabla 2 o la Tabla 10A. En ciertos aspectos, los dominios de unión de dedos de zinc comprenden una secuencia seleccionada a partir del grupo que consiste en proteínas con dedos de zinc que comprenden los dominios de reconocimiento (por ejemplo, una fila simple) que se muestran en la Tabla 1 o la Tabla 10B. Los dominios de escisión y los medio dominios de escisión se pueden obtener, por ejemplo, a partir de varias endonucleasas de restricción y/o endonucleasas de migración dirigida. En un aspecto, los medio dominios de escisión se obtienen a partir de una endonucleasa de restricción de tipo IIS (por ejemplo, FokI).

En una realización preferida, la presente invención proporciona un polinucleótido que codifica una o varias proteínas con dedos de zinc que comprenden las regiones de reconocimiento de hélices mostradas en una fila simple de la Tabla 10B o proteínas de fusión que comprenden un dominio funcional que es un dominio regulador transcripcional o un dominio de escisión y una proteína con dedos de zinc que no está presente en la naturaleza que comprende las regiones de reconocimiento de hélices mostradas en una fila simple de la Tabla 10B, en donde dicha proteína de fusión modula la expresión de al menos un gen FatB endógeno.

En otros aspectos, se describen en esta memoria polinucleótidos que codifican cualquiera de los dominios que se unen a ADN (por ejemplo, proteínas con dedos de zinc) y/o proteínas de fusión descritas en el presente documento. En ciertos aspectos, se describe en este documento un vector de expresión de ZFP que comprende un polinucleótido, que codifica una o varias ZFPs descritas en este documento, ligado funcionalmente a un promotor. En una realización, una o varias de las ZFPs son ZFNs.

Las ZFPs y las proteínas de fusión que comprenden estas ZFPs se pueden unir a y/o pueden escindir un gen impli-

5 cado en la síntesis de ácidos grasos dentro de la región codificadora del gen o en una secuencia no codificadora dentro o próxima al gen, tal como, por ejemplo, una secuencia líder, una secuencia remolque o un intrón, o dentro de una región no transcrita, ya sea aguas arriba o aguas abajo de la región codificadora. En ciertos aspectos, las ZFPs o ZFNs se unen y/o escinden una secuencia codificadora o una secuencia reguladora de un gen implicado en la biosíntesis de ácidos grasos.

10 En otro aspecto, en este documento se describen composiciones que comprenden una o varias proteínas, proteínas de fusión o polinucleótidos como se describen en el presente documento. Las células vegetales pueden contener un gen diana único o múltiples copias parálogas del mismo gen. Por lo tanto, las composiciones pueden comprender una o varias proteínas que contienen ZFP (y polinucleótidos que las codifican) que se dirigen a uno o varios genes implicados en la síntesis de ácidos grasos en una célula vegetal. Las ZFPs se pueden dirigir a todos los genes parálogos o homeólogos y a genes parálogos o homeólogos seleccionados particulares en una célula vegetal o una combinación de los mismos.

15 En otro aspecto, se proporciona en este documento una célula hospedadora vegetal que comprende una o varias proteínas o polinucleótidos (por ejemplo, vectores de expresión de ZFP) como se describe en el presente documento. Para los polinucleótidos, la célula hospedadora vegetal se puede transformar de manera estable o de forma transitoria o una combinación de las mismas, con uno o varios vectores de expresión de ZFP. En una realización, uno o varios de los vectores de expresión de ZFP expresan una o varias ZFNs en la célula hospedadora vegetal. En otra realización, uno o varios de los vectores de expresión de ZFP expresan uno o varios ZFP TFs en la célula hospedadora vegetal.

20 En otro aspecto, se describe en esta memoria un método para modificar uno o varios genes implicados en la biosíntesis de ácidos grasos en una célula vegetal, comprendiendo el método: (a) introducir, en la célula vegetal, uno o varios vectores de expresión que codifican uno o varios ZFP TFs que se unen a un sitio diana en uno o varios de los genes implicados en la biosíntesis de ácidos grasos, en condiciones tales que los ZFP TFs se expresan y se modifica la expresión de uno o varios de los genes, y más particularmente, se expresan una o varias de las proteínas con  
25 dedos de zinc que comprenden las regiones de reconocimiento de hélices mostradas en una fila simple de la Tabla 10B, y se modifica la expresión de uno o varios de los genes FatB. La modificación comprende modular la expresión de al menos un gen FatB. La modulación puede ser una activación o represión de la expresión de al menos un gen FatB. En ciertos aspectos, al menos un sitio diana está en una ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, FATB5, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK y/o el o los genes S-ACP-DES. En otros  
30 aspectos, se modula más de un gen implicado en la biosíntesis de ácidos grasos. En cualquiera de los métodos para la modificación de uno o varios genes implicados en la biosíntesis de ácidos grasos como se describe en el presente documento, los métodos dan como resultado células vegetales con un contenido en ácidos grasos modificado, por ejemplo, una reducción de la cantidad de grasas saturadas en las células vegetales. En algunos aspectos, el contenido en ácidos grasos modificados en las células vegetales da como resultado un contenido en ácidos grasos modifi-  
35 cados en las semillas de la planta, por ejemplo, una reducción de la cantidad de grasas saturadas en las semillas de la planta. En algunos aspectos, la planta es una planta productora de aceite que permite ácidos grasos modificados en las semillas de la planta productora de aceite (p. ej., una reducción de la cantidad de grasas saturadas). En algunos aspectos particulares, la planta es una planta *Brassica napus* que permite ácidos grasos modificados en las semillas de la planta *Brassica napus* (por ejemplo, grasas saturadas reducidas).

40 En otro aspecto, se proporciona en esta memoria un método para escindir uno o varios genes implicados en la biosíntesis de ácidos grasos en una célula vegetal, comprendiendo el método: (a) introducir, en la célula vegetal, uno o varios vectores de expresión que codifican una o varias nucleasas (por ejemplo, ZFNs) que se unen a un sitio diana en uno o varios de los genes implicados en la biosíntesis de ácidos grasos, en condiciones tales que las nucleasas (por ejemplo, la o las ZFNs) se expresan y uno o varios de los genes se escinden, y más particularmente al menos  
45 un gen FatB se escinde. En ciertos aspectos, al menos un sitio diana está en un gen que codifica ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, FATB5, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK y/o S-ACP-DES. En otros aspectos, más de un gen implicado en la biosíntesis de ácidos grasos se escinde. Además, en cualquiera de los métodos descritos en el presente documento, la escisión de uno o varios de los genes puede dar como resultado una delección, adición y/o sustitución de nucleótidos en la región escindida.

50 En aún otro aspecto, se describe en esta memoria un método para introducir una secuencia exógena en el genoma de una célula vegetal, comprendiendo el método las etapas de: (a) poner en contacto la célula con una secuencia exógena (vector donante); y (b) expresar una o varias nucleasas (por ejemplo, nucleasas con dedos de zinc) como se describe en el presente documento en la célula, en donde una o varias de las nucleasas escinden el ADN cromosómico; de tal modo que la escisión del ADN cromosómico en la etapa (b) estimula la incorporación del vector donante en el genoma mediante recombinación homóloga. En ciertos aspectos, una o varias de las nucleasas son  
55 fusiones entre el dominio de escisión de una endonucleasa de restricción de tipo IIs y un dominio de unión a dedos de zinc modificado genéticamente. En otros aspectos, la nucleasa comprende una endonucleasa de migración dirigida, por ejemplo, una endonucleasa de migración dirigida con un dominio que se une a ADN modificado. En cualquiera de los métodos descritos en el presente documento, la secuencia exógena puede codificar un producto proteico.  
60

En un aspecto adicional más, también se proporciona una célula vegetal transgénica o no transgénica obtenida de

acuerdo con cualquiera de los métodos descritos en este documento.

En otro aspecto, se proporciona en este documento una planta que comprende una célula vegetal transgénica o no transgénica obtenida como se describe en el presente documento.

5 En otro aspecto, se proporciona en este documento una semilla de una planta que comprende una célula vegetal transgénica o no transgénica obtenida como se describe en el presente documento.

En otro aspecto, se describe en este documento aceite procedente de una semilla extraída a partir de una planta que comprende una célula vegetal transgénica o no transgénica obtenida como se describe en el presente documento.

### Breve descripción de los dibujos

10 La **Figura 1** es un esquema que representa rutas de biosíntesis de ácido graso en canola (*B. napus*). Este se adaptó según John Shanklin, Brookhaven National Laboratory, Upton, NY (Thelen y Ohlrogge, 2002, Metabolic engineering 4:12-21).

15 La **Figura 2** es un esquema que representa las ubicaciones de los sitios diana dentro del gen KASII para diversos ejemplares de ZFP TFs dirigidos a KASII. Los números indican los sitios diana para las ZFPs contenidas dentro de las estructuras artificiales mostradas en la Tabla III. Los diseños de ZFP se muestran en las Tablas I y II.

La **Figura 3** es un gráfico que representa la expresión del ARNm de KASII en hojas de plantas T0 transformadas con la estructura artificial 4695 de ZFP TF que activa KASII, tal y como se sometió a ensayo mediante qRT-PCR. Se empleó un promedio de 27 plantas de control para calcular el nº de veces de regulación al alza que se muestra en el gráfico. En algunos eventos, se observó una regulación al alza del ARNm de KASII de más de tres veces.

20 La **Figura 4** es un gráfico que representa relaciones entre la expresión de KASII-ZFP TF/tubulina en hojas de plantas T1, detectada a través de qRT-PCR. Tres eventos se compararon junto con los nulos correspondientes.

La **Figura 5** es un gráfico que representa la relación promedio entre la expresión del ARNm de KASII/tubulina en plantas T1 nulas que contienen y segregan ZFP TF de cada uno de tres eventos, tal y como se determina mediante qRT-PCR. Se observó una regulación al alza de 2-3 veces del ARNm de KASII en estas hojas de plantas T1.

25 Las **Figuras 6A y 6B** son gráficos, dibujados con el programa informático estadístico JMP para el análisis unidireccional de ácido(s) graso(s) diana, que representan una disminución consistente y significativa ( $p = <0,001$ ) de C16 total (6A) y una disminución en las relaciones entre C16 total/C18 total (6B) en plantas nulas hermanas y positivas para ZFP TF dentro de cada evento. C16 total comprendía contenidos de C16:0 y C16:1, y C18 total comprendía contenidos de C18:0, 18:1, 18:2 y 18:3.

30 La **Figura 7** muestra una alineación de la secuencia de AF318307 y AF244520. El sombreado indica las regiones de homología exacta.

35 La **Figura 8** representa una alineación de la secuencia de SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 46, SEQ ID NO: 30, AC189461 y BH723504. Las secuencias de cebadores directos e inversos (SEQ ID NOs 28 y 29) para la amplificación del ADNc de la sintetasa II de  $\beta$ -cetoacil-ACP se destacan mediante líneas de trazos por encima de la secuencia correspondiente. El sombreado indica las regiones de homología exacta.

La **Figura 9** representa la expresión de los genes FatB4 y FatB5 en líneas de callos transgénicos de *B. napus* para los diferentes diseños de ZFP TF presentes en las estructuras artificiales pDAB4690 - pDAB4692. Se analizaron 17 líneas de control y 25 líneas para cada una de las estructuras artificiales de ZFP. Barras negras = expresión del ARNm de FatB4; barras grises = expresión del ARNm de FatB5.

40 La **Figura 10** muestra la expresión de FatB4 y FatB5 en hojas de plantas transgénicas T0 de *B. napus* analizada con qRT-PCR. Tres estructuras artificiales que comprendían ZFP TFs, pDAB4689 - pDAB4691 se sometieron a ensayo en plantas transgénicas. El número total de plantas transgénicas T0 independientes analizadas para este experimento era 40, 62, 41 y 22 para pDAB4689 - pDAB4691 y pDAB8210, respectivamente. El control Nex710 estaba compuesto de 10 plantas. Noventa y siete por ciento de los eventos transgénicos para 3 estructuras artificiales de ZFP TF eran positivos para la expresión de ZFP TF, tal como se determinó mediante el ensayo de la expresión de ZFP TF (Ejemplo 8.3). Se obtuvieron resultados similares cuando se utilizó la expresión de ARNm de tubulina natural como referencia para normalizar las expresiones del gen FatB. Barra negra = expresión de FatB4; barra gris = expresión de FatB5.

45 La **Figura 11** es un gráfico que muestra una relación lineal entre los genes FatB y la expresión de tubulina cuando se analizaron a partir de plantas T1. Cuadrados negros = expresión de FatB4; rombos grises = expresión de FatB5.

50 La **Figura 12** muestra una qRT-PCR para la expresión de FatB4 y FatB5 en hojas de plantas T1 transgénicas para la estructura artificial de ZFP TF, pDAB4691. Cuadrados negros = expresión de FatB4; rombos grises = expresión de FatB5.

La **Figura 13** es un análisis unidireccional del contenido en C18:0 por muestra, utilizando el programa informático estadístico JMP.

La **Figura 14** muestra un análisis unidireccional del contenido total en C16/C18 por muestra, usando el programa informático estadístico JMP.

5 La **Figura 15** representa un análisis de los perfiles de T1 FA de una semilla madura que comprende 4 eventos de estructuras artificiales de ZFP TF, mostrado con el programa informático estadístico JMP. Los paneles A y B representan un análisis unidireccional por muestra (estructura artificial) del contenido en C18:0 y el contenido en C16:0/C18:0, respectivamente.

10 La **Figura 16** representa un análisis de los perfiles T1 FA de una semilla madura que comprende 4 eventos de estructuras artificiales de ZFP TF, mostrado con el programa informático estadístico JMP. Los paneles A y B representan un análisis unidireccional por muestra (estructura artificial) del contenido en C14:0 y C16:1, respectivamente, destacando el atributo diferenciador de la estructura artificial pDAS5227. El análisis de los perfiles T1 FA de una semilla madura que comprende 4 eventos de estructuras artificiales de ZFP TF, se muestra con el programa informático estadístico JMP. Los paneles A y B representan un análisis unidireccional por muestra (estructura artificial) del contenido en C14:0 y C16:1, respectivamente, destacando el atributo diferenciador de la estructura artificial pDAS5227.

15 La **Figura 17** muestra la regulación a la baja del ARNm de FatB5 (barras negras) en una semilla T2 inmadura transformada con la estructura artificial de ZFP TF, pDAS5227. La expresión de ZFP TF está representada por la expresión de ERF3 (barras grises). Se analizaron 25 semillas inmaduras DAF procedentes de cuatro plantas T1 nulas, tres heterocigotas (5227\_12ZF-1) y cuatro homocigotas (5227\_12ZF-2) del evento 5227-12.

20 La **Figura 18** muestra la regulación a la baja del ARNm de FatB5 (barras negras) en una semilla T2 inmadura transformada con la estructura artificial de ZFP TF, pDAS5212. La expresión del ARNm de ZFP TF está representada por la expresión de KRAB1 (barras grises). Se analizaron 25 semillas inmaduras DAF procedentes de cinco plantas nulas, tres heterocigotas y cuatro homocigotas del evento 5212-4.

25 La **Figura 19** muestra un análisis unidireccional de ácidos grasos saturados totales (sats.) mediante cigosidad en la semilla T2 del evento de ZFP TF, 5227-12. Las semillas T2 obtenidas a partir de plantas parentales T1 nulas, heterocigotas y homocigotas están marcadas como 5227-12ZF nula, 5227-12ZF(1) y 5227-12ZF(2), respectivamente.

30 La **Figura 20** muestra un análisis unidireccional de ácidos grasos saturados totales (sats.) mediante cigosidad en la semilla T2 del evento de ZFP TF, 5212-4. Las semillas T2 obtenidas a partir de plantas parentales T1 nulas, heterocigotas y homocigotas están marcadas como 5212-4ZF nula, 5212-4ZF(1) y 5212-4(2), respectivamente.

### Descripción detallada

35 En el presente documento se describen composiciones y métodos útiles para la modulación de la expresión y la escisión y la alteración dirigida de uno o varios genes implicados en la síntesis de ácidos grasos en las plantas. La regulación de tales genes se puede modular, por ejemplo, mediante el uso de factores de transcripción de ZFP modificados genéticamente o modificando regiones reguladoras de genes. Los genes se pueden alterar, por ejemplo, mediante una escisión dirigida seguida de recombinación homóloga intracromosómica o mediante una escisión dirigida seguida de recombinación homóloga entre un polinucleótido exógeno (que comprende una o varias regiones de homología con la secuencia de nucleótidos del gen) y una secuencia genómica.

40 Las secuencias genómicas incluyen aquellas presentes en cromosomas, episomas, genomas de orgánulos (por ejemplo, mitocondrias, plástidos), cromosomas artificiales y cualquier otro tipo de ácido nucleico presente en una célula de ese tipo, tal como, por ejemplo, secuencias amplificadas, cromosomas dobles diminutos 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, delecciones, translocaciones, reordenamientos y/o mutaciones puntuales. Una secuencia genómica puede comprender también uno entre  
45 una serie de alelos diferentes.

Las composiciones descritas en este documento comprenden una o varias ZFPs que comprenden dominios que se unen a dedos de zinc modificados genéticamente, polinucleótidos que codifican estos polipéptidos y combinaciones de ZFPs y polinucleótidos que codifican ZFPs. Un dominio que se une a dedos de zinc puede comprender uno o  
50 varios dedos de zinc (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de zinc) y puede estar modificado genéticamente para unirse a cualquier secuencia genómica de un gen, incluyendo secuencias reguladoras ligadas funcionalmente al gen, que participan en la biosíntesis de ácidos grasos.

55 Las ZFPs tal y como se describen en esta memoria, se pueden utilizar para regular la expresión génica, ya sea a través de la activación o la represión de la transcripción génica. Las ZFPs que comprenden fusiones de dominios de dedos de zinc ligados a dominios reguladores, se pueden construir para crear factores de transcripción quiméricos que activan o inhiben la transcripción. Las ZFPs también se pueden utilizar para una escisión dirigida de una región genómica de interés mediante la unión de dominios de dedos de zinc con dominios de escisión de nucleasas (o

medio dominios de escisión) para producir nucleasas de dedos de zinc. Por lo tanto, mediante la identificación de una región genómica diana de interés en la que se desea una regulación, escisión o recombinación de los genes, de acuerdo con los métodos descritos en el presente documento, se puede construir una proteína con dedos de zinc que comprende una o varias proteínas de fusión que comprenden uno o varios dominios reguladores y/o dominios de escisión (o medio dominios de escisión) ligados a un dominio de dedos de zinc modificado genéticamente para reconocer una secuencia de genes en esa región genómica. La presencia de una ZFP de este tipo que comprende una proteína (o proteínas) de fusión en una célula, dará lugar a la unión de la o las proteínas de fusión con sus sitios de unión y una regulación o escisión alterada dentro o cerca de la región genómica. Además, si una región genómica se escinde y un polinucleótido exógeno homólogo a esa región genómica también está presente en la célula, se produce una recombinación homóloga con una tasa elevada entre la región genómica y el polinucleótido exógeno.

Como se muestra en la Figura 1, hay varios genes que están implicados en la biosíntesis de ácidos grasos. Por tanto, las composiciones descritas en esta memoria se pueden dirigir a uno o varios de esos genes en una célula vegetal, incluyendo, pero no limitados a los genes de ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, FATB5, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK o S-ACP-DES y ortólogos, parálogos y homeólogos de esos genes. Por ejemplo, 1, 2, 3, 4, 5 o más genes implicados en la biosíntesis de ácidos grasos pueden ser el blanco de proteínas (por ejemplo, ZFPs) descritas en esta memoria. Por tanto, una o varias ZFPs o vectores de expresión que codifican ZFPs de diferentes especificidades, se pueden combinar para dirigirse a los genes deseados de interés en una planta.

#### General

La práctica de los métodos, así como la preparación y el uso de las composiciones descritas en la presente memoria emplean, salvo que se indique otra cosa, técnicas convencionales en biología molecular, bioquímica, estructura y análisis de cromatina, química computacional, cultivo celular, ADN recombinante y campos relacionados que pertenecen a las habilidades de la técnica. Estas técnicas se explican a fondo en la bibliografía. Véanse, por ejemplo, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, segunda edición, Cold Spring Harbor Laboratory y Press, 1989 y tercera edición, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 y actualizaciones periódicas; la serie 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, compiladores), Academic Press, San Diego, 1999; y METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, compilador) Humana Press, Totowa, 1999.

#### Definiciones

Las expresiones "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan de forma intercambiable y hacen referencia a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en forma monocatenaria o bicatenaria. Para los propósitos de la presente descripción, estos términos no deben entenderse como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en los restos de la base, el azúcar y/o el fosfato (por ejemplo, estructuras principales de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de apareamiento de bases; es decir, un análogo de A se emparejará con T.

Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable para hacer referencia a un polímero de residuos de aminoácidos. El término también se aplica a polímeros de aminoácidos en donde uno o varios aminoácidos son análogos químicos o derivados modificados de un aminoácido de origen natural correspondiente.

"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 tienen que ser específicos de secuencia (por ejemplo, contactos con residuos fosfato en una estructura principal de ADN), siempre que la interacción en su totalidad sea específica de secuencia. Tales interacciones se caracterizan generalmente por una constante de disociación ( $K_d$ ) de  $10^{-6} \text{ M}^{-1}$  o inferior. "Afinidad" se refiere a la fuerza de la unión: correlacionándose una afinidad de unión aumentada con una  $K_d$  inferior.

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 que se une a ADN), una molécula de ARN (una proteína que se une a ARN) y/o una molécula proteica (una proteína que se une a proteína). En el caso de una proteína que se une a proteína, puede unirse a sí misma (para formar homodímeros, homotrimeros, etc.) y/o puede unirse a una o varias moléculas de una proteína o proteínas diferentes. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas con dedos de zinc tienen actividad de unión a ADN, unión a ARN y unión a proteína.

Una "proteína que se une a ADN de dedos de zinc" (o dominio de unión) es una proteína, o un dominio dentro de una proteína más grande, que se une a ADN de un modo específico de secuencia a través de uno o varios dedos de zinc, que son regiones de una secuencia de aminoácidos dentro del dominio de unión, cuya estructura se estabiliza a través de la coordinación de un ion zinc. La expresión proteína que se une a ADN de dedos de zinc, frecuentemen-

te se abrevia como proteína con dedos de zinc o ZFP.

Los dominios de unión de dedos de zinc pueden "diseñarse" para que se unan a una secuencia de nucleótidos pre-determinada (p. ej., una secuencia diana en cualquier gen implicado en la biosíntesis de ácidos grasos). Ejemplos no limitantes de métodos para diseñar proteínas de dedos de zinc son el diseño y la selección. Una proteína con dedos  
5 de zinc diseñada es una proteína que no existe en la naturaleza cuyo diseño/composición es el resultado principal-mente de criterios racionales. Los criterios racionales para el diseño incluyen la aplicación de normas de sustitución y algoritmos informatizados para procesar información en una base de datos que almacena información de diseños y datos de unión existentes de ZFP. Véanse, por ejemplo, los documentos de patentes de EE.UU. 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; 10 WO 02/016536 y WO 03/016496; y los documentos de Patentes de EE.UU. 6.746.838; 6.866.997; y 7.030.215. Por lo tanto, una proteína con dedos de zinc "diseñada" o "de origen no natural" con dedos de zinc es una en la que uno o varios de los componente de dominios (hélices de reconocimiento) de unión a ADN de dedos de zinc no son de origen natural y han sido diseñados para unirse a una sitio diana seleccionado previamente.

Una proteína con dedos de zinc "seleccionada" es una proteína no hallada en la naturaleza cuya producción es el resultado principalmente de un procedimiento empírico, tal como presentación en fagos, retención por interacción o 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.

El término "secuencia" se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser ADN o ARN; puede ser lineal, circular o ramificada y puede ser monocatenaria o bicatenaria. La expresión "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede tener cualquier longitud, por ejemplo, entre 2 y 25.000 nucleótidos de longitud (o cualquier valor entero entre ellos o por encima de ellos), preferiblemente entre aproximadamente 100 y 5.000 nucleótidos de longitud (o cualquier número entero entre ellos), más preferiblemente entre aproximadamente 200 y 2.500 nucleótidos de longitud.

Una "secuencia homóloga" se refiere a una primera secuencia que comparte un grado de identidad de secuencia con una segunda secuencia, y cuya secuencia puede ser idéntica a la de la segunda secuencia. Una "secuencia no idéntica homóloga" se refiere a una primera secuencia que comparte un grado de identidad de secuencia con una segunda secuencia, pero cuya secuencia no es idéntica a la de la segunda secuencia. Por ejemplo, un polinucleótido que comprende la secuencia de tipo silvestre de un gen mutante, es homólogo y no idéntico a la secuencia del gen mutante. En ciertas realizaciones, el grado de homología entre las dos secuencias es suficiente para permitir una 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 de un único nucleótido (por ejemplo, para la corrección de una mutación puntual genómica por recombinación homóloga dirigida) o tan larga como de 10 o más kilobases (por ejemplo, para la inserción de un gen en un sitio predeterminado en un cromosoma). Dos polinucleótidos que comprenden las secuencias no idénticas homólogas no tienen que tener 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.

Las técnicas para determinar la identidad de la secuencia de ácidos nucleicos y aminoácidos son conocidas en la técnica. Típicamente, tales técnicas incluyen determinar la secuencia de nucleótidos del ARNm de un gen y/o determinar la secuencia de aminoácidos codificada por la misma, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también pueden determinarse y compararse de este modo. En general, la identidad se refiere a una correspondencia exacta nucleótido a nucleótido o aminoácido a aminoácido de dos secuencias de polinucleótidos o polipéptidos, respectivamente. Se pueden comparar dos o más secuencias (polinucleótidos o aminoácidos) determinando su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácido nucleico o aminoácidos, es la cantidad de coincidencias exactas entre dos secuencias alineadas, dividida por la longitud de la secuencia más corta y multiplicada por 100. Una alineación aproximada para secuencias de ácido nucleico se proporciona por 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 valoración desarrollada por Dayhoff. *Atlas of Protein Sequences and Structure*. M.O. Dayhoff compilador, 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 por el Genetics Computer Group (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 Genetics Computer Group, Madison, WI). Un método preferido para establecer el porcentaje de identidad en el contexto de la presente descripción, es usar el paquete MPSRCH de programas con derechos de autor de la University of Edinburgh, desarrollado por John F. Collins y Shane S. Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). Para este paquete de contenidos puede emplearse el algoritmo de Smith-Waterman, en donde se usan parámetros por defecto para la tabla de puntuaciones (por ejemplo, penalización de 12 por apertura de hueco, penalización de 1 por extensión de hueco, y de seis por un hueco). A partir de los datos generados el valor "Coincidencia" refleja la identidad de secuencia. En la técnica se conocen en líneas generales otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias, por ejemplo, otro programa de alineación es BLAST, usado con parámetros

por defecto. Por ejemplo, pueden usarse BLASTN y BLASTP usando los siguientes parámetros por defecto: código genético = convencional; filtro = ninguno; hebra = ambas; punto de corte = 60; expectativa = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; clasificado por = HIGH SCORE; Bases de datos = no redundantes, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Pueden hallarse detalles de estos programas en internet. Con respecto a secuencias descritas en la presente memoria, el intervalo de grados deseados de identidad de secuencia es de aproximadamente el 35% al 100% y cualquier valor entero entre los mismos. Típicamente, los porcentajes de identidad entre secuencias son de al menos el 35%-40%; 40%-45%; 45%-50%; 50%-60%; 60%-70%; 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, aún más preferiblemente 95% y lo más preferiblemente una identidad de secuencia del 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, seguida por digestión con una o más nucleasas específicas de una hebra, y la determinación del tamaño de los fragmentos digeridos. Dos ácidos nucleicos, o dos secuencias polipeptídicas son sustancialmente homólogas entre sí cuando las secuencias muestran al menos aproximadamente el 70%-75%, preferiblemente 80%-82%, más preferiblemente 85%-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y mucho más preferiblemente 98% de identidad de secuencia sobre una longitud definida de las moléculas, determinada usando los métodos anteriores. Como se usa en la presente memoria, sustancialmente homólogas 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, definidas para ese sistema particular. La definición de las condiciones apropiadas de hibridación pertenece a las habilidades de la técnica. Véase, por ejemplo, Sambrook et al., supra; *Nucleic Acid Hybridization: A Practical Approach*, compiladores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

La hibridación selectiva de dos fragmentos de ácido nucleico puede determinarse del siguiente modo. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta a la eficacia y la fuerza de los eventos de hibridación entre tales moléculas. Una secuencia de ácido nucleico parcialmente idéntica inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica con una molécula diana. La inhibición de la hibridación de la secuencia completamente idéntica puede evaluarse usando ensayos de hibridación que son bien conocidos en la técnica (por ejemplo, transferencia Southern (ADN), transferencia Northern (ARN), 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 grados variables de selectividad, por ejemplo, usando condiciones que varían de baja a alta rigurosidad. Si se emplean condiciones de baja rigurosidad, puede evaluarse la ausencia de unión no específica usando una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda que tiene menos de aproximadamente el 30% de identidad de secuencia con la molécula diana), de modo que, en ausencia de eventos de unión no específica, la sonda secundaria no se hibridará con la diana.

Cuando se utiliza un sistema de detección basado en hibridación, se elige una sonda de ácido nucleico que es complementaria a una secuencia de ácido nucleico referencia, y después mediante una selección de condiciones apropiadas, la sonda y la secuencia de referencia se hibridaran de forma selectiva, o se unirán, entre sí para formar una molécula dúplex. Una molécula de ácido nucleico que es capaz de hibridarse de forma selectiva con una secuencia de referencia en condiciones de hibridación moderadamente rigurosas, típicamente se hibridará en condiciones que permitan la detección de una secuencia de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tiene al menos aproximadamente un 70% de identidad de secuencia con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación rigurosas típicamente permiten la detección de secuencias de ácido nucleico diana de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencia de más de aproximadamente el 90-95% con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de la sonda/secuencia de referencia, en donde la sonda y la secuencia de referencia tienen un grado específico de identidad de secuencia, pueden determinarse de forma conocida en la técnica (véase, por ejemplo, *Nucleic Acid Hybridization: A Practical Approach*, compiladores B.D. Hames y S.J. Higgins, (1985) Oxford; Washington, DC; ERL Press).

Las condiciones de hibridación son bien conocidas por los expertos en la técnica. Rigurosidad de hibridación se refiere al grado en el cual las condiciones de hibridación desfavorecen la formación de híbridos que contienen nucleótidos desapareados, correlacionándose una mayor rigurosidad con una inferior tolerancia hacia híbridos desapareados. Los factores que afectan a la rigurosidad de la hibridación son bien conocidos por los expertos en la técnica e incluyen, pero sin limitarse a, la temperatura, el pH, la 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 superiores, con fuerza iónica inferior y menores concentraciones de disolvente.

Con respecto a las condiciones de rigurosidad para la hibridación, es bien sabido en la técnica que pueden emplearse numerosas condiciones equivalentes para establecer una rigurosidad particular variando, por ejemplo, los siguientes factores: la longitud y la naturaleza de las secuencias, la composición de bases de las diversas secuencias, las concentraciones de sales y otros componentes de la solución de hibridación, la presencia o ausencia de agentes bloqueantes en las soluciones de la hibridación (por ejemplo, sulfato de dextrano y polietilenglicol), la temperatura de reacción de la hibridación y parámetros temporales, así como, variando las condiciones lavado. La selección de un

conjunto particular de condiciones de hibridación se hace siguiendo métodos convencionales en la técnica (véase, por ejemplo, Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, segunda edición, (1989) Cold Spring Harbor, N.Y.).

5 "Recombinación" se refiere a un procedimiento de intercambio de información genética entre dos polinucleótidos. Para los propósitos de esta descripción, "recombinación homóloga (HR)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de la cadena doble en células. Este proceso requiere homología de la secuencia de nucleótidos, usa una molécula "donante" para reparar el molde de una molécula "diana" (es decir, la que ha experimentado la rotura de la cadena doble), y se conoce de forma variada como "conversión génica no cruzada" o "conversión génica a corto plazo", porque conduce a la transferencia de información genética del donante a la diana. Sin el deseo de limitación con ninguna teoría particular, dicha transferencia puede implicar la corrección de desapareamientos de ADN heterodúplex que se forma entre la diana rota y el donante, y/o "reasociación de hebra dependiente de síntesis", en la que el donante se usa para resintetizar una información genética que llegará a formar parte de la diana, y/o procesos relacionados. Dicha HR especializada frecuentemente provoca una alteración de la secuencia de la molécula diana, de modo que parte o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

10 "Escisión" se refiere a la rotura de la estructura principal covalente de una molécula de ADN. La escisión puede iniciarse por una diversidad de métodos incluyendo, pero sin limitarse a, hidrólisis enzimática o química de un enlace fosfodiéster. Es posible tanto una escisión monocatenaria como una escisión bicatenaria, y la escisión bicatenaria puede suceder como resultado de dos eventos de escisión monocatenaria distintos. La escisión del ADN puede provocar la producción de extremos romos o extremos escalonados. En ciertas realizaciones, se usan polipéptidos de fusión para una escisión dirigida del ADN bicatenario.

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

25 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 (preferiblemente actividad de escisión bicatenaria).

30 "Cromatina" es la estructura nucleoproteica que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteínas, incluyendo histonas y proteínas cromosómicas no histonas. La mayor parte de la cromatina celular eucariota existe en forma de nucleosomas, en donde un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociadas con un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; un ADN enlazador (de longitud variable dependiendo del organismo) se extiende entre núcleos de nucleosoma. Una molécula de histona H1 generalmente se asocia con el ADN enlazador. Para los propósitos de la presente descripción, el término "cromatina" se entiende que incluye todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye cromatina tanto cromosómica como episómica.

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

40 Un "episoma" es un ácido nucleico replicante, un complejo nucleoproteico u otra estructura que comprende un ácido nucleico que no forma parte del cariotipo cromosómico de una célula. Ejemplos de episomas incluyen plásmidos y ciertos genomas virales.

45 Una "región accesible" es un sitio en la cromatina celular en donde un sitio diana presente en el ácido nucleico puede unirse con una molécula exógena que reconoce el sitio diana. Sin el deseo de limitarse con ninguna teoría particular, se cree que una región accesible es una que no está empaquetada en una estructura nucleosómica. La estructura distinta de una región accesible puede detectarse frecuentemente por su sensibilidad a sondas químicas y enzimáticas, por ejemplo, nucleasas.

Un "sitio diana" o "secuencia diana" es una secuencia de ácido nucleico que define una porción de un ácido nucleico a la cual se unirá una molécula de unión, con la condición de que existan suficientes condiciones para la unión. Por ejemplo, la secuencia 5'-GAATTC-3' es un sitio diana para la endonucleasa de restricción *Eco RI*.

50 Una molécula "exógena" es una molécula que no está normalmente presente en una célula, pero que puede introducirse en una célula por uno o más métodos genéticos, bioquímicos u otros métodos. La "presencia normal en la célula" se determina con respecto a la fase de desarrollo y las condiciones ambientales particulares de la célula. Por tanto, por ejemplo, una molécula que está presente solamente durante el desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular adulta. Así mismo, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula sin choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión funcional de una molécula endógena que no funciona bien o una versión que no funciona bien de una molécula endógena que funciona normalmente.

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

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

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

Una molécula "de fusión" es una molécula en la que están unidas dos o más moléculas de una subunidad, preferiblemente de forma covalente. Las moléculas de una subunidad pueden ser el mismo tipo químico de molécula, o pueden ser tipos químicos diferentes de moléculas. Ejemplos del primer tipo de molécula de fusión incluyen, pero no se limitan a, proteínas de fusión (por ejemplo, una ZFN que comprende una fusión entre un dominio que se une a 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 *supra*). Ejemplos del segundo tipo de molécula de fusión incluyen, pero no se limitan a, una fusión entre un ácido nucleico que forma un tríplex y un polipéptido, y una fusión entre un agente de unión al surco menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede ser el resultado del suministro de la proteína de fusión a la célula o mediante el suministro de un polinucleótido que codifica la proteína de fusión para una célula, en donde el polinucleótido se transcribe, y el transcrito se traduce, para generar la proteína de fusión. El corte y empalme con ajuste en trans, la escisión de polipéptidos y la ligación de polipéptidos también pueden estar implicados en la expresión de una proteína en una célula. Se presentan métodos para el suministro de polinucleótidos y polipéptidos a células en otra parte de esta descripción.

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

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

"Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitarse a, activación génica y represión génica.

Las células "vegetales" incluyen, pero sin limitarse a, células vegetales monocotiledóneas (monocots) o dicotiledóneas (dicots). Ejemplos no limitantes de monocots incluyen plantas de cereales tales como maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, cebolla, banana y coco. Ejemplos no limitantes de dicots incluyen tabaco, tomate, girasol, algodón, remolacha azucarera, patata, lechuga, melón, soja, canola (semilla de colza) y alfalfa. Las células vegetales pueden proceder de cualquier parte de la planta y/o de cualquier fase de desarrollo de la planta. Por lo tanto, las células vegetales pueden ser células procedentes de semillas de la planta. En algunas realizaciones, la planta o la célula vegetal es, o se obtiene a partir de, una planta que participa en la producción de

aceites vegetales para usos comestibles y/o industriales (es decir, una "planta productora de aceite"). Plantas productoras de aceite ejemplares incluyen, pero no se limitan a, *Brassica* sp. (por ejemplo, *Brassica napus* que incluye cultivares de canola), maíz, soja, crambe, mostaza, semilla de ricino, cacahuete, sésamo, algodón, semillas de lino, cártamo, aceite de palma, lino, girasol y coco.

5 Una "región de interés" es cualquier región de cromatina celular tal como, por ejemplo, un gen o una secuencia no codificante dentro de un gen o adyacente al mismo, en la que es deseable la unión a una molécula exógena. La unión puede ser con fines de escisión dirigida del ADN y/o recombinación dirigida. Una región de interés puede estar presente en un cromosoma, un episoma, un genoma de un orgánulo (por ejemplo, de mitocondria, cloroplasto), o un genoma vírico infeccioso, por ejemplo. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias líder, secuencias de remolque o intrones, o dentro de regiones no transcritas, aguas arriba o aguas abajo de la región codificante. Una región de interés puede ser tan pequeña como una única pareja de nucleótidos o de hasta 25.000 parejas de nucleótidos de longitud, o cualquier valor entero de parejas de nucleótidos.

15 Las expresiones "enlace funcional" y "ligado de forma funcional" (o "ligado funcionalmente") se usan de forma intercambiable con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en donde los componentes están dispuestos de tal modo que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar en una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está ligada funcionalmente a una secuencia codificante, si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificante, como respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción generalmente está ligada funcionalmente en *cis* con una secuencia codificante, pero no necesita estar directamente adyacente a la misma. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está ligada funcionalmente a una secuencia codificante, aunque no sean contiguas.

25 Con respecto a polipéptidos de fusión, la expresión "ligado funcionalmente" puede hacer referencia al hecho de que cada uno de los componentes realiza la misma función en el enlace con el otro componente como si no estuvieran ligados de este modo. Por ejemplo, con respecto a un polipéptido de fusión en el que un dominio de unión a ADN de ZFP está fusionado con un dominio regulador, el dominio de unión a ADN de ZFP y el dominio regulador están ligados funcionalmente si en el polipéptido de fusión, la porción del dominio de unión a ADN de ZFP es capaz de unirse a su sitio diana y/o su sitio de unión, mientras que el dominio regulador es capaz de regular el ADN en las cercanías del sitio diana.

35 Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero conserva la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede poseer más, menos o la misma cantidad de residuos que la molécula natural correspondiente, y/o puede contener una o más sustituciones de aminoácidos o nucleótidos. Los métodos para determinar la función de un ácido nucleico (por ejemplo, función codificante, capacidad de hibridarse con otro ácido nucleico) son bien conocidos en la técnica. Así mismo, los métodos para determinar la función proteica son bien conocidos. Por ejemplo, la función de unión a ADN de un polipéptido puede determinarse, por ejemplo, por unión a un filtro, desplazamiento de la movilidad electroforética o ensayos de inmunoprecipitación. La escisión del ADN se puede someter a ensayo mediante electroforesis en gel. Véase Ausubel et al., *supra*. La capacidad de una proteína para interactuar con otra proteína puede determinarse, por ejemplo, mediante coimmunoprecipitación, ensayos de doble híbrido o complementación, tanto genéticos como bioquímicos. Véanse, por ejemplo, Fields et al. (1989) Nature 340:245-246; documento de patente de EE.UU. nº 5.585.245 y documento PCT WO 98/44350.

45 Tal y como se usa en este documento, "ácidos grasos saturados" incluye, pero no se limita a, ácido láurico (C12:0), mirístico (C14:0), palmítico (C16:0) y esteárico (C18:0). De forma similar, "ácidos grasos monoinsaturados" y "ácidos grasos poliinsaturados" incluyen, pero no se limitan a, ácidos grasos C18, tales como ácido oleico (C18:1), ácido linoleico reducido (C18:2) y ácido linoléico reducido (C18:3).

#### Sitios diana

50 Los métodos y las composiciones descritas incluyen proteínas de fusión que comprenden un dominio que se une a ADN (por ejemplo, ZFP) y un dominio regulador o dominio de escisión (nucleasa) (o un medio dominio de escisión), en donde el dominio que se une a ADN (por ejemplo, un dominio de dedos de zinc), mediante la unión a una secuencia en la cromatina celular en un gen implicado en la síntesis de ácidos grasos, dirige la actividad del dominio regulador de la transcripción o del dominio de escisión (o medio dominio de escisión) en el entorno de la secuencia y, por lo tanto, modula la transcripción o induce la escisión en el entorno de la secuencia diana.

55 Como se expone en esta descripción en otra parte, un dominio de dedos de zinc se puede diseñar para unirse a prácticamente cualquier secuencia deseada. En consecuencia, después de identificar una región de interés que contiene una secuencia en la que se desea una regulación, escisión o recombinación de genes, uno o varios dominios que se unen a dedos de zinc se pueden diseñar para unirse a una o varias secuencias en la región de interés.

La selección de un sitio diana en una región genómica de interés en la cromatina celular de cualquier gen implicado en la biosíntesis de ácidos grasos (por ejemplo, véase la Figura 1) para la unión mediante un dominio de dedos de zinc (por ejemplo, un sitio diana) se puede lograr, por ejemplo, de acuerdo con los métodos descritos en el documento de patente en copropiedad de EE.UU. nº 6.453.242 (17 de septiembre de 2002), que también describe métodos para el diseño de ZFPs para unirse a una secuencia seleccionada. Será evidente para los expertos en la técnica que la simple inspección visual de una secuencia de nucleótidos también se puede utilizar para la selección de un sitio diana. En consecuencia, cualquier medio para la selección de un sitio diana se puede usar en los métodos reivindicados.

Los sitios diana se componen generalmente de una pluralidad de subsitios diana adyacentes. Un subsitio diana se refiere a la secuencia (por lo general ya sea un triplete de nucleótidos, o un cuadruplete de nucleótidos que se puede solapar en un nucleótido con un cuadruplete adyacente) unida por un dedo de zinc individual. Véase, por ejemplo, el documento WO 02/077227. Si la hebra con la que una proteína con dedos de zinc tiene la mayoría de los contactos, se designa la hebra diana, "hebra de reconocimiento primaria" o "hebra de contacto principal", algunas proteínas con dedos de zinc se unen a un triplete de tres bases en la hebra diana y una cuarta base en la hebra no diana. Un sitio diana tiene generalmente una longitud de al menos 9 nucleótidos y, por consiguiente, se une mediante un dominio de unión de dedos de zinc que comprende al menos tres dedos de zinc. Sin embargo la unión, por ejemplo, de un dominio de unión de 4 dedos a un sitio diana de 12 nucleótidos, un dominio de unión de 5 dedos a un sitio diana de 15 nucleótidos o un dominio de unión de 6 dedos a un sitio diana de 18 nucleótidos, también es posible. Como será evidente, la unión de dominios de unión de mayor tamaño (por ejemplo, 7, 8, 9 dedos y más) a sitios diana más largos, también es posible.

No es necesario que un sitio diana sea un múltiplo de tres nucleótidos. Por ejemplo, en casos en los que tienen lugar interacciones cruzadas entre hebras (véase, por ejemplo, el documento de Patente de EE.UU. 6.453.242 y el documento WO 02/077227), uno o varios de los dedos de zinc individuales de un dominio de unión de múltiples dedos, se puede unir a subsitios cuádruples solapantes. Como resultado, una proteína de tres dedos se puede unir a una secuencia de 10 nucleótidos, en donde el décimo nucleótido forma parte de un cuadruplete unido por un dedo terminal, una proteína de cuatro dedos se puede unir a una secuencia de 13 nucleótidos, en donde el nucleótido decimotercero forma parte de un cuadruplete unido por un dedo terminal, etc.

La longitud y la naturaleza de las secuencias enlazadoras de aminoácidos entre los dedos de zinc individuales en un dominio de unión con múltiples dedos, también afecta a la unión a una secuencia diana. Por ejemplo, la presencia de un llamado "enlazador no canónico", "enlazador largo" o "enlazador estructurado" entre los dedos de zinc adyacentes en un dominio de unión con múltiples dedos, puede permitir que esos dedos se unan a subsitios que no son inmediatamente adyacentes. Ejemplos no limitantes de tales enlazadores se describe, por ejemplo, en el documento de Patente de EE.UU. nº 6.479.626 y en el documento WO 01/53480. En consecuencia, uno o varios subsitios, en un sitio diana para un dominio de unión de dedos de zinc, pueden estar separados entre sí por 1, 2, 3, 4, 5 o más nucleótidos. Para proporcionar un ejemplo, un dominio de unión de cuatro dedos se puede unir a un sitio diana de 13-nucleótidos que comprende, en secuencia, dos subsitios contiguos de 3 nucleótidos, un nucleótido intercalado y dos subsitios de tripletes contiguos. Véase, también, el documento de solicitud de Patente de EE.UU. nº 61/130.099 para composiciones y métodos para enlazar nucleasas artificiales que se unen a sitios diana separados por diferentes cantidades de nucleótidos. La distancia entre secuencias (por ejemplo, sitios diana) se refiere al número de nucleótidos o de parejas de nucleótidos que se intercalan entre dos secuencias, tal y como se mide desde los bordes de las secuencias más cercanas entre sí.

En ciertas realizaciones, se diseñan ZFPs con la función de factor de transcripción. Para la función de factor de transcripción, una unión y una proximidad suficiente al promotor, es simplemente todo lo que se necesita generalmente. Un posicionamiento exacto con relación al promotor, la orientación y estar dentro de los límites, la distancia no importa en gran medida. Esta característica permite una gran flexibilidad en la elección de los sitios diana para la construcción de factores de transcripción artificiales. El sitio diana reconocido por la ZFP puede ser, por lo tanto, cualquier sitio adecuado en el gen diana que permita la activación o la represión de la expresión génica a través de una ZFP, opcionalmente ligada a un dominio regulador. Los sitios diana preferidos incluyen regiones adyacentes, aguas abajo o aguas arriba del sitio de inicio de la transcripción. Además, los sitios diana que se encuentran en regiones potenciadoras, sitios represores, sitios de pausa de una ARN polimerasa y sitios reguladores específicos (por ejemplo, sitios SP-1, elementos de respuesta a la hipoxia, elementos de reconocimiento de receptores nucleares, sitios de unión de p53), sitios en la región que codifica el ADNc o en una región que codifica un marcador de secuencia expresado (EST).

En otras realizaciones, se diseñan ZFPs con actividad nucleasa. La expresión de una ZFN que comprende una proteína de fusión que comprende un dominio de unión de dedos de zinc y un dominio de escisión (o de dos proteínas de fusión, comprendiendo cada una un dominio de unión de dedos de zinc y un medio dominio de escisión), en una célula, efectúa una escisión en el entorno de la secuencia diana. En ciertas realizaciones, la escisión depende de la unión de moléculas de fusión de un dominio de dos dedos de zinc/medio dominio de escisión a sitios diana distintos. Los dos sitios diana pueden estar en hebras de ADN opuestas, o alternativamente, los dos sitios diana pueden estar en la misma hebra de ADN.

## Dominios de unión a ADN

Puede usarse cualquier dominio de unión a ADN en la práctica de la presente memoria descriptiva. En ciertos aspectos, el dominio de unión a ADN comprende un dominio de unión de dedos de zinc con uno o varios dedos de zinc (Miller et al. (1985) EMBO J. 4:1609-1614; Rhodes (1993) Scientific American Feb.: 56-65; documento de Patente de EE.UU. nº 6.453.242). Típicamente, un único dominio de dedos de zinc tiene aproximadamente 30 aminoácidos de longitud. Estudios estructurales han demostrado que cada dominio de dedos de zinc (motivo) contiene dos láminas beta (sostenidas en un giro beta que contiene los dos residuos de cisteína invariables) y una hélice alfa (que contiene los dos residuos de histidina invariables), que se sostienen en una conformación particular mediante la coordinación de un átomo de zinc por las dos cisteínas y las dos histidinas.

Los dedos de zinc incluyen tanto dedos de zinc  $C_2H_2$  canónicos (es decir, aquellos en que el ion zinc está coordinado por dos residuos de cisteína y dos de histidina) y dedos de zinc no canónicos tales como, por ejemplo, dedos de zinc  $C_3H$  y  $C_2HC$  (aquellos en los que el ion zinc está coordinado por tres residuos de cisteína y un residuo de histidina, véase, por ejemplo, el documento de publicación de Patente de EE.UU. nº 2008/0182332) y dedos de zinc  $C_4$  (aquellos en los que el ion de zinc está coordinado por cuatro residuos de cisteína). Véase también el documento WO 02/057293.

Los dominios de unión de dedos de zinc se pueden diseñar 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 Biotechnology. 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 dedos de zinc modificado genéticamente o que se produce de forma no natural) puede tener una especificidad de unión novedosa, en comparación con una proteína con dedos de zinc de origen natural. Métodos para modificar genéticamente incluyen, pero no se limitan a, un diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, usar bases de datos que comprenden secuencias de nucleótidos de tripletes (o cuadrupletes) y secuencias de aminoácidos de dedos de zinc individuales, en donde cada secuencia de nucleótidos de tripletes o cuadrupletes se asocia con una o varias secuencias de aminoácidos de dedos de zinc que se unen a la secuencia del triplete o cuadruplete particular. Véanse, por ejemplo, los documentos de patentes en copropiedad de EE.UU. 6.453.242 y 6.534.261. Métodos de diseño adicionales se describen, por ejemplo, en los documentos de Patente de EE.UU. 6.746.838; 6.785.613; 6.866.997; y 7.030.215.

Métodos de selección ejemplares, incluyendo presentación en fagos y sistemas de dos híbridos, se describen en los documentos de Patente de EE.UU. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como los documentos WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.

Una mejora de la especificidad de la unión para los dominios de unión de dedos de zinc ha sido descrita, por ejemplo, en el documento de Patente en co-propiedad de EE.UU. nº 6.794.136.

Puesto que un dedo de zinc individual se une a una secuencia de tres nucleótidos (es decir, triplete) (o una secuencia de cuatro nucleótidos que se puede solapar, en un nucleótido, con el sitio de unión de cuatro nucleótidos de un dedo de zinc adyacente), la longitud de una secuencia para la cual se diseña un dominio de unión de dedos de zinc para que se una a la misma (por ejemplo, una secuencia diana) determinará el número de dedos de zinc en un dominio de unión de dedos de zinc modificado genéticamente. Por ejemplo, para ZFPs en las que los motivos de dedos no se unen a subsitios solapantes, una secuencia diana de seis nucleótidos se une mediante un dominio de unión de dos dedos; una secuencia diana de nueve nucleótidos se une mediante un dominio de unión de tres dedos, etc. Como se indica en la presente memoria, los sitios de unión para dedos de zinc individuales (es decir, subsitios) en un sitio diana no tienen que ser contiguos, sino que pueden estar separados por uno o varios nucleótidos, dependiendo de la longitud y naturaleza de las secuencias de aminoácidos entre los dedos de zinc (es decir, los enlazadores entre los dedos) en un dominio de unión de múltiples dedos.

En un dominio de unión de dedos de zinc de múltiples dedos, los dedos de zinc adyacentes pueden estar separados por secuencias enlazadoras de aminoácidos de aproximadamente 5 aminoácidos (los llamados enlazadores entre dedos "canónicos") o, como alternativa, por uno o más enlazadores no canónicos. Véanse, por ejemplo, los documentos de las patentes en copropiedad de Estados Unidos nº 6.453.242 y 6.534.261. Para dominios de unión de dedos de zinc diseñados por ingeniería que comprenden más de tres dedos, puede preferirse la inserción de enlazadores entre dedos más largos ("no canónicos") entre ciertos dedos de zinc, ya que esto puede aumentar la afinidad y/o la especificidad de la unión por el dominio de unión. Véase, por ejemplo, el documento de Patente de EE.UU. nº 6.479.626 y el documento WO 01/53480. Por consiguiente, los dominios de unión de dedos de zinc de múltiples dedos también pueden caracterizarse con respecto a la presencia y localización de enlazadores entre dedos no canónicos. Por ejemplo, un dominio de unión de dedos de zinc de seis dedos que comprende tres dedos (unidos por dos enlazadores entre dedos canónicos), un enlazador largo y tres dedos adicionales (unidos por dos enlazadores entre dedos canónicos) se indica como una configuración 2x3. Asimismo, un dominio de unión que comprende dos dedos (con un enlazador canónico entre los mismos), un enlazador largo y dos dedos adicionales (unidos por un enlazador canónico) se indica como una proteína 2x2. Una proteína que comprende tres unidades de dos dedos (en cada una de las cuales los dos dedos están unidos por un enlazador canónico), y en donde cada unidad de dos dedos está unida a la unidad de dos dedos adyacentes por un enlazador largo, se indica como una proteína 3x2.

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

Como se ha mencionado previamente, un subsitio diana es una secuencia de tres o cuatro nucleótidos que se une por un único dedo de zinc. Para ciertos propósitos, una unidad de dos dedos se indica como un módulo de unión. Un módulo de unión puede obtenerse, por ejemplo, seleccionando dos dedos adyacentes en el contexto de una proteína de múltiples dedos (generalmente tres dedos) que se une a una secuencia diana particular de seis nucleótidos. Como alternativa, pueden construirse módulos mediante el ensamblaje de dedos de zinc individuales. Véanse también los documentos WO 98/53057 y WO 01/53480.

En un aspecto, se describe en el presente documento un dominio de unión de dedos de zinc que comprende una secuencia de aminoácidos como se muestra en la Tabla 1. En otro aspecto, la descripción proporciona un polinucleótido que codifica un dominio de unión de dedos de zinc, en el que el dominio de unión de dedos de zinc comprende una secuencia de aminoácidos como se muestra en la Tabla 1.

Alternativamente, el dominio que se une a ADN se puede obtener a partir de una nucleasa. Por ejemplo, las secuencias de reconocimiento de endonucleasas de migración dirigida y meganucleasas tales como I-Scel, I-CeuI, PI-PspI, PI-Sce, I-ScelV, I-CsmI, I-PanI, I-Scell, I-PpoI, I-ScellI, I-Crel, I-TevI, I-TevII y I-TevIII son conocidas. Véanse también los documentos de Patente de EE.UU. n° 5.420.032; Patente de EE.UU. n° 6.833.252; Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388; Dujon et al. (1989), Gene 82:115-118; Perler et al. (1994) Nucleic Acids Res. 22, 1125-1127; Jasin (1996) Trends Genet. 12:224-228; Gimble et al. (1996) J. Mol. Biol. 263:163-180; Argast et al. (1998) J. Mol. Biol. 280:345-353 y el catálogo de New England Biolabs. Además, la especificidad de la unión a ADN de endonucleasas de migración dirigida y meganucleasas se puede diseñar para unirse a sitios diana no naturales. Véase, por ejemplo, Chevalier et al. (2002) Molec. Cell 10:895-905; Epinat et al. (2003) Nucleic Acids Res. 31:2952-2962; Ashworth et al. (2006) Nature 441:656-659; Paques et al. (2007), Current Gene Therapy 7:49-66; documento de publicación de Patente de EE.UU. n° 20070117128.

#### 30 Dominios reguladores

Los dominios que se unen a ADN (por ejemplo, ZFPs) descritos en este documento, opcionalmente pueden estar asociados con dominios reguladores para una modulación de la expresión génica. La ZFP puede estar asociada de forma covalente o no covalente con uno o varios dominios reguladores, alternativamente dos o más dominios reguladores, siendo los dos o más dominios dos copias del mismo dominio, o dos dominios diferentes. Los dominios reguladores pueden enlazarse covalentemente a la ZFP, por ejemplo, a través de un enlazador de aminoácidos, como parte de una proteína de fusión. La ZFP también se puede asociar con un dominio regulador por medio de un dominio de dimerización no covalente, por ejemplo, una cremallera de leucina, un dominio N-terminal de proteína STAT o una proteína de unión a FK506 (véase, por ejemplo, O'Shea, Science 254: 539 (1991), Barahmand-Pour et al., Curr. Top. Microbiol. Immunol. 211:121-128 (1996); Klemm et al., Annu. Rev. Immunol. 16:569-592 (1998); Klemm et al., Annu. Rev. Immunol. 16:569-592 (1998); Ho et al., Nature 382:822-826 (1996); y Pomeranz et al., Biochem. 37:965 (1998)). El dominio regulador puede asociarse con la ZFP en cualquier posición adecuada, incluyendo el extremo C-terminal o N-terminal de la ZFP.

Dominios reguladores comunes para adición a la ZFP incluyen, por ejemplo, dominios efectores de factores de transcripción (activadores, represores, coactivadores, correpresores), silenciadores, receptores de hormonas nucleares, factores de transcripción de oncogenes (por ejemplo, miembros de la familia myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos, etc.); enzimas de reparación del ADN y sus factores y modificadores asociados; enzimas de reordenamiento de ADN y sus factores y modificadores asociados; proteínas asociadas a la cromatina y sus modificadores (por ejemplo, cinasas, acetilasas y desacetilasas); y enzimas modificadoras de ADN (por ejemplo, metiltransferasas, topoisomerasas, helicasas, ligasas, cinasas, fosfatasas, polimerasas, endonucleasas) y sus factores y modificadores asociados.

Los polipéptidos del factor de transcripción a partir del cual se puede obtener un dominio regulador incluyen aquéllos que están involucrados en la transcripción regulada y basal. Tales polipéptidos incluyen factores de transcripción, sus dominios efectores, coactivadores, silenciadores, receptores nucleares de hormonas (véase, por ejemplo, Goodrich et al., Cell. 84:825-30 (1996) para una revisión de proteínas y elementos de ácido nucleico implicados en la transcripción; los factores de transcripción en general se revisan en Barnes y Adcock, Clin. Exp. Allergy 25 Supl 2:46-9 (1995) y Roeder, Methods Enzymol. 273:165-71 (1996)). Factores de transcripción de receptores de hormonas nucleares se describen, por ejemplo, en Rosen et al., J Med. Chem. 38:4855-74 (1995). La familia C/EBP de factores de transcripción se revisan en Wedel et al., Immunobiology 193:171-85 (1995). Coactivadores y correpresores que median en la regulación de la transcripción mediante receptores nucleares de hormonas, se revisan, por ejemplo, en Meier, Eur. J. Endocrinol. 134(2):158-9 (1996); Kaiser et al., Trends Biochem. Sci. 21:342-5 (1996); y

Uttley et al., *Nature* 394:498-502 (1998)). Los factores de transcripción GATA, que están implicados en la regulación de la hematopoyesis, se describen, por ejemplo, en Simon, *Nat. Genet.* 11:9-11 (1995); Weiss et al., *Exp. Hematol.* 23:99-107. La proteína de unión a una caja TATA (TBP) y sus polipéptidos TAP asociados (que incluyen TAF30, TAF55, TAF80, TAF10, TAFI 50 y TAF250) se describen en Goodrich y Tjian, *Curr. Opin. Cell Biol.* 6:403-9 (1994) y Hurlley, *Curr. Opin. Struct. Biol.* 6:69-75 (1996). La familia STAT de factores de transcripción se revisa, por ejemplo, en Barahmand-Pour et al., *Curr. Top. Microbiol. Immunol.* 211:121-8 (1996). Los factores de transcripción implicados en enfermedades se revisan en Aso et al., *J Clin. Invest.* 97:1561-9 (1996).

En un aspecto, el dominio de represión KRAB de la proteína KOX-1 humana se utiliza como un represor de la transcripción (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)). En otro aspecto, KAP-1, un correpresor de KRAB se emplea con KRAB (Friedman et al., *Genes Dev.* 10:2067-2078 (1996)). Alternativamente, KAP-1 puede utilizarse solo con una ZFP. Otros factores de transcripción preferidos y dominios de factores de transcripción que actúan como represores de la transcripción incluyen MAD (véase, p. ej., Sommer et al., *J. Biol. Chem.* 273:6632-6642 (1998); Gupta et al., *Oncogene* 16:1149-1159 (1998); Queva et al., *Oncogene* 16:967-977 (1998); Larsson et al., *Oncogene* 15:737-748 (1997); Laherty et al., *Cell* 89:349-356 (1997); y Cultraro et al., *Mol. Cell. Biol.* 17:2353-2359 (1997)); FKHR (forkhead in rhabdosarcoma gene; Ginsberg et al., *Cancer Res.* 15:3542-3546 (1998); Epstein et al., *Mol. Cell. Biol.* 18:4118-4130 (1998)); EGR-1 (early growth response gene product-1; Yan et al., *Proc. Natl. Acad. Sci USA* 95:8298-8303 (1998); y Liu et al., *Cancer Gene Ther.* 5:3-28 (1998)); el dominio del represor del factor represor de ets2 (ERD; Sgouras et al., *EMBO J* 14:4781-4793 (1995)); el dominio de interacción MAD smSIN3 (SID; Ayer et al., *Mol. Cell. Biol.* 16:5772-5781 (1996)); y el dominio de represión anfíllico ERF3 (factor-3 de respuesta a etileno), EAR (Ohta, M., et al. (2001), *Plant Cell* 13:1959-1968).

En un aspecto, el dominio de activación de VP16 de HSV se utiliza como un activador de la transcripción (véase, por ejemplo, Hagmann et al., *J. Virol.* 71:5952-5962 (1997)). Otros factores de transcripción preferidos que podrían suministrar dominios de activación incluyen el dominio de activación VP64 (Seipel et al., *EMBO J.* 11:4961-4968 (1996)); receptores de hormonas nucleares (véase, por ejemplo, Torchia et al., *Curr. Opin. Cell Biol.* 10:373-383 (1998)); la subunidad p65 del factor nuclear kappa B (Bitko y Barik, *J. Virol* 72:5.610-5618 (1998) y Doyle y Hunt, *Neuroreport* 8:2937-2942 (1997)); EGR-1 (producto-1 del gen de respuesta de crecimiento temprano; Yan et al., *Proc. Natl. Acad. Sci. USA* 95:8298-8303 (1998); y Liu et al., *Cancer Gene Ther.* 5:3-28 (1998)); y la proteína reguladora de la vía biosintética de la antocianina de maíz, C1 (S.A. Goff, et al., (1991), *Genetics and Development* 5: 298-309).

Las cinasas, fosfatasa y otras proteínas que modifican polipéptidos implicados en la regulación génica son también útiles como dominios reguladores para las ZFPs. Tales modificadores están implicados a menudo en la activación o desactivación de la transcripción mediada, por ejemplo, por hormonas, cinasas implicadas en la regulación transcripcional se revisan en Davis, *Mol. Reprod. Dev.* 42:459-67 (1995), Jackson et al., *Adv. Second Messenger Phosphoprotein Res.* 28:279-86 (1993), y Bouliskas, *Crit. Rev. Eukaryot. Gene Expr.* 5:1-77 (1995), mientras que las fosfatasa se revisan, por ejemplo, en Schonthal y Semin, *Cancer Biol.* 6:239-48 (1995). Las tirosina cinasas nucleares se describen en Wang, *Trends Biochem. Sci.* 19:373-6 (1994)

Como se ha descrito, los dominios útiles pueden también obtenerse a partir de los productos génicos de oncogenes (por ejemplo, miembros de la familia myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos) y sus factores y modificadores asociados. Los oncogenes se describen, por ejemplo, en Cooper, *Oncogenes*, 2ª ed., The Jones and Bartlett Series in Biology, Boston, Mass., Jones and Bartlett Publishers, 1995. Los factores de transcripción ets se revisan en Wasylk et al., *Eur. J. Biochem.* 211:7-18 (1993) y Crepieux et al., *Crit. Rev. Oncog.* 5:615-38 (1994). Los oncogenes myc se revisan, por ejemplo, en Ryan et al., *Biochem. J.* 314:713-21 (1996). Los factores de transcripción jun y fos se describen, por ejemplo, en The Fos and Jun Families of Transcription Factors, Angel y Herrlich, compiladores (1994). El oncogén max se revisa en Hurlin et al., *Cold Spring Harb. Symp. Quant. Biol.* 59:109-16. La familia de genes myb se revisa en Kanei-Ishii et al., *Curr. Top. Microbiol. Immunol.* 211:89-98 (1996). La familia mos se revisa en Yew et al., *Curr. Opin. Genet. Dev.* 3:19-25 (1993).

Las ZFPs pueden incluir dominios reguladores obtenidos a partir de enzimas de reparación del ADN y sus factores y modificadores asociados. Sistemas de reparación del ADN se revisan, por ejemplo, en Vos, *Curr. Opin. Cell Biol.* 4:385-95 (1992); Sancar, *Ann. Rev. Genet.* 29:69-105 (1995); Lehmann, *Genet. Eng.* 17:1-19 (1995); y Wood, *Ann. Rev. Biochem.* 65:135-67 (1996). Las enzimas de reordenamiento de ADN y sus factores y modificadores asociados también se pueden utilizar como dominios reguladores (véase, por ejemplo, Gangloff et al., *Experientia* 50:261-9 (1994); Sadowski, *FASEB J.* 7:760-7 (1993)).

Del mismo modo, dominios reguladores se pueden obtener a partir de enzimas modificadoras de ADN (por ejemplo, metiltransferasas, topoisomerasas, helicasas, ligasas, cinasas, fosfatasa, polimerasas de ADN) y sus factores y modificadores asociados. Las helicasas se revisan en Matson et al., *Bioessays*, 16: 3-22 (1994) y las metiltransferasas se describen en Cheng, *Curr. Opin. Struct. Biol.* 5:4-10 (1995). Las proteínas asociadas a la cromatina y sus modificadores (por ejemplo, cinasas, acetilasas y desacetilasas) tales como la histona desacetilasa (Wolfe, *Science* 272:371-2 (1996)) también son útiles como dominios para añadir a la ZFP de elección. En un aspecto preferido, el dominio regulador es una metil transferasa de ADN que actúa como un represor de la transcripción (véase, por

ejemplo, Van den Wyngaert et al., FEBS Lett. 426:283-289 (1998); Flynn et al., J. Mol. Biol. 279:101-116 (1998); Okano et al., Nucleic Acids Res. 26:2536-2540 (1998); y Zardo y Caiafa, J. Biol. Chem. 273:16517-16520 (1998)). En otro aspecto preferido, endonucleasas tales como *Fok1* se utilizan como represores de la transcripción, que actúan a través de la escisión de genes (véanse, por ejemplo, los documentos WO95/09233; y PCT/US94/01201).

5 Los factores que controlan la cromatina y la estructura del ADN, el movimiento y la localización y sus factores y modificadores asociados; factores obtenidos a partir de microbios (por ejemplo, procariotas, eucariotas y virus) y factores que se asocian con ellos o los modifican, también se pueden utilizar para obtener proteínas quiméricas. En un aspecto, las recombinasas y las integrasas se utilizan como dominios reguladores. En un aspecto, la acetiltransferasa de histonas se utiliza como un activador de la transcripción (véase, por ejemplo, Jin y Scotto, Mol. Cell. Biol. 18:4377-4384 (1998); Wolffe, Science 272:371-372 (1996); Taunton et al., Science 272:408-411 (1996); y Hassig et al., Proc. Natl. Acad. Sci USA 95:3519-3524 (1998)). En otro aspecto, la desacetilasa de histonas se utiliza como un represor de la transcripción (véase, por ejemplo, Jin y Scotto, Mol. Cell. Biol. 18:4377-4384 (1998); Syntichaki y Thireos, J. Biol. Chem. 273:24414-24419 (1998); Sakaguchi et al., Genes Dev. 12:2831-2841 (1998); y Martinez et al, J. Biol. Chem. 273:23781-23785 (1998)).

15 Los dominios enlazadores entre dominios de polipéptidos, por ejemplo, entre dos ZFPs o entre una ZFP y un dominio regulador, se pueden incluir. Tales enlazadores son típicamente secuencias polipeptídicas, tales como secuencias poli gli de entre aproximadamente 5 y 200 aminoácidos. Los enlazadores pueden ser subsecuencias de aminoácidos flexibles o rígidas que se sintetizan como parte de una proteína de fusión recombinante. Véase, por ejemplo, el documento de Patente de EE.UU. nº 6.534.261; Liu et al., Proc. Nat. Acad. Sci. USA, 95:5525-5530 (1997); Pomerantz et al., Proc. Nat. Acad. Sci. USA 92:9752-9756 (1995); Kim et al., Proc. Nat. Acad. Sci. USA 93:1156-1160 (1996).

20 Alternativamente, los enlazadores flexibles pueden ser diseñados racionalmente usando un programa de ordenador capaz de modelar ambos sitios de unión al ADN y los propios péptidos (Desjarlais y Berg, Proc. Nat. Acad. Sci. USA 90:2256-2260 (1993), Desjarlais y Berg, Proc. Nat. Acad. Sci. USA 91:11099-11103 (1994) o por métodos de presentación en fagos.

25 Un enlazador químico se utiliza para conectar secuencias de dominios producidas de forma sintética o recombinante. Tales enlazadores flexibles son conocidos por los expertos en la técnica. Por ejemplo, enlazadores de polietilenglicol están disponibles de Shearwater Polymers, Inc. Huntsville, Ala. Estos enlazadores tienen opcionalmente enlaces amida, enlaces sulfhidrilo o enlaces heterofuncionales. Además del enlace covalente de las ZFPs a los dominios reguladores, se pueden utilizar métodos no covalentes para producir moléculas con ZFPs asociadas con dominios reguladores.

#### Dominios de escisión

35 Como se ha señalado anteriormente, el dominio que se une a ADN también puede estar asociado con un dominio de escisión (nucleasa). Por ejemplo, las endonucleasas de migración dirigida se pueden modificar en su especificidad de la unión a ADN, al tiempo que se conserva la función de nucleasa. Además, las proteínas con dedos de zinc también pueden estar fusionadas a un dominio de escisión para formar una nucleasa con dedos de zinc (ZFN). La porción del dominio de escisión de las proteínas de fusión descritas en este documento se puede obtener a partir de cualquier endonucleasa o exonucleasa. Endonucleasas ejemplares a partir de las cuales se puede obtener un dominio de escisión incluyen, pero no se limitan a, endonucleasas de restricción y endonucleasas de migración dirigida. Véase, por ejemplo, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; y Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388. Se conocen enzimas adicionales que escinden el ADN (por ejemplo, la nucleasa S1, la nucleasa de frijol mungo; la ADNasa pancreática I; la nucleasa microcócica, la endonucleasa de levadura HO; véase también Linn et al. (compiladores) Nucleases, Cold Spring Harbor Laboratory Press, 1993). Una o varias de estas enzimas de síntesis (o fragmentos funcionales de las mismas) se pueden utilizar como fuente de dominios de escisión y medio dominios de escisión.

45 Del mismo modo, un medio dominios de escisión se puede obtener a partir de cualquier nucleasa o una porción de la misma, como se ha descrito anteriormente, lo que requiere una dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden medio dominios de escisión. Alternativamente, se puede utilizar una sola proteína que comprende dos medio dominios de escisión. Los dos medio dominios de escisión se pueden obtener a partir de la misma endonucleasa (o fragmentos funcionales de la misma), o cada uno de los medio dominios de escisión se puede obtener a partir de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios diana para las dos proteínas de fusión se disponen preferiblemente, con respecto uno a otro, de modo que la unión de las dos proteínas de fusión a sus respectivos sitios diana coloca los medio dominios de escisión en una orientación espacial entre sí que permite que los medio dominios de escisión formen un dominio de escisión funcional, por ejemplo, mediante dimerización. Por lo tanto, en ciertos aspectos, los márgenes cerca de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número entero de nucleótidos o parejas de nucleótidos se puede intercalar entre dos sitios diana (por ejemplo, de 2 a 50 parejas de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios diana.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de una unión específica de secuencia con el ADN (en un sitio de reconocimiento) y la escisión de ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (p. ej., Tipo IIS) escinden el ADN en sitios separados del sitio de reconocimiento y tienen dominios de unión y de escisión separables. Por ejemplo, la enzima Tipo IIS *FokI* cataliza la escisión de la cadena doble de ADN, a 9 nucleótidos de su sitio de reconocimiento en una hebra y a 13 nucleótidos de su sitio de reconocimiento en la otra. Véanse, por ejemplo, los documentos de Patentes de EE.UU. 5.356.802; 5.436.150 y 5.487.994; así como Li et al. (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim et al. (1994b) J. Biol. Chem. 269:31.978-31.982. Por tanto, en un aspecto, las proteínas de fusión comprenden el dominio de escisión (o el medio dominio de escisión) de al menos una enzima de restricción de Tipo IIS y uno o varios dominios de unión de dedos de zinc, que pueden estar modificados genéticamente o no.

Una enzima de restricción ejemplar de Tipo IIS, cuyo dominio de escisión es separable del dominio de unión, es *FokI*. Esta enzima particular está activa como un dímero. Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA 95: 10.570-10.575. Por consiguiente, para los fines de la presente memoria descriptiva, la porción de la enzima *FokI* usada en las proteínas de fusión descritas se considera un medio dominio de escisión. Por lo tanto, para una escisión de cadena doble dirigida y/o una sustitución dirigida de secuencias celulares usando fusiones de dedos de zinc-*FokI*, dos proteínas de fusión, que comprende cada una un medio dominio de escisión de *FokI*, se pueden utilizar para reconstituir un dominio de escisión catalíticamente activo. Alternativamente, una única molécula de polipéptido que contiene un dominio de unión de dedos de zinc y dos medio dominios de escisión de *FokI*, también se puede utilizar. Parámetros para una escisión dirigida y una alteración de la secuencia dirigida utilizando fusiones de dedos de zinc-*FokI* se proporcionan en otro lugar en esta memoria descriptiva.

Un dominio de escisión o un medio dominio de escisión puede ser cualquier porción de una proteína que conserva la actividad de escisión, o que conserva la capacidad de multimerizar (por ejemplo, dimerizan) para formar un dominio de escisión funcional.

Enzimas de restricción de Tipo IIS ejemplares se describen en el documento de publicación internacional WO 07/014275. Enzimas de restricción adicionales también contienen dominios de unión y escisión separables y estos se contemplan en la presente memoria descriptiva. Véase, por ejemplo, Roberts et al. (2003) Nucleic Acids Res. 31:418-420.

En ciertos aspectos, el dominio de escisión comprende uno o varios medio dominios de escisión modificados genéticamente (también referidos como mutantes de dominio de dimerización) que minimizan o evitan una homodimerización, como se describe, por ejemplo, en los documentos de publicación de Patente de EE.UU. nº 20050064474; 20060188987; 20080131962. Los residuos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de *FokI* son todos dianas para influir en la dimerización de los medio dominios de escisión de *FokI*.

Medio dominios de escisión ejemplares modificados genéticamente de *FokI* que forman heterodímeros obligados, incluyen una pareja en donde un primer medio dominio de escisión incluye mutaciones en residuos de aminoácidos en las posiciones 490 y 538 de *FokI* y un segundo medio dominio de escisión incluye mutaciones en los residuos de aminoácidos 486 y 499.

Por lo tanto, en ciertos aspectos, la mutación en 490 sustituye Glu (E) por Lys (K); la mutación en 538 reemplaza Iso (I) por Lys (K); la mutación en 486 reemplaza Gln (Q) por Glu (E); y la mutación en la posición 499 reemplaza Iso (I) por Lys (K). Específicamente, los medio dominios de escisión modificados genéticamente descritos en este documento se prepararon mutando las posiciones 490 (E → K) y 538 (I → K) en un medio dominio de escisión, para producir un medio dominio de escisión modificado genéticamente designado "E490K:I538K" y mutando las posiciones 486 (Q → E) y 499 (I → L) en otro medio dominio de escisión para producir un medio dominio de escisión modificado genéticamente designado "Q486E:I499L". Los medio dominios de escisión modificados genéticamente descritos en este documento son mutantes heterodímeros obligados en los que se minimiza o suprime una escisión aberrante. Véase, por ejemplo, el Ejemplo 1 del documento de publicación de Patente de EE.UU. nº 2008/0131962. Véase, también, Szczepek et al. (2007) Nat Biotechnol 25:786-793. En ciertos aspectos, el medio dominio de escisión modificado genéticamente comprende mutaciones en las posiciones 486, 499 y 496 (numeradas en relación con *FokI* de tipo silvestre), para mutaciones ilustrativas que sustituyen el residuo Gln de tipo silvestre (Q) en la posición 486 con un residuo Glu (E), el residuo Iso (I) de tipo silvestre en la posición 499 con un residuo Leu (L) y el residuo de tipo silvestre Asn (N) en la posición 496 con Asp (D) o Glu (E) (también denominados dominios "ELD" y "ELE", respectivamente). En otros aspectos, el medio dominio de escisión modificado genéticamente comprende mutaciones en las posiciones 490, 538 y 537 (numeradas en relación con *FokI* de tipo silvestre), para mutaciones ilustrativas que sustituyen al residuo de tipo silvestre Glu (E) en la posición 490 con un residuo Lys (K), el residuo de tipo silvestre Iso (I) en la posición 538 con un residuo Lys (K) y el residuo de tipo silvestre His (H) en la posición 537 con un residuo Lys (K) o un residuo Arg (R) (también referidos como dominios "KKK" y "KKR", respectivamente). En otros aspectos, el medio dominio de escisión modificado genéticamente comprende mutaciones en las posiciones 490 y 537 (numeradas en relación con *FokI* de tipo silvestre), para mutaciones ilustrativas que sustituyen al residuo de tipo silvestre Glu (E) en la posición 490 con un residuo Lys (K) y el residuo de tipo silvestre His (H) en la posición 537 con un residuo Lys (K) o un residuo Arg (R) (también referidos como dominios "KIK" y "KIR", respectivamente).

(Véase la solicitud provisional de Estados Unidos 61/337.769, presentada el 8 de febrero de 2010). En otros aspectos, el medio dominio de escisión modificado genéticamente comprende "Sharkey" y/o mutaciones "Sharkey" (véase Guo et al., (2010) J. Mol. Biol. doi:10.1016/j.jmb.2010.04.060).

5 Los medio dominios de escisión modificados genéticamente descritos en esta memoria, se pueden preparar usando cualquier método adecuado, por ejemplo, mediante mutagénesis dirigida al sitio de medio dominios de escisión de tipo silvestre (*Fok I*) como se describe en los documentos de Publicación de Patente de EE.UU. nº. 20050064474 y 20080131962.

10 Otra enzima de restricción preferida de Tipo IIS es Bfil (véase, Zaremba et al, (2004) J. Mol Biol. 336(1):81-92). El dominio de escisión de esta enzima se puede separar de su dominio que se une a ADN y ligarlo funcionalmente a un dominio que se une a ADN de dedos de zinc para crear una ZFN.

#### Proteínas de fusión

15 Los métodos para el diseño y la construcción de proteínas de fusión (y polinucleótidos que las codifican) son conocidos por los expertos en la técnica. Por ejemplo, los métodos para el diseño y la construcción de proteínas de fusión que comprenden dominios de unión a ADN (p. ej., dominios de dedos de zinc) y dominios reguladores o de escisión (o medio dominios de escisión) y polinucleótidos que codifican tales proteínas de fusión, se describen en los documentos de Patente en co-propiedad de EE.UU. 6.453.242 y 6.534.261 y las publicaciones de solicitud de patente de EE.UU. 2007/0134796 y 2005/0064474.

20 En ciertas realizaciones, se construyen polinucleótidos que codifican las proteínas de fusión. Estos polinucleótidos se pueden insertar en un vector y el vector se puede introducir en una célula (véase más abajo para una descripción adicional en cuanto a los vectores y los métodos para la introducción de polinucleótidos en células).

Como se ha señalado anteriormente, en ciertas realizaciones, la proteína de fusión comprende una proteína con dedos de zinc que se une a un sitio diana en un gen implicado en la biosíntesis de ácidos grasos y al menos un dominio regulador transcripcional, por ejemplo, un dominio de activación o represión.

25 En otros aspectos de los métodos descritos en el presente documento, una nucleasa de dedos de zinc comprende una proteína de fusión que comprende un dominio de unión de dedos de zinc y un medio dominio de escisión de la enzima de restricción *FokI* y dos de proteínas de fusión de este tipo se expresan en una célula. La expresión de dos proteínas de fusión en una célula puede dar como resultado la entrega en la célula de las dos proteínas; la entrega en la célula de una proteína y un ácido nucleico que codifica una de las proteínas; la entrega en la célula de dos ácidos nucleicos que codifican cada uno una de las proteínas; o mediante la entrega en la célula de un único ácido nucleico, que codifica ambas proteínas. En aspectos adicionales, una proteína de fusión comprende una cadena polipeptídica única que comprende dos medio dominios de escisión y un dominio de unión de dedos de zinc. En este caso, una proteína de fusión aislada se expresa en una célula y, sin desear estar ligado por la teoría, se cree que se escinde el ADN como resultado de la formación de un dímero intramolecular de los medio dominios de escisión.

35 En ciertos aspectos, los componentes de las nucleasas de dedos de zinc (por ejemplo, las fusiones ZFP- *Fok I*) están dispuestos de tal manera que el dominio de dedos de zinc está más cerca del extremo amino-terminal de la proteína de fusión y el medio dominio de escisión está más cerca del extremo carboxi-terminal. Esto refleja la orientación relativa del dominio de escisión en dominios de escisión de dimerización de origen natural, tales como los derivados de la enzima *FokI*, en la que el dominio que se une a ADN está más cerca del extremo amino-terminal y el medio dominio de escisión está más cerca del extremo carboxi-terminal. En estos aspectos, la dimerización de los medio dominios de escisión para formar una nucleasa funcional está provocada por la unión de las proteínas de fusión a sitios en hebras de ADN opuestas, estando los extremos 5' de los sitios de unión próximos entre sí.

45 En aspectos adicionales, los componentes de las proteínas de fusión (por ejemplo, fusiones ZFP-*Fok I*) están dispuestos de tal manera que el medio dominio de escisión está más cerca del extremo amino-terminal de la proteína de fusión y el dominio de dedos de zinc está más cerca del extremo carboxi-terminal. En estos aspectos, la dimerización de los medio dominios de escisión para formar una nucleasa funcional está provocada por la unión de las proteínas de fusión a sitios en hebras de ADN opuestas, siendo los extremos 3' de los sitios de unión proximales entre sí.

50 En aspectos todavía adicionales, una primera proteína de fusión contiene el medio dominio de escisión más próximo al extremo amino-terminal de la proteína de fusión y el dominio de dedos de zinc más cerca del extremo carboxi-terminal y una segunda proteína de fusión se dispone de manera que el dominio de dedos de zinc está más cerca del extremo amino-terminal de la proteína de fusión y el medio dominio de escisión está más cerca del extremo carboxi-terminal. En estos aspectos, ambas proteínas de fusión se unen a la misma hebra de ADN, en donde el sitio de unión de la primera proteína de fusión que contiene el dominio de dedos de zinc más cercano al extremo carboxi-terminal está situado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio de dedos de zinc más cerca del extremo amino-terminal.

55 En ciertos aspectos, en las proteínas de fusión dadas a conocer, la secuencia de aminoácidos entre el dominio de dedos de zinc y el dominio de escisión (o el medio dominio de escisión) se denomina el "enlazador ZC". El enlazador

ZC se debe distinguir de los enlazadores entre los dedos descritos anteriormente. Véanse, por ejemplo, los documentos de publicaciones de Patente de EE.UU. 20050064474A1 y 20030232410 y la publicación de patente internacional WO05/084190, para más detalles sobre la obtención de enlazadores ZC que optimizan la escisión.

5 En una realización, la descripción proporciona una ZFN que comprende una proteína con dedos de zinc que tiene las secuencias de aminoácidos de reconocimiento de hélice mostradas en una fila simple de la Tabla 10B. En otro aspecto, se describe en esta memoria un vector de expresión de ZFN que comprende una secuencia de nucleótidos que codifica una ZFP que tiene hélices de reconocimiento que se muestran en la Tabla 1 o la Tabla 10.

#### Regulación de la expresión génica

10 Se puede emplear una variedad de ensayos para determinar si una ZFP modula la expresión génica. La actividad de una ZFP particular se puede evaluar utilizando una variedad de ensayos *in vitro* e *in vivo*, midiendo, por ejemplo, los niveles de proteínas o de ARNm, los niveles de producto, la actividad enzimática; la activación o inhibición de la transcripción de un gen informador, utilizando, por ejemplo, inmunoensayos (por ejemplo, ELISA y ensayos inmunohistoquímicos con anticuerpos), ensayos de hibridación (por ejemplo, protección de ARNasa, Northern, hibridación *in situ*, estudios de matrices de oligonucleótidos), ensayos colorimétricos, ensayos de amplificación, ensayos de actividad enzimática, ensayos fenotípicos y similares.

15 Las ZFPs se sometieron a ensayo típicamente para estudiar la actividad *in vitro* usando ensayos de ELISA y, a continuación, utilizando células renales. La ZFP se somete frecuentemente a ensayo primero utilizando un sistema de expresión transitorio con un gen informador y luego la regulación del gen endógeno diana se somete a ensayo en las células y en plantas enteras, tanto *in vivo* como *ex vivo*. La ZFP puede expresarse recombinantemente en una célula, expresada de forma recombinante en células trasplantadas en una planta, o expresada de manera recombinante en una planta transgénica, así como administrar como una proteína a la planta o la célula utilizando vehículos de administración descritos a continuación. Las células se pueden inmovilizar, estar en solución, inyectarse en una planta o ser de origen natural en una planta transgénica o no transgénica.

20 La modulación de la expresión génica se somete a ensayo utilizando uno de los ensayos *in vitro* o *in vivo* descritos en el presente documento. Las muestras o los ensayos se tratan con una ZFP y se comparan con muestras de control sin el compuesto de ensayo, para examinar el grado de modulación. Para la regulación de la expresión génica endógena, la ZFP tiene típicamente una  $K_d$  de 200 nM o menos, más preferiblemente 100 nM o menos, más preferiblemente 50 nM, lo más preferiblemente 25 nM o menos.

25 Los efectos de las ZFPs pueden medirse mediante un examen de cualquiera de los parámetros descritos anteriormente. Cualquier expresión génica, cambio fenotípico o fisiológico adecuado se puede usar para evaluar la influencia de una ZFP. Cuando las consecuencias funcionales se determinan usando células o plantas intactas, también se puede medir una variedad de efectos tales como el crecimiento de las plantas, cambios transcripcionales de marcadores genéticos conocidos y no caracterizados (por ejemplo, transferencias Northern o estudios de matrices de oligonucleótidos), cambios en el metabolismo celular tales como crecimiento celular o cambios de pH, y cambios en segundos mensajeros intracelulares tales como GMPc.

30 Los ensayos preferidos para la regulación mediante ZFP de la expresión génica endógena se pueden realizar *in vitro*. En un formato de ensayo preferido *in vitro*, la regulación mediante ZFP de la expresión génica endógena en células cultivadas se mide mediante el examen de la producción de proteína, utilizando un ensayo ELISA. La muestra del ensayo se compara con células de control tratadas con un vector vacío o una ZFP no relacionada que se dirige a otro gen.

35 En otro aspecto, la regulación mediante ZFP de la expresión génica endógena se determina *in vitro* midiendo el nivel de expresión del ARNm del gen diana. El nivel de expresión génica se mide utilizando la amplificación, por ejemplo, utilizando PCR, LCR o ensayos de hibridación, por ejemplo, hibridación Northern, protección de ARNasa, transferencia de manchas. El nivel de proteína o ARNm se detecta utilizando agentes de detección marcados directa o indirectamente, por ejemplo, ácidos nucleicos marcados con fluorescencia o radiactivamente, anticuerpos marcados radiactiva o enzimáticamente y similares, como se describe en el presente documento.

40 Alternativamente, un sistema de gen informador se puede diseñar usando el promotor del gen diana ligado funcionalmente a un gen informador tal como luciferasa, proteína fluorescente verde, CAT o  $\beta$ -gal. La estructura artificial informadora, se cotransfecta normalmente en una célula cultivada. Después del tratamiento con la ZFP de elección, la cantidad de transcripción, traducción o actividad del gen informador se mide de acuerdo con técnicas estándar conocidas por los expertos en la técnica.

45 Las plantas transgénicas y no transgénicas también se utilizan para examinar la regulación de la expresión de genes endógenos *in vivo*. Las plantas transgénicas pueden expresar de forma estable la ZFP de elección. Alternativamente, las plantas que expresan transitoriamente la ZFP de elección, o a las que ha administrado la ZFP en un vehículo de suministro, se pueden utilizar. La regulación de la expresión de genes endógenos se somete a ensayo utilizando uno cualquiera de los ensayos descritos en este documento.

Métodos para la escisión dirigida

- Los métodos y composiciones descritos se pueden usar para escindir el ADN en una región de interés en la cromatina celular (por ejemplo, en un sitio deseado o predeterminado en un genoma, por ejemplo, dentro o cerca de un gen implicado en la biosíntesis de ácidos grasos). Para tal escisión dirigida del ADN, un dominio de unión de dedos de zinc está diseñado para unirse a un sitio diana en o cerca del sitio de escisión predeterminado, y una proteína de fusión que comprende el dominio de unión de dedos de zinc modificado genéticamente y un dominio de escisión, se expresa en una célula. Tras la unión de la porción de dedos de zinc de la proteína de fusión en el sitio diana, el ADN se escinde cerca del sitio diana mediante el dominio de escisión. El sitio exacto de la escisión puede depender de la longitud del enlazador ZC.
- Alternativamente, dos proteínas de fusión, que comprenden cada una un dominio de unión de dedos de zinc y un medio dominio de escisión, se expresan en una célula y se unen a sitios diana que se yuxtaponen, de tal manera que un dominio de escisión funcional se reconstituye y el ADN se escinde en las proximidades de los sitios diana. En un aspecto, la escisión se produce entre los sitios diana de dos dominios de unión de dedos de zinc. Uno o ambos de los dominios de unión de dedos de zinc pueden estar modificados genéticamente.
- Para la escisión dirigida usando un polipéptido de fusión de dominio de unión de dedos de zinc-dominio de escisión, el sitio de unión puede abarcar el sitio de escisión, o el borde cerca 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 ubicación 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, que comprenden cada uno un dominio de unión de dedos de zinc y un medio dominio de escisión, los sitios de unión generalmente se extienden a ambos lados del sitio de escisión. Por lo tanto, el borde cercano al 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 corte y el borde cercano al 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 métodos para cartografiar sitios de escisión *in vitro* e *in vivo* son conocidos por los expertos en la técnica.
- Por lo tanto, los métodos descritos en el presente documento pueden emplear un dominio de unión de dedos de zinc modificado genéticamente, fusionado con un dominio de escisión. En estos casos, el dominio de unión está modificado genéticamente para unirse a una secuencia diana, en o cerca de donde se desea la escisión. La proteína de fusión, o un polinucleótido que la codifica, se introduce en una célula vegetal. Una vez introducida o expresada en la célula, la proteína de fusión se une a la secuencia diana y escinde la secuencia diana o cerca de la misma. El sitio exacto de la escisión depende de la naturaleza del dominio de escisión y/o de la presencia y/o la naturaleza de las secuencias enlazadoras entre los dominios de unión y escisión. En los casos en que se emplean dos proteínas de fusión, que comprenden cada una un medio dominio de escisión, la distancia entre los bordes cercanos de los sitios de unión puede ser de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos). Los niveles óptimos de escisión también pueden depender tanto de la distancia entre los puntos de unión de las dos proteínas de fusión (véase, por ejemplo, Smith et al. (2000) *Nucleic Acids Res.* 28:3361-3369; Bibikova et al. (2001) *Mol. Cell. Biol.* 21:289-297) y la longitud del enlazador ZC en cada proteína de fusión. Véanse, también, el documento de publicación de Patente de EE.UU. 20050064474A1 y las publicaciones de patentes internacionales WO05/084190, WO05/014791 y WO03/080809.
- En ciertos aspectos, el dominio de escisión comprende dos medio dominios de escisión, en donde ambos forman parte de un único polipéptido que comprende un dominio de unión, un primer medio dominio de escisión y un segundo medio dominio de escisión. Los medio dominios de escisión pueden tener la misma secuencia de aminoácidos o secuencias de aminoácidos diferentes, siempre que funcionen para escindir el ADN.
- Los medio dominios de escisión también se pueden proporcionar en moléculas distintas. Por ejemplo, dos polipéptidos de fusión pueden introducirse en una célula, en donde cada polipéptido comprende un dominio de unión y un medio dominio de escisión. Los medio dominios de escisión pueden tener la misma secuencia de aminoácidos o secuencias de aminoácidos diferentes, siempre que funcionen para escindir el ADN. Además, los dominios de unión se unen a secuencias diana que normalmente están dispuestas de tal manera que, tras la unión de los polipéptidos de fusión, los dos medio dominios de escisión se presentan en una orientación espacial entre sí que permite la reconstitución de un dominio de escisión (por ejemplo, mediante dimerización de los medio dominios), posicionando de este modo los medio dominios de escisión uno 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 ubicado entre las dos secuencias diana. Una o ambas de las proteínas puede estar modificada genéticamente para unirse a su sitio diana.
- Las dos proteínas de fusión se pueden unir en la región de interés en la misma polaridad u opuesta y sus sitios de unión (es decir, sitios diana) pueden estar separados por cualquier número de nucleótidos, por ejemplo, de 0 a 200 nucleótidos o cualquier valor entero entre los mismos. En ciertos aspectos, los sitios de unión para dos proteínas de fusión, en donde cada uno comprende un dominio de unión de dedos de zinc y un medio dominio de escisión, pueden estar situados con una separación de 5 y 18 nucleótidos, por ejemplo, a 5-8 nucleótidos de distancia, o 15-18 nucleótidos de distancia o 6 nucleótidos de distancia, o 16 nucleótidos de distancia, tal y como se mide desde el

borde de cada sitio de unión más cercano del otro sitio de unión y la escisión se produce entre los sitios de unión.

El sitio en el que se escinde el ADN generalmente se encuentra entre los sitios de unión para las dos proteínas de fusión. La rotura de la cadena doble de ADN tiene lugar frecuentemente a partir de dos roturas de hebras simples, o "muescas", compensadas por 1, 2, 3, 4, 5, 6 o más nucleótidos, (por ejemplo, una escisión de ADN de cadena doble mediante *Fok I* natural es el resultado de roturas de una cadena sencilla compensada por 4 nucleótidos). Por lo tanto, la escisión no se produce necesariamente en los sitios exactamente opuestos en cada cadena de ADN. Además, la estructura de las proteínas de fusión y la distancia entre los sitios diana pueden influir en si se produce la escisión de forma adyacente a una sola pareja de nucleótidos, o si la escisión se produce en varios sitios. Sin embargo, para muchas aplicaciones, incluyendo la recombinación dirigida y la mutagénesis dirigida (véase *infra*), la escisión dentro de un intervalo de nucleótidos generalmente es suficiente y no se requiere una escisión entre parejas de bases particulares.

Como se ha señalado anteriormente, la o las proteínas de fusión se pueden introducir como polipéptidos y/o polinucleótidos. Por ejemplo, dos polinucleótidos, en donde cada uno comprende secuencias que codifican uno de los polipéptidos mencionado anteriormente, se pueden introducir en una célula y cuando los polipéptidos se expresan y cada uno se une a su secuencia diana, se produce la escisión en o cerca de la secuencia diana. Alternativamente, un solo polinucleótido que comprende secuencias que codifican ambos polipéptidos de fusión, se introduce en una célula. Los polinucleótidos pueden ser ADN, ARN o cualquier forma o análogo modificado o ADN y/o ARN.

Para mejorar la especificidad de la escisión, las composiciones adicionales también se pueden emplear en los métodos descritos en el presente documento. Por ejemplo, medio dominios de escisión aislados pueden mostrar una actividad de escisión de cadena doble limitada. En métodos en los que dos proteínas de fusión, que contienen cada una un dominio de dedos de zinc de tres dedos y un medio dominio de escisión, se introducen en la célula, cada proteína especifica un sitio diana de 9 nucleótidos aproximadamente. Aunque es probable que la secuencia diana agregada de 18 nucleótidos sea única en un genoma de mamífero y vegetal, cualquier sitio diana dado de 9 nucleótidos aparece, en promedio, aproximadamente 23.000 veces en el genoma humano. Por lo tanto, una escisión no específica, debido a la unión específica del sitio de un solo medio dominio, se puede producir. En consecuencia, los métodos descritos en esta memoria contemplan el uso un mutante dominante negativo de un medio dominio de escisión tal como *Fok I* (o un ácido nucleico que lo codifica) que se expresa en una célula junto con las dos proteínas de fusión. El mutante dominante negativo es capaz de una dimerización, pero es incapaz de escindir y también bloquea la actividad de escisión de un medio dominio con el que está dimerizado. Al proporcionar el mutante dominante negativo en un exceso molar a las proteínas de fusión, solo las regiones en las que ambas proteínas de fusión están unidas tendrán una concentración local suficientemente alta de medio dominios de escisión funcionales para que se produzca la dimerización y la escisión. En los sitios en donde solo se une una de las dos proteínas de fusión, su medio dominio de escisión forma un dímero con el medio dominio mutante dominante negativo, y no tiene lugar una escisión no deseable, no específica.

### 35 Vectores de expresión

Un ácido nucleico que codifica una o varias proteínas (por ejemplo, ZFPs) como se describe en el presente documento, se puede clonar en un vector para la transformación en células procariotas o eucariotas para una replicación y/o expresión. Los vectores pueden ser vectores procariotas, por ejemplo, plásmidos, o vectores lanzadera, vectores de insectos, vectores víricos o vectores eucariotas. Un ácido nucleico que codifica una ZFP también se puede clonar en un vector de expresión para la administración a una célula vegetal.

Para expresar las ZFPs, las secuencias que codifican las ZFPs se subclonan típicamente en un vector de expresión que contiene un promotor para dirigir la transcripción. Promotores bacterianos y eucariotas adecuados son bien conocidos en la técnica 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* Stockton Press, New York (1990); y *Current Protocols in Molecular Biology* (Ausubel et al., supra. Bacterial expression systems for expressing the ZFP are available in, e.g., *E. coli*, *Bacillus* sp., and *Salmonella* (Palva et al., *Gene* 22:229-235 (1983)). Los kits para tales sistemas de expresión están disponibles comercialmente. Los sistemas de expresión eucariotas para células de mamíferos, plantas, levaduras e insectos son bien conocidos por los expertos en la técnica y por lo tanto están disponibles comercialmente.

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

En contraste, cuando se administra una ZFP *in vivo* para la regulación de un gen vegetal (véase, "Nucleic Acid Delivery to Plant Cells", a continuación), se utiliza ya sea un promotor constitutivo o inducible, dependiendo del uso particular de la ZFP. Ejemplos no limitativos de promotores de plantas incluyen secuencias promotoras obtenidas a partir de la ubicuitina-3 (*ubi-3*) de *A. thaliana* (Callis, et al., 1990, *J. Biol. Chem.* 265-12486-12493); la sintasa de manopina ( $\Delta$ mas) de *A. tumefaciens* (Petolino et al., documento de Patente de EE.UU. nº 6.730.824); y/o el virus del mosaico de las nervaduras de mandioca (Verdaguer et al., (1996), *Plant Molecular Biology* 31:1129-1139). Véase, también los Ejemplos.

Además del promotor, el vector de expresión contiene típicamente una unidad de transcripción o una casete de expresión que contiene todos los elementos adicionales requeridos para la expresión del ácido nucleico en células hospedadoras, procariotas o eucariotas. Una casete de expresión típico contiene por tanto un promotor ligado funcionalmente, por ejemplo, a una secuencia de ácido nucleico que codifica la ZFP y las señales requeridas, por ejemplo, para una poliadenilación eficiente del transcrito, terminación de la transcripción, sitios de unión a ribosoma o terminación de la traducción. Elementos adicionales de la casete pueden incluir, por ejemplo, potenciadores, señales de corte y empalme heterólogas y/o una señal de localización nuclear (NLS).

El vector de expresión particular utilizado para transportar la información genética en la célula se selecciona con respecto al uso pretendido de la ZFP, por ejemplo, una expresión en plantas, animales, bacterias, hongos, protozoos, etc. (véanse los vectores de expresión descritos más adelante). Los vectores de expresión bacteriana y animal convencionales son conocidos en la técnica y se describen en detalle, por ejemplo, en los documentos de Patente de EE.UU. nº de Publicación 20050064474A1 y las publicaciones de patentes internacionales WO05/084190, WO05/014791 y WO03/080809.

Los métodos de transfección convencionales se pueden emplear para producir líneas celulares bacterianas, de mamífero, vegetal, levadura o insecto que expresan grandes cantidades de proteína, que se pueden purificar después usando técnicas estándar (véase, por ejemplo, Colley et al., J. Biol. Chem. 264:17619-17622 (1989); Guide to Protein Purification, in Methods in Enzymology, vol. 182 (Deutscher, compilador, 1990)). La transformación de células eucariotas y procariotas se realiza según técnicas convencionales (véase, p. ej., Morrison, J. Bact. 132:349-351 (1977); Clark-Curtiss y Curtiss, Methods in Enzymology 101:347-362 (Wu et al., compiladores, 1983).

Se puede utilizar cualquiera de los procedimientos bien conocidos para introducir secuencias de nucleótidos extraños en tales células hospedadoras. Estos incluyen el uso de transfección con fosfato de calcio, polibreno, fusión de protoplastos, electroporación, métodos ultrasónicos (por ejemplo, sonoporación), liposomas, microinyección, ADN desnudo, vectores plasmídicos, vectores víricos, tanto episomales como integradores y cualquier otro método bien conocido para la introducción de ADN clonado genómico, ADNc, ADN sintético u otro material genético extraño en una célula hospedadora (véase, por ejemplo, Sambrook et al., *supra*). Solo es necesario que el procedimiento de modificación genética particular utilizado sea capaz de introducir con éxito al menos un gen en la célula hospedadora capaz de expresar la proteína de elección.

#### Entrega de ácido nucleico a células vegetales

Como se ha señalado anteriormente, las estructuras artificiales de ADN se pueden introducir (por ejemplo, en el genoma de) una planta hospedadora deseada a través de una diversidad de técnicas convencionales. Para revisiones de tales técnicas véase, por ejemplo, Weissbach y Weissbach, Methods for Plant Molecular Biology (1988, Academic Press, N.Y.) sección VIII, p. 421-463; y Grierson y Corey, Plant Molecular Biology (1988, 2ª ed.), Blackie, London, Cap. 7-9.

Por ejemplo, la estructura artificial de ADN puede introducirse directamente en el ADN genómico de la célula vegetal utilizando técnicas tales como electroporación y microinyección de protoplastos de células vegetales, o las estructuras artificiales de ADN se pueden introducir directamente en el tejido vegetal utilizando métodos biolísticos, tales como bombardeo de partículas de ADN (véase, por ejemplo, Klein et al., (1987) Nature 327: 70-73). Alternativamente, las estructuras artificiales de ADN pueden combinarse con regiones flanqueantes de T-ADN adecuadas e introducirse en un vector hospedador convencional de *Agrobacterium tumefaciens*. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, incluyendo el desarme y el uso de vectores binarios, están bien descritas en las publicaciones científicas. Véase, por ejemplo Horsch et al., (1984) Science 233:496-498, y Fraley et al., (1983) Proc. Natl. Acad. Sci. USA 80:4803.

Además, la transferencia de genes se puede conseguir usando bacterias que no sean *Agrobacterium tumefaciens* 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 nervaduras de mandioca 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 estructura artificial y el marcador adyacente en el ADN celular de la planta, cuando la célula es infectada por la bacteria utilizando un vector binario T-ADN (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 de *Agrobacterium* se utiliza para diseñar plantas dicotiledóneas (Bevan et al., (1982) Ann. Rev. Genet. 16:357-384; Rogers et al., (1986) Methods Enzymol. 118:627-641). El sistema de transformación de *Agrobacterium* también se puede usar para transformar, así como para transferir, ADN a plantas monocotiledóneas y células vegetales. Véase el documento de Patente de EE.UU. nº 5 591.616; Hernalsteen et al., (1984) EMBO J3:3039-3041; 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.

Una transferencia de genes y métodos de transformación alternativos incluyen, pero no se limitan a, una transformación de protoplastos a través de captación mediada por calcio, polietilenglicol (PEG) o electroporación de ADN des-

nudo (véase, Paszkowski et al. (1984) EMBO J3:2717-2722, Potrykus et al. (1985) Molec. Gen. Genet. 199:169-177; Fromm et al. (1985) Proc. Nat. Acad. Sci. USA 82:5824-5828; y Shimamoto (1989) Nature 338:274-276)) y electroporación de tejidos vegetales (D'Halluin et al. (1992) Plant Cell 4:1495-1505). Los métodos adicionales para la transformación de células vegetales incluyen microinyección, captación de ADN mediada por carburo de silicio (Kaepler et al. (1990) Plant Cell Reporter 9:415-418) y bombardeo con microproyectiles (véase, Klein et al. (1988) Proc. Nat. Acad. Sci. USA 85:4305-4309; y Gordon-Kamm et al. (1990) Plant Cell 2:603-618), o nanopartículas.

Los métodos y composiciones se pueden utilizar para insertar secuencias exógenas en una localización predeterminada en un genoma de célula vegetal. Esto es útil en la medida que la expresión de un transgén introducido en un genoma vegetal depende decisivamente de su sitio de integración. En consecuencia, los genes que codifican, por ejemplo, nutrientes, antibióticos o moléculas terapéuticas se pueden insertar mediante recombinación dirigida, en regiones de un genoma vegetal favorable para su expresión.

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

Los ácidos nucleicos introducidos en una célula vegetal se pueden utilizar para conferir rasgos deseados en esencialmente cualquier planta. Una amplia variedad de plantas y sistemas de células vegetales se puede modificar genéticamente para obtener las características fisiológicas y agronómicas deseadas, descritas en este documento usando las estructuras artificiales de ácido nucleico de la presente memoria y diversos métodos de transformación mencionados anteriormente. Las plantas diana y las células vegetales para modificar genéticamente incluyen, pero no se limitan a, aquellas plantas monocotiledóneas y dicotiledóneas, tales como cultivos que incluyen cultivos de cereales (por ejemplo, trigo, maíz, arroz, mijo, cebada), cultivos de frutas (por ejemplo, tomate, manzana, pera, fresa, naranja), cultivos forrajeros (por ejemplo, alfalfa), cultivos de hortalizas de raíz (por ejemplo, zanahoria, patata, remolacha de azúcar, ñame), cultivos de hortalizas de hoja (por ejemplo, lechuga, espinaca); plantas con flores (por ejemplo, petunia, rosa, crisantemo), árboles coníferas y pinos (por ejemplo, pino, abeto); plantas utilizadas en la fitorremediación (por ejemplo, plantas que acumulan metales pesados); cultivos de aceite (por ejemplo, girasol, semilla de colza, soja) y plantas utilizadas con fines experimentales (por ejemplo, *Arabidopsis*). Por lo tanto, los métodos y composiciones descritos emplean una amplia gama de plantas, incluyendo, pero no limitadas a, especies de los géneros Asparagus, Avena, Brassica, Citrus, Citrullus, Capsicum, Cucurbita, Daucus, Erigeron, Glycine, Gossypium, Hordeum, Lactuca, Lolium, Lycopersicon, Malus, Manihot, Nicotiana, Orychophragmus, Oryza, Persea, Phaseolus, Pisum, Pyrus, Prunus, Raphanus, Secale, Solanum, Sorghum, Triticum, Vitis, Vigna y Zea.

Un experto en la técnica reconocerá que después de incorporar la casete de expresión de manera estable en plantas transgénicas y confirmar que es operativa, se puede introducir en otras plantas por cruzamiento sexual. Cualquiera entre una serie de técnicas de reproducción estándar se puede utilizar, dependiendo de la especie que se va a cruzar.

Una célula vegetal, callo, tejido o planta transformada se puede identificar y aislar mediante la selección o el escrutinio del material vegetal modificado genéticamente, por los rasgos codificados por los genes marcadores presentes en el ADN transformante. Por ejemplo, la selección puede llevarse a cabo haciendo crecer el material vegetal modificado genéticamente en medios que contienen una cantidad inhibitoria del antibiótico o el herbicida frente al que la estructura artificial del gen transformando confiere resistencia. Además, las plantas transformadas y las células vegetales también pueden identificarse mediante el escrutinio de las actividades de cualquier gen marcador visible (por ejemplo, la  $\beta$ -glucuronidasa, luciferasa, genes B o C1) que pueden estar presentes en las estructuras artificiales de ácidos nucleicos recombinantes. Tales metodologías de selección y escrutinio son bien conocidas por los expertos en la técnica.

También se pueden emplear métodos físicos y bioquímicos para identificar transformantes de plantas o de células vegetales que contienen estructuras artificiales de genes insertadas. Estos métodos incluyen, pero no se limitan a: 1) el análisis con Southern o la amplificación con PCR para la detección y determinación de la estructura del inserto de ADN recombinante; 2) transferencia Northern, protección de ARNasa S1, amplificación por PCR con extensión del cebador o transcriptasa inversa para la detección y el examen de los transcritos de ARN de las estructuras artificiales génicas; 3) ensayos enzimáticos para la detección de la actividad enzimática o de ribozimas, en donde tales productos génicos están codificados por la estructura artificial del gen; 4) electroforesis en gel de proteínas, técnicas de transferencia Western, inmunoprecipitación o inmunoensayos ligados a enzimas, en donde los productos de la estructura artificial del gen son proteínas. Técnicas adicionales, tales como hibridación *in situ*, tinción enzimática e inmunotinción, también se pueden usar para detectar la presencia o la expresión de la estructura artificial recombinante en órganos y tejidos específicos de la planta. Los métodos para realizar todos esos ensayos son bien conoci-

dos por los expertos en la técnica.

Los efectos de la manipulación de genes utilizando los métodos descritos en esta memoria se pueden observar, por ejemplo, mediante transferencias Northern del ARN (por ejemplo, ARNm) aislado a partir de los tejidos de interés. Típicamente, si la cantidad de ARNm se ha incrementado, se puede suponer que el gen endógeno correspondiente se expresa con una tasa mayor que antes. Otros métodos para medir la actividad del gen y/o CYP74B, se pueden utilizar. Se pueden utilizar diferentes tipos de ensayos enzimáticos, dependiendo del sustrato utilizado y del método para detectar el aumento o disminución de un producto de reacción o subproducto. Además, los niveles de y/o proteína CYP74B expresada se pueden medir inmunoquímicamente, es decir, ELISA, RIA, EIA y otros ensayos basados en anticuerpos, bien conocidos por los expertos en la técnica, tales como mediante ensayos de detección electroforética (ya sea con tinción o transferencia Western). El transgén se puede expresar de forma selectiva en algunos tejidos de la planta o en algunas etapas de desarrollo, o el transgén se puede expresar en sustancialmente todos los tejidos de la planta, sustancialmente a lo largo de todo su ciclo vital. Sin embargo, cualquier modo de expresión combinatoria es también aplicable.

La presente memoria descriptiva también incluye semillas de las plantas transgénicas descritas anteriormente, en donde la semilla tiene el transgén o la estructura artificial génica, e incluye semillas con perfiles de aceite modificados deseados. La presente memoria descriptiva incluye además la progenie, clones, líneas celulares o células de las plantas transgénicas descritas anteriormente, en donde dicha progenie, clon, línea celular o célula tiene el transgén o la estructura artificial génica.

La administración de cantidades eficaces se realiza por cualquiera de las rutas utilizadas normalmente para introducir ZFPs en contacto final con la célula de la planta que se va a tratar. Las ZFPs se administran de cualquier manera adecuada, preferiblemente con vehículos aceptables. Los métodos adecuados de administración de tales moduladores están disponibles y son bien conocidos por los expertos en la técnica, y, aunque se puede utilizar más de una ruta para administrar una composición particular, una ruta en particular puede proporcionar frecuentemente una reacción más inmediata y más efectiva que otra ruta.

Los vehículos también se pueden utilizar y se determinan en parte por la composición particular que se administra, así como por el método particular usado para administrar la composición. En consecuencia, existe una amplia variedad de formulaciones adecuadas de composiciones que están disponibles.

#### Aplicaciones

Como se ha señalado anteriormente, una modulación dirigida de genes implicados en la síntesis de ácidos grasos se puede utilizar para generar rápida y eficientemente aceites vegetales con el perfil de ácidos grasos deseado.

La alteración de las composiciones de ácidos grasos de los aceites vegetales procedentes de plantas productoras de aceite, tales como, por ejemplo, canola, tiene profundas implicaciones para la producción de alimentos y, en consecuencia, sobre la salud de la dieta. Por ejemplo, variedades de semillas oleaginosas de *Brassica* de calidad canola con niveles reducidos de ácidos grasos saturados en el aceite de la semilla, se pueden utilizar para producir productos alimenticios que favorecen la salud cardiovascular. Los métodos y las composiciones descritas en este documento pueden usarse para generar plantas productoras de aceite con bajo contenido en ácido palmítico y/o esteárico, lo que reduce los niveles de ácidos grasos saturados en el aceite vegetal. Del mismo modo, la generación de plantas productoras de aceite con niveles incrementados de contenido en ácido oleico, también reducirá la cantidad de ácidos grasos saturados en el aceite vegetal.

Además, los métodos y las composiciones descritas en el presente documento se pueden usar para aumentar el contenido en ácido palmítico, por ejemplo, en aceites producidos a partir de tales plantas. Aceites con alto contenido en ácido palmítico son particularmente útiles en la formulación de margarinas.

Además, una alteración dirigida de los perfiles de ácidos grasos como se describe en el presente documento, se puede usar para generar plantas (y aceites vegetales) que son bajos en ácido linolénico. Los aceites con bajo contenido en ácido linolénico muestran un aumento de la estabilidad; alimentos preparados usando estos aceites no se estropean en términos de sabor u olor tan rápido como los alimentos elaborados a partir de materiales vegetales con concentraciones elevadas de ácido linolénico.

#### Ejemplos

A continuación se presentan ejemplos de aspectos específicos de la presente memoria descriptiva.

Se han realizado esfuerzos para asegurar la exactitud con respecto a los números utilizados (por ejemplo, cantidades, temperaturas, etc.), pero se deben permitir por supuesto algunos errores y desviaciones experimentales.

Ejemplo 1: Identificación de la secuencia diana en *Brassica napus* L

##### 1.1 Identificación de la secuencia diana

En este ejemplo, se identificaron las secuencias de ADN endógenas que codifican la enzima natural sintetasa II de

$\beta$ -cetoacil-ACP (KAS II) para *Brassica napus* L (canola). Estos genes se seleccionaron como dianas ejemplares para demostrar una regulación transcripcional por medio de factores de transcripción de proteínas de dedos de zinc (ZFP TF) modificadas genéticamente, dando como resultado la modificación deseada de la biosíntesis de ácidos grasos y los perfiles de aceite de semilla alterados concomitantes. La enzima sintetasa II de  $\beta$ -cetoacil-ACP cataliza la conversión de 16:0-ACP a 18:0-ACP y la posterior formación de ácido oleico (Ohlrogge y Browse, 1995, *The Plant Cell*, 7: 957-970). Un mutante *fab1* de *Arabidopsis thaliana* indicado del gen de la sintetasa II de  $\beta$ -cetoacil-ACP producía una reducción del 65% de la actividad de la enzima y un aumento concomitante del contenido en ácido esteárico del 7% y el 3% en las hojas y las raíces, respectivamente (Wu et al., 1994, *Plant Physiology*, 106: 143-150). Un transgén con expresión estable transformado de la sintetasa II de  $\beta$ -cetoacil-ACP de soja en canola, reducía el contenido en ácido palmítico de las semillas en un 0,8% y la introducción del mismo gen en el tabaco disminuía el contenido en ácido palmítico en un 2% (documento de Publicación de Patente Japonesa n° 501446/1995).

Las secuencias del ADNc de la sintetasa II de  $\beta$ -cetoacil-ACP se han descrito en múltiples especies de plantas, incluyendo *Arabidopsis thaliana* (GenBank: AF318307) y *Brassica napus* (GenBank: AF244520). La alineación de las secuencias de ADNc de *A. thaliana* y *B. napus* (GenBank: AF318307 y AF244520, respectivamente) mostró que la secuencia AF244520 era incompleta (Figura 7). En esta secuencia de ADN truncado de *B. napus* faltaban varios cientos de pares de bases en el extremo 5'. Además, puesto que *B. napus* es una especie anfidiplode resultante de la combinación de los conjuntos de cromosomas de *B. rapa* ( $2n = 20$ , AA) y *B. oleracea* ( $2n = 18$ , CC) (Morinaga, 1934 *Cytologia*, 6: 62-67; U.N., 1935, *Japanese J. Bot.*, 7: 389-452), se predice que habría más de un gen de la sintetasa II de  $\beta$ -cetoacil-ACP en esta especie. Las secuencias 5' UTR adicionales presentes en las secuencias de ADNc fueron identificadas y se obtuvieron. Estas secuencias 5' de genes de la sintetasa II de  $\beta$ -cetoacil-ACP sirven como dianas para una regulación al alza de la transcripción a través de ZFP TFs en los presentes ejemplos.

### 1.2 Aislamiento del ARN total

El ARN total se aisló a partir de semillas inmaduras de *Brassica napus* (canola) genotipo Nex710 (Crop Certificate 99-7049208-501), 15 días después de la floración (DAF), utilizando el minikit de Qiagen, RNEASY® Plant Mini Kit (Qiagen, Valencia, CA). El ARNm total se trató con ADNasa exenta de ARNasa según las recomendaciones del fabricante, para eliminar cualquier ADN contaminante que se pudiera amplificar durante la RT-PCR cuantitativa

### 1.3 RACE 5' y análisis de la secuencia

La rápida amplificación de los extremos de ADNc (RACE) específica para el extremo 5' del ADNc de la sintetasa II de  $\beta$ -cetoacil-ACP de *B. napus* (GenBank AF244520) se realizó empleando el kit FirstChoice® RLM-RACE de Ambion (Ambion, Austin, TX) según las recomendaciones del fabricante. Para obtener la secuencia del ADNc 5', un adaptador de ARN sintético se asoció con la región 5' del ARNm sin bloquear. Un cebador suministrado con el kit (5'-CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATGAAA - 3') (SEQ ID NO:1) y un segundo cebador (5'-CTCGAGCTGCTACTGCTAGTTCGGTGGAGGAGCC - 3') (SEQ ID NO:2), que se había diseñado para unirse dentro de la secuencia parcial de ADNc de *B. napus* (GenBank AF244520) se utilizaron para amplificar un fragmento para determinar la secuencia aguas arriba desconocida. La amplificación RACE 5' identificó aproximadamente 500 pares de bases de nuevas secuencias de *B. napus*. Cántigos de cuatro secuencias de ADNc 5' claramente diferentes de genes de sintetasa II de  $\beta$ -cetoacil-ACP (SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 46 y SEQ ID NO: 30) se identificaron cuando las secuencias amplificadas fueron alineadas (Figura 8). Estas secuencias 5' mostraron altos niveles de homología con secuencias de *B. rapa* (GenBank: AC189461) y *B. oleracea* (GenBank: BH723504) de la misma región.

Las secuencias de unión de ZFP TF fueron identificadas en las secuencias de ADNc 5' de los genes de la sintetasa II de  $\beta$ -cetoacil-ACP. Las secuencias procedentes de la región aguas arriba de los genes de la sintetasa II de  $\beta$ -cetoacil-ACP (Tabla 2) servían como dianas para la unión de ZFP TF (Tabla 1). Estructuras artificiales de plásmidos: pDAB4695, pDAB4696, pDAB4697 y pDAB4698 que contenían los diseños de los ZFP TFs modificados genéticamente se utilizaron para transformar de manera estable *B. napus* como se describe en las siguientes secciones. Se planteó la hipótesis de que los ZFP TFs modificados genéticamente, tras la expresión en células vegetales, se unirían a las dianas endógenas de la sintetasa II de  $\beta$ -cetoacil-ACP, dando como resultado una expresión modificada del ARNm de la diana.

Ejemplo 2: Diseño de dominios que se unen a ADN de ZFP específicos del gen de la sintetasa II de  $\beta$ -cetoacil-ACP

Se diseñaron proteínas con dedos de zinc para varios sitios dianas en la región promotora del gen de la sintetasa II de  $\beta$ -cetoacil-ACP de *B. napus* y en la región 5' no traducida y traducida. Véase, Figura 2. Las hélices de reconocimiento para los diseños de ZFP representativos se muestran a continuación en la Tabla 1. Los sitios diana de los diseños de dedos de zinc se muestran a continuación en la Tabla 2.

Tabla 1: Diseños de los dedos de zinc de la sintetasa II de  $\beta$ -cetoacil-ACP

ZFP	F1	F2	F3	F4	F5	F6
14025	RSDNLSV (SEQ ID NO:5)	QKINLQV (SEQ ID NO: 6)	RSDTLSE (SEQ ID NO:7)	TRSSRIN (SEQ ID NO:8)	RSDALAR (SEQ ID NO:9)	N/A
14033	RSDHLSA (SEQ ID NO:10)	TSSSRIN (SEQ ID NO:11)	RSDNLAR (SEQ ID NO:12)	DRSHLAR (SEQ ID NO:13)	RSDNLSE (SEQ ID NO:14)	RNAHRTT (SEQ ID NO:15)
14035	QSGNLAR (SEQ ID NO:16)	RSDHLSE (SEQ ID NO:17)	QKANRTK (SEQ ID NO:18)	RSDDLTR (SEQ ID NO:19)	TSANLSR (SEQ ID NO:20)	N/A
14047	RSDDLK (SEQ ID NO:21)	RSANLTR (SEQ ID NO:22)	RSDDLTR (SEQ ID NO:19)	RSDHLSE (SEQ ID NO:17)	DKSNRKK (SEQ ID NO:23)	N/A

Tabla 2: Sitios diana de los dedos de zinc de la sintetasa II de  $\beta$ -cetoacil-ACP

ZFP	Sitio diana (5' a 3')	Sitio diana/unión a ZFP presente en SEQ ID Nos.
14025	cGTGGAGACGtCAAAGa (SEQ ID NO: 24)	3, 4, 46 y 30
14033	aAGGAAGGGCGAGAAAAGGg (SEQ ID NO: 25)	3 y 4
14035	aGATGCGTAACAGGAAg (SEQ ID NO: 26)	3, 4, 46 y 30
14047	cTACCGGGCGGAGTCGt (SEQ ID NO: 27)	3, 4 y 30

5 Los diseños de la sintetasa II de  $\beta$ -cetoacil-ACP se incorporaron en vectores de expresión de dedos de zinc que codificaban una proteína que tenía al menos un dedo con una estructura CCHC. Véase, el documento de patente de EE.UU. n° de publicación 2008/0182332. En particular, el último dedo en cada proteína tenía una estructura CCHC (arquitectura).

10 Las secuencias que codificaban los dedos de zinc se fusionaron después con secuencias que codificaban un dominio de activación VP16 y una señal de localización nuclear *opaque-2*. La expresión de las proteínas de fusión estaba impulsada por un promotor constitutivo relativamente fuerte, tal como un promotor obtenido a partir del promotor de ubicuitina 10 de *Arabidopsis thaliana* (AtUbi10). Los vectores ejemplares se muestran en la Tabla 3 a continuación.

Ejemplo 3: Regulación al alza mediada por ZFP TF del o los genes de la sintetasa II de  $\beta$ -cetoacil-ACP en *B. napus*

15 Con el fin de evaluar la funcionalidad de las proteínas con dedos de zinc diseñadas en células vegetales, se utilizaron métodos para la expresión de tales proteínas en células vegetales vivas. El ADN que codifica proteínas con dedos de zinc se puede entregar en las células vegetales en condiciones en las que el ADN no se incorpora en el genoma de la célula vegetal. Por lo tanto, la molécula de ADN se mantiene de forma transitoria en las células vegetales y actúa como un molde para la expresión génica. Alternativamente, el ADN que codifica las proteínas con dedos de zinc se puede entregar en células vegetales en condiciones que permiten que el ADN se incorpore en el genoma de la célula vegetal, lo que produce una transgénesis de los genes que codifican la proteína de dedos de zinc, de tal manera que la molécula de ADN se mantiene de forma estable en las células vegetales y actúa como un molde para la expresión génica. Un experto en la técnica puede utilizar una expresión transitoria o transgénica de los ADNs que codifican las proteínas de dedos de zinc, con el fin de determinar la funcionalidad de esas proteínas en células vegetales vivas.

### 25 3.1 Diseño de estructuras artificiales

Los plásmidos binarios diseñados y construidos para este proyecto se especifican en la Tabla 3.

Tabla 3: Descripción de estructuras artificiales para ZFP TFs dirigidos a KAS II de *B. napus*

S.N.	ZFP	Estructura artificial nº	Casete génico
1	14025	pDAB4695	Atubi10/ZFP1-Vp16/AtuORF23/CsVMV/ <i>pat</i> /AtuORF1
2	14033	pDAB4696	Atubi10/ZFP2-Vp16/AtuORF23/CsVMV/ <i>pat</i> /AtuORF1
3	14035	pDAB4697	Atubi10/ZFP3-Vp16/AtuORF23/CsVMV/ <i>pat</i> /AtuORF1
4	14047	pDAB4698	Atubi10/ZFP4-Vp16/AtuORF23/CsVMV/ <i>pat</i> /AtuORF1

AtUbi10 = promotor de ubicuitina 10 de *Arabidopsis thaliana*, CsVMV = promotor del virus del mosaico de las nervaduras de mandioca, ZFP = gen de la proteína de dedos de zinc, *pat* = gen de la transferasa de fosfinotricin acilo, AtuORF1 = 3' UTR 1 de *Agrobacterium tumefaciens* y AtuORF23 = 3' UTR 23 de *Agrobacterium tumefaciens*.

5 pDAB4695 es un plásmido binario que contiene las casetes de expresión génica 14025 v3 / VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; RB7 MAR v3 (región de fijación a la matriz (Thompson et al., 1997, documento W09727207)) :: AtUbi10 Promotor v2 (promotor de ubicuitina 10 de *Arabidopsis thaliana* (Callis, et al., 1990, J. Biol. Chem. 265-12486-12493)) :: fusión 14025 v3 dedo de zinc / VP16 :: Atu ORF23 3'UTR v1 (marco de lectura abierto 23 de *Agrobacterium tumefaciens*, región 3' no traducida (Gelvin et al., 1987, documento EP222493)) :: CsVMV Promotor v2 (promotor del virus del mosaico de las nervaduras de mandioca (Verdaguer et al., 1996, Plant Molecular Biology 31: 1129-1139)) :: *pat* v5 (transferasa de fosfinotricin acetilo (Wohlleben et al., 1988, Gene 70: 25-37)) :: AtuORF 1 3'UTR v4 (marco de lectura abierto 1 de *Agrobacterium tumefaciens*, región 3' no traducida (Huang et al., J. Bacteriol. 172: 1814-1822)). El binario se construyó clonando un fragmento de ADN que contenía la fusión 14025 v3 dedo de zinc / VP16 en pDAB3916 a través de los sitios de restricción *NcoI-SacI*. La estructura artificial resultante que se etiquetó como pDAB8221 contenía la casete de expresión génica AtUbi10 Promotor v2 :: fusión 14025 v3 dedo de zinc / VP16 :: Atu ORF23 3'UTR v1. pDAB8221 se clonó en el binario pDAB7309 a través de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción producía pDAB4695 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

20 pDAB4696 es un plásmido binario que contiene las casetes de expresión génica 14033 v3 / VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; RB7 MAR v3 :: AtUbi10 Promotor v2 :: fusión 14033 v3 dedo de zinc / VP16 :: Atu ORF23 3'UTR v1 :: CsVMV Promotor v2 :: *pat* v5 :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión 14033 v3 dedo de zinc / VP16 en pDAB3916 a través de los sitios de restricción *NcoI-SacI*. La estructura artificial resultante que se etiquetó como pDAB8222 contenía la casete de expresión génica AtUbi10 Promotor v2 :: fusión 14033 v3 dedo de zinc / VP16 :: Atu ORF3 3'UTR v1. pDAB8222 se clonó en el binario pDAB7309 a través de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción producía Pdad4696 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

25 pDAB4697 es un plásmido binario que contiene las casetes de expresión génica 14035 v3 / VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; RB7 MAR v3 :: AtUbi10 Promotor v2 :: fusión 14035 v3 dedo de zinc / VP16 :: Atu ORF23 3'UTR v1 :: CsVMV Promotor v2 :: *pat* v5 :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión 14035 v3 dedo de zinc / VP16 en pDAB3916 a través de los sitios de restricción *NcoI-SacI*. La estructura artificial resultante que se etiquetó como pDAB8223 contenía la casete de expresión génica AtUbi10 Promotor v2 :: fusión 14035 v3 dedo de zinc / VP16 :: Atu ORF3 3'UTR v1. pDAB8223 se clonó en el binario pDAB7309 a través de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción producía pDAB4697 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

35 pDAB4698 es un plásmido binario que contiene las casetes de expresión génica 14047 v3 / VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; RB7 MAR v3 :: AtUbi10 Promotor v2 :: fusión 14047 v3 dedo de zinc / VP16 :: Atu ORF23 3'UTR v1 :: CsVMV Promotor v2 :: *pat* v5 :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión 14047 v3 dedo de zinc / VP16 en pDAB3916 a través de los sitios de restricción *NcoI-SacI*. La estructura artificial resultante que se etiquetó como pDAB8224 contenía la casete de expresión génica AtUbi10 Promotor v2 :: fusión 14047 v3 dedo de zinc / VP16 :: Atu ORF3 3'UTR v1. pDAB8224 se clonó en el binario pDAB7309 a través de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción producía pDAB4698 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

### 3.2 Transformación de *Agrobacterium*

Las células de *Agrobacterium* se prepararon para la electroporación utilizando el protocolo descrito en Weigel D., Glazebrook J. *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2002: p. 123-124. Se hicieron modificaciones menores del protocolo que permitieron un crecimiento óptimo de *Agrobacterium* (es decir, el

medio YEP se sustituyó por LB). Independientemente, se añadieron 1,5-3 µg de ADN plasmídico a cada estructura artificial a 50 µl de células de *Agrobacterium tumefaciens* C58 :: Z707s y se mezclaron suavemente. La mezcla se transfirió a cubetas frías de 0,2 cm de GENE PULSER® (BioRad, Hercules, CA) y se colocaron sobre hielo. Las cubetas se colocaron después en un estante frío de GENE PULSER® (BioRad, Hercules, CA) y se electroporaron con las siguientes condiciones: salida de capacitancia 25 µFaradios, extensor de capacitancia 960 µFaradios, resistencia 200 ohmios y voltaje 2,5 kV. Inmediatamente después de la electroporación, se añadieron 950 µl de medio SOC (Invitrogen, Carlsbad, CA) y la mezcla se transfirió a un tubo Falcon 2059 (Becton Dickinson y Co., Franklin Lakes, NJ) o uno equivalente. Las células transformadas se incubaron a continuación a 28°C durante 1 hora. Después de la incubación, 50 µl, 100 µl y 200 µl de células se sembraron en placas distintas con medio YEP (10 g de extracto de levadura, 10 g de peptona, 5 g de NaCl, 10 g de sacarosa y 15 g de agar en 1 litro de agua) que contenía antibióticos de forma apropiada. Las placas se cultivaron invertidas a 28°C durante aproximadamente 36-48 horas. Las colonias individuales se seleccionaron y propagaron en 50 ml de YEP líquido (10 g de extracto de levadura, 10 g de peptona, 5 g de NaCl y 10 g de sacarosa en 1 litro de agua), que contenía antibióticos de forma apropiada, a 28°C durante aproximadamente 36 horas.

La cepa de *Agrobacterium tumefaciens* se confirmó a través de una prueba de cetolactosa. Las colonias transformadas de forma putativa se sembraron en estrías sobre agar lactosa y se incubaron a 28°C durante 48 horas. Después las placas se inundaron después con solución de Benedict. Las placas se controlaron; los aislados de estrías que cambiaban la solución de Benedict de azul a amarillo se confirmaron como *Agrobacterium* (Bouzar, H., Jones, J., Bishop, A. "Simple Cultural Tests for Identification of *Agrobacterium* Biovars." Methods in Molecular Biology, Volumen 44. Humana Press, 1995: 9-13).

Después de completar un protocolo de mini-prep de pocas copias de QIAGEN® (Qiagen, Valencia, CA), el ADN plasmídico purificado se preparó a partir de cultivos bacterianos. La integridad del ADN se evaluó mediante digestión con restricción. Los clones con los patrones de bandas esperados se identificaron y se prepararon reservas en glicerol mediante la adición de 500 µl de cultivo bacteriano a 500 µl de glicerol estéril (Sigma Chemical Co., St. Louis, MO) e invirtiendo para mezclar. Las reservas de glicerol se congelaron sobre hielo seco y se almacenaron a -80°C.

### 3.3 Transformación de *B. napus* con ZFP TFs

Preparación de segmentos de hipocótilo: semillas de *B. napus* con genotipo Nex710, se esterilizaron en superficie con 10% de lejía comercial durante 10 minutos y se lavaron 3 veces con agua destilada estéril. Las semillas se secaron con una toalla de papel estéril, después se colocaron en una bandeja Phyta que contenía 'medio de germinación' que consistía en media concentración de medio basal MS (Murashige y Skoog, Physiol Plant 15(3):473-497, 1962), 20 g/L de sacarosa y 8 g/L de agar TC (PhytoTechnology Laboratories, Shawnee Mission, KS) y se mantuvieron con un régimen de crecimiento fijado en 23°C y un fotoperiodo de 16 h de luz/8 h oscuridad.

El día 5, se comprobó la esterilidad de las plántulas y la bandeja Phyta se colocó dentro de una campana de flujo laminar EDGE GARD® (The Baker Company, Sanford, ME) para mantener la esterilidad. Mediante el uso de pinzas estériles y tijeras de disección, las plantas se retiraron de la bandeja Phyta y la región aérea (meristemo y cotiledón) y las raíces se separaron y se desecharon. Los hipocótilos se colocaron en una placa de Petri de 100 x 25 mm que contenía agua destilada estéril que se quería para evitar el secado. Los hipocótilos se colocaron sobre la tapa de una placa de Petri de 100 x 25 mm y se cortaron transversalmente en segmentos de 3 mm utilizando un bisturí del nº 10. Los segmentos de hipocótilo se colocaron en "medio de inducción de callos" que consistía en medio MS que contenía 30 g/L de sacarosa, 1 mg/L de cinetina y 1 mg/L de 2,4-D solidificada con 7 g/L de agar, TC incluyendo un papel de filtro estéril. Las placas se colocaron en una cubeta transparente STERILITE® y se mantuvieron con el mismo régimen de crecimiento durante 3 días como tratamiento previo.

El día 8, los segmentos se transfirieron después a una placa de Petri estéril de 100 x 25 que contenía 20 ml de 'medio de cultivo líquido' que consistía en medio basal Skoog y Linsmaier (1965) que contenía 30 g/L de glucosa, vitaminas Gamborg de ½ fuerza (1968), 215,2 mg/L de cinetina y 221,04 mg/L de 2,4-D para un tratamiento previo de 1 hora. El 'medio de cultivo líquido' se eliminó de los segmentos de hipocótilo y 40 ml de suspensión de *Agrobacterium* (que contenía pDAB4695, pDAB4696, pDAB4697 o pDAB4698 en Z707s) a una 50 Klett, se sometió a vórtice brevemente y se vertió en la placa de Petri de 100 x 25 mm que contenía segmentos de hipocótilo para un tratamiento de 30 minutos. Después de 30 minutos, toda la suspensión de *Agrobacterium* se retiró usando una pipeta doble apilada. Los hipocótilos tratados se colocaron de nuevo en el 'medio de inducción de callo' con papeles de filtro, se devolvieron a la cubeta STERILITE®, se cubrieron con una tapa oscura y se devolvieron a la sala de cultivo con el mismo régimen de crecimiento que anteriormente, durante un periodo de cocultivo de 3 días. Después de 3 días, los hipocótilos se colocaron directamente sobre 'medio de inducción de callo' que contenía 1 mg/L de HERBIACE®, se colocaron de nuevo en la cubeta con una tapa transparente y se devolvieron a la sala de cultivo, manteniendo el mismo régimen de crecimiento que anteriormente. Después de 1 semana, los hipocótilos se transfirieron directamente a un 'medio de inducción de callo' con selección con 3 mg/L de HERBIACE® durante 2 semanas para el desarrollo adicional de los callos y se transfirieron a 'medio de inducción de callo' con selección a 5 mg/L de HERBIACE® durante 2-8 semanas para el desarrollo adicional de los callos. Una vez que estaba disponible una cantidad suficiente de callos, se sometió a un análisis molecular.

Regeneración de plantas a partir del callo Aged Canola: el tejido del callo Canola se colocó sobre 'medio de regene-

ración de brotes' que consistía en medio MS que contenía 30 g/L de sacarosa, 3 mg/L de benciaminopurina, 0,5 g/L de MES [ácido 2-(N-morfolio)etano sulfónico], 5 mg/L de nitrato de plata, 1 mg/L de zeatina, 250 mg/L de carbenicilina, 300 mg/L de timentina y 7 g/L de agar TC con selección a 5 mg/L de HERBIACE<sup>®</sup>, las placas se envolvieron con cinta 3M y se colocaron bajo un régimen de crecimiento establecido en 23°C y un fotoperiodo de 16 h de luz/8 h de oscuridad. Los tejidos se trasladaron después a un 'medio elongación de los brotes' que consistía en medio MS que contenía 0,5 g/L de MES, 300 mg/L de timentina, 20 g/L de sacarosa y 7 g/L de agar TC con selección a 5 mg/L de HERBIACE<sup>®</sup>. Las plántulas se transfirieron a 'medio de enraizamiento' que consistía en ½ medio MS que contenía 10 g/L de sacarosa, 0,5 mg/L de ácido indolbutírico, 300 mg/L de timentina y 8 g/L de agar TC con selección a 5 mg/L de HERBIACE<sup>®</sup>.

Una vez que se establecieron las raíces en la plántula *in vitro*, se trasplantaron a 5 macetas de 13,3 cm que contenían Metro Mix 360. Las plantas se cubrieron con un vaso transparente de la marca SOLO y se colocaron en un Conviron para la aclimatación. Después de 48-72 horas, los vasos se retiraron para permitir la circulación de aire. Después de un total de siete días, las macetas se desplazaron desde la Conviron a una nave de invernadero para un desarrollo adicional. Las plantas se cultivaron con un fotoperiodo de 16:8 horas, con temperatura diurna y nocturna entre 22-24°C. Cuando el vástago de floración principal comenzó a alargarse y formar brotes de flores, la planta entera se cubrió con una bolsa de autofecundación para evitar la polinización cruzada. Las semillas obtenidas a partir de las autopolinizaciones se recogieron aproximadamente cuatro meses después de trasplantar.

Preparación de *Agrobacterium*: *Agrobacterium* procedente de una reserva en glicerol se sembró con estrías, cuatro días antes del tratamiento, sobre un 'medio de crecimiento bacteriano semisólido' que consistía en 10 g/L de peptona, 10 g/L de extracto de levadura, 5 g/L de NaCl, 10 g/L de sacarosa más 100 mg/L de espectinomina y 250 mg/L de estreptomina y se solidificó con 15 g/L de agar Bacto y se dejó crecer durante dos días en una incubadora (incubadora Fisher Scientific Isotemp) a 28°C. Después de 2 días, un pequeño bucle de *Agrobacterium* se colocó en un matraz con deflectores desechable, estéril de 500 ml que contenía 150 ml de 'medio de crecimiento líquido bacteriano' (igual que anteriormente, menos agente solidificante), 250 mg/L de estreptomina y 100 mg/L de espectinomina, creció durante 16 horas durante una noche a 28°C en la oscuridad en un agitador cerrado (agitador con incubadora refrigerada New Brunswick Scientific Innova 4330) a 200 rpm. Después de 16 horas, el cultivo de *Agrobacterium* se retiró del agitador y se dividió en partes alícuotas en tubos de centrifuga de 50 ml (uno contenía 35 ml para la preparación y dos contenían 50 ml para la revalidación). Los tubos de centrifuga se colocaron en una centrifuga (centrifuga Beckman Modelo J2-21) y se centrifugaron a 6000 rpm durante 15 minutos y posteriormente se resuspendieron en el 'medio líquido de cultivo' hasta una densidad final de 50 Klett con un filtro rojo.

#### Ejemplo 4: Análisis de muestras de callos transformados

Muestras de callos transformados putativamente de *B. napus* se analizaron para estudiar alteraciones en la expresión del ARNm del gen endógeno de la sintetasa II de  $\beta$ -cetoacil-ACP, el gen endógeno de tubulina y el transgén de ZFP TF. Los niveles de ARNm de tubulina se utilizaron como control interno para normalizar la expresión de ZFP TF y de los ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP. Los datos de la expresión del ARNm de ZFP TF se emplearon para confirmar la presencia de al menos un transgén funcional de ZFP TF en los callos transformados.

#### 4.1 Preparación de la muestra de callo

Aproximadamente 40-50 muestras de callos de ocho semanas de edad de *B. napus* fueron obtenidos después de la transformación de tejido de hipocótilo Nex710. Se transformaron individualmente con las estructuras artificiales de ZFP TF, pDAB4695, pDAB4696, pDAB4697 y pDAB4698, como se ha descrito anteriormente en el Ejemplo 3.3. Las muestras de control se obtuvieron mediante transformación de una estructura artificial binaria de control que contenía una casete de expresión génica *pat* (AtUbi10 Promotor v2 :: *pat* v3 :: Atu ORF1 3'UTR v3). Todas las muestras se cultivaron en medio de cultivo celular Herbiace complementado, hasta su cosecha.

El ARN total se preparó a partir de tejido de callo fresco usando el kit QIAGEN RNEASY 96 (Qiagen, Valencia, CA). El ARN se trató con ADNasa exenta de ARNasa según las instrucciones del kit para eliminar cualquier contaminante de ADN genómico. La síntesis de la primera hebra se estableció según las instrucciones del fabricante con la enzima transcriptasa inversa Superscript<sup>®</sup> III (Invitrogen, Carlsbad, CA) y se cebó usando hexámeros aleatorios. Las hebras de ADNc sintetizadas se diluyeron en agua en proporciones de 1:10 y 1:50 (esto proporciona un molde suficiente para amplificar con PCR múltiples dianas). Cada parte alícuota se guardó indefinidamente a -20°C.

#### 4.2 Análisis de la expresión del ARNm de la sintetasa II de $\beta$ -cetoacil-ACP

Las mezclas de la reacción qRT-PCR se establecieron para la amplificación del ADNc de la sintetasa II de  $\beta$ -cetoacil-ACP del modo siguiente: 7,5  $\mu$ l de 2X tampón LC480 Probes Master (Roche Diagnostic, Indianapolis, IN), 0,3  $\mu$ l de cebador directo específico del gen (SEQ ID NO: 28: 5'- TTGACTCGAGCTGCTACTGC -3'; las posiciones de nucleótidos 544 - 563 de SEQ ID NO: 3 en la alineación de la Figura 8) a partir de una reserva 10  $\mu$ M, 0,3  $\mu$ l de cebador inverso específico del gen (SEQ ID NO: 29: 5'- TTTCCATATCCATCGCAACA -3'; las posiciones de nucleótidos 588 - 607 de SEQ ID NO: 3 en la alineación de la Figura 8) a partir de una reserva 10  $\mu$ M, 0,15  $\mu$ l de la sonda UPL n° 25 de LightCycler<sup>®</sup> 480 Probes Master, (Roche Diagnostic, Indianapolis, IN), 1,5  $\mu$ l de 10 % (p/v) de polivinilpirrolidona-40 (PVP-40) y 3,9  $\mu$ l de agua. La sonda UPL (Roche Diagnostics, Indianapolis, EE.UU.) es un ácido nu-

cleico bloqueado y por lo tanto tiene una  $T_m$  mayor que si se calcula de otro modo. Todos los componentes se colocaron de nuevo en el congelador antes de manipular los patrones y los desconocidos. Una microplaca de 384 pocillos se separó y se marcó, se añadieron 13,5  $\mu$ l de mezcla maestra por pocillo. Una lámina de sellado se fijó suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrifuga de microplacas de Qiagen. Se añadieron 1,5  $\mu$ l de hebras de ADNc descongeladas, sintetizadas diluidas. Además, se añadieron 1,5  $\mu$ l de patrones de número de copias de ADN plasmídico en pocillos separados en una serie de diluciones, de menor a mayor concentración, estos patrones se compararon con el ADNc de la sintetasa II de  $\beta$ -cetoacil-ACP (sintetizado a partir del ARNm total) para cuantificar el número de copias. Los patrones de número de copias de la sintetasa II de  $\beta$ -cetoacil-ACP se prepararon mediante una clonación del amplicón diana en un plásmido pCR2.1 (Invitrogen, Carlsbad, CA) y haciendo una serie de diluciones para la cuantificación del número de copias. Una lámina de sellado se fijó firmemente a la placa y se centrifugó como se ha descrito previamente. Un programa de PCR se llevó a cabo del modo siguiente: i. Activar a 95°C durante 5 minutos; ii. Desnaturalizar a 95°C durante 10 s @ 4,8°C/s; iii. Reasociar/Extender a 60°C durante 25 s @ 2,5°C/s; iv. Adquirir a 72°C durante 1 s @ 4,8°C/s; Repetir las etapas ii - iv, 40-50 veces más; v. Enfriar a 38°C durante 5 s. El ADN se amplificó con un aparato de PCR en tiempo real, LC480 (Roche, Indianapolis, IN) o uno equivalente. El tamaño del amplicón era de 64 pares de bases. Las secuencias de cebadores directos e inversos empleados en este ensayo de PCR se emparejaban perfectamente con las secuencias correspondientes presentes en dos de las diana de genes de sintetasa II de  $\beta$ -cetoacil-ACP, SEQ ID NOs: 3 y 30 (Figura 8). Por lo tanto, este ensayo representaba la expresión cuantitativa de dos dianas del gen de la sintetasa II de  $\beta$ -cetoacil-ACP.

#### 4.3 Análisis de la expresión del ARNm de tubulina

El gen de tubulina, un gen natural de *Brassica napus* (GenBank: AF258790 y GenBank: DU106489), se utilizó como patrón de referencia para normalizar con precisión la señal de la expresión del ARNm a través de los genes en los ensayos de qRT-PCR. El ADNc sintetizado para el ensayo qRT-PCR de la sintetasa II de  $\beta$ -cetoacil-ACP (descrito en el Ejemplo 4.1), también se utilizó en el ensayo de qRT-PCR para tubulina como se describe a continuación.

Una qRT-PCR se preparó con 0,3  $\mu$ l de un cebador directo específico del gen (SEQ ID NO: 31: 5'- ACAGCGATT-GCCTACAAGG -3') (reserva 10  $\mu$ M) y 0,3  $\mu$ l de un cebador inverso específico del gen (SEQ ID NO: 32: 5'- AGATGGTTAAGATCACCAAAGG -3') (reserva 10  $\mu$ M), 1,5  $\mu$ l de 10% (p/v) de PVP-40, 3,9  $\mu$ l de agua y 7,5  $\mu$ l de 2X la mezcla maestra LightCycler<sup>®</sup> 480 SYBR Green I Master Mix (Roche, Indianapolis, IN) para detectar y cuantificar el ADN. Una microplaca de 384 pocillos se separó y se marcó, se añadieron 13,5  $\mu$ l de mezcla maestra por pocillo. Una lámina de sellado se fijó suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrifuga de microplacas de Qiagen. La lámina de sellado se retiró y se añadieron 1,5  $\mu$ l de hebras de ADNc descongeladas, sintetizadas diluidas. Además, se añadieron 1,5  $\mu$ l de patrones de número de copias de ADN plasmídico en pocillos separados en una serie de diluciones, de menor a mayor concentración, estos patrones se compararon con el ADNc de la tubulina (sintetizado a partir del ARNm total) para cuantificar el número de copias. Las series de patrones de número de copias de ADN de la tubulina se prepararon mediante una clonación del amplicón diana en un plásmido pCR2.1 (Invitrogen, Carlsbad, CA) y haciendo una serie de diluciones para la cuantificación del número de copias. Una lámina de sellado se fijó firmemente a la placa y se centrifugó como se ha descrito previamente. Un programa de PCR se realizó del modo siguiente: i. Activar a 95°C durante 10 minutos; ii. Desnaturalizar a 95°C durante 10 s @ 4,8°C/s; iii. Reasociar/Extender a 55,5°C durante 20 s @ 2,5°C/s; iv. Adquirir a 72°C durante 20 s @ 4,8°C/s; repetir las etapas ii - iv, 39 veces más; vi. Enfriar a 38°C durante 5 s. El ADN se amplificó con un aparato de PCR en tiempo real, LC480 o uno equivalente. Un amplicón de 307 pares de bases se amplificó en esta reacción. Este cebador inverso incluye un intrón de 78 pb basándose en la secuencia de GenBank. Por lo tanto, los amplicones amplificados a partir del ADN genómico no estaban favorecidos y los contaminantes de ADN genómico tendrían un mayor peso molecular y se pueden diferenciar fácilmente del ADNc mediante la migración de los amplicones sobre un gel de agarosa.

#### 4.4 Análisis de la expresión del ARNm de ZFP TF

La expresión del ARNm de ZFP TF se cuantificó a partir de muestras de ADNc que se sintetizaron originalmente para el ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP (descrito anteriormente en el Ejemplo 4.1). Un ensayo de PCR TaqMan fue diseñado a partir dla casete de ZFP mediante el anclaje de cebadores a la secuencia *opaque-2* NLS en el extremo 5' y la secuencia de VP16 en el extremo 3'. El ensayo de PCR se realizó usando: 0,5  $\mu$ l de cebador directo "OP2\_NLS\_F1" (SEQ ID NO: 33: 5'- AAGGAAGAGGAAGGAGTCTAACAG -3') a partir de una solución de reserva 10  $\mu$ M, 0,5  $\mu$ l de cebador inverso "VP16\_R1" (SEQ ID NO: 34: 5'- CTTCTGCTCTCCACCGTA -3') a partir de una solución de reserva 10  $\mu$ M, 0,25  $\mu$ l de la sonda "VP16\_MGB\_185" (SEQ ID NO: 45: 5'- TTGATGGTGAAGATGT -3') a partir de una solución de reserva 5  $\mu$ M, 1,5  $\mu$ l de 10% (p/v) de PVP-40, 1,5  $\mu$ l de tampón 10x Hot Start PCR, 1,0  $\mu$ l de MgCl<sub>2</sub> 25 mM, 1,2  $\mu$ l de dNTP (cada uno 2,5 mM), 0,15  $\mu$ l de polimerasa Hot Start Taq (Qiagen, Valencia, CA) y 6,9  $\mu$ l de agua. La mezcla se amplificó usando LightCycler<sup>®</sup> 480 (Roche Diagnostics, USA). Una microplaca de 384 pocillos se separó y se marcó, se añadieron 13,5  $\mu$ l de mezcla maestra por pocillo. Una lámina de sellado se fijó suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrifuga de microplacas de Qiagen. Se añadieron 1,5  $\mu$ l de hebras de ADNc descongeladas, sintetizadas diluidas. Una lámina de sellado se fijó firmemente a la microplaca y se centrifugó como se ha descrito previamente. Un programa de PCR se llevó a cabo del modo siguiente: i. Activar a 95°C durante 15 minutos; ii. Desnaturalizar a 95°C durante 20 s @ 4,8°C/s; iii. Reasociar/Extender a 60°C durante 20 s @ 2,5°C/s; iv. Adquirir a 72°C durante 55 s @ 4,8°C/s; repetir las etapas ii -

iv 44 veces más; vi. Enfriar a 38°C durante 5 s. El ADN se amplificó con un aparato de PCR en tiempo real, LC480 o uno equivalente. El tamaño del amplicón era de 775 pares de bases. El número de copias del ADNc/ARNm de ZFP TF se determinó utilizando una serie estándar en la que el amplicón diana se clonó en un plásmido y se realizó una serie de diluciones para los ensayos de PCR como se ha descrito para sintetasa II de  $\beta$ -cetoacil-ACP y la tubulina anteriormente en los ejemplares.

#### 4.5 Análisis de la expresión de muestras de callos

La expresión del ARNm de ZFP TF se midió en muestras de callos transgénicos de 8 semanas, transformadas putativamente con ZFP TFs y que habían crecido sobre medio de selección HERBIACE® (véase el Ejemplo 3). La expresión del ARNm cuantitativo de ZFP TF, tal como se midió con el ensayo de qRT-PCR, se detectó en las muestras transformadas solo con estructuras artificiales que contenían ZFP TF, pero no en las muestras que contenían la estructura artificial de control. Estos datos indicaron que el ensayo era específico de la expresión de ZFP TF y que los callos de control contenían el transgén de ZFP TF.

Se calcularon relaciones de la expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP y de tubulina, basándose en los resultados de la qRT-PCR, para discernir las diferencias de la expresión entre las muestras de callos. La mayor regulación al alza del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP se observó en muestras de callos de canola transformados con la estructura artificial pDAB4695 (Tabla 4). Este 27% de regulación al alza era estadísticamente significativo ( $p = 0,05$ ) sobre el de las muestras de control.

Tabla 4: Expresión del ARNm de KASII en muestras de callos de *B. napus* transformados con diferentes estructuras artificiales de ZFP TF

Estructura artificial	Número de callos analizados	Expresión media de KASII/tubulina	DS/DE de la expresión de KASII/tubulina
4695	23	3,69	1,20
4696	20	2,81	0,84
4697	24	3,18	0,93
4698	24	3,03	0,68
Control	14	2,90	1,12

#### Ejemplo 5: Análisis de muestras de plantas T0 transformadas

##### 5.1 Análisis de la expresión del ARNm de sintetasa II de $\beta$ -cetoacil-ACP, tubulina y ZFP TF

Plantas T0 de canola que contenían la estructura artificial transgénica pDAB4695 se cultivaron hasta la etapa de planta con 6 hojas y se analizaron para estudiar la expresión del ARNm del gen endógeno de la sintetasa II de  $\beta$ -cetoacil-ACP, el gen endógeno de tubulina y el transgén de ZFP TF. Las plantas transgénicas se compararon con muestras de control que se habían transformado con una estructura artificial binaria que contenía una casete de expresión génica *pat* (AtUbi10 Promotor v2 :: *patv3* :: AtU ORF1 3'UTR v3). Seis perforaciones de las hojas, realizadas con una perforadora de papel de tamaño estándar, se tomaron sobre hielo a partir de cada planta y el ARNm total se extrajo con el kit de extracción de ARN de Qiagen para 96 pocillos RNeasy, de acuerdo con las instrucciones del fabricante. (Qiagen, Carlsbad, CA). Los análisis de la expresión del ARNm de sintetasa II de  $\beta$ -cetoacil-ACP, tubulina y ZFP TF se realizaron utilizando los protocolos descritos anteriormente en el Ejemplo 4.

La expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP variaban entre plantas T0 transgénicas independientes. Una regulación al alza del ARNm mayor de 3 veces se observó entre los 16 eventos analizados (Figura 3). Todas las plantas eran positivas para la expresión de ZFP TF. Además, 27 plantas de control se utilizaron para calcular la expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP. La expresión media del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP desde el control de la línea de base, se comparó con la expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP de eventos de plantas ZFP TF individuales, para calcular el factor de aumento de la expresión (Figura 3). Estas plantas se cultivaron hasta la madurez para obtener semillas T1. Antes de la floración, todas las plantas estaban cubiertas con bolsas de autofecundación individuales para facilitar la autopolinización de las flores dentro de una planta. Las semillas T1 de esas plantas se recogieron aproximadamente 4 meses después de su transferencia a un invernadero

Se seleccionaron los eventos 3, 6 y 12.2 (en lo sucesivo denominado el evento 12), que representaban diferentes

intervalos de expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP, para un estudio adicional.

Ejemplo 6: Análisis de muestras de plantas T1 transformadas

6.1 Análisis de ácidos grasos de semillas T1

5 El análisis de ácidos grasos de una semilla individual se realizó en 24 semillas T1 individuales, por evento, para estudiar los efectos de ZFP TF sobre la alteración en el contenido de ácidos grasos (Tabla 5). Un método de análisis de éster metílico de ácidos grasos (del inglés, fatty acid methyl ester, FAME) basado en el método AOCS Ce2 - 66(97), se empleó y todos los números en la Tabla 5 se muestran como un porcentaje del total de ácidos grasos presentes en las semillas de canola.

Tabla 5: Perfil de ácidos grasos de semillas T1 individuales medidas con el análisis FAME

Eventos	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	Total C18
3 bajo	2,8	0,3	1,6	76,1	12,2	3,7	93,6
3 bajo	2,9	0,3	1,4	76,7	11,8	4,0	93,9
3 bajo	2,9	0,3	1,4	75,7	13,0	3,8	93,9
3 alto	3,6	0,3	1,4	75,1	13,0	3,5	93,1
3 alto	3,8	0,4	1,4	75,6	12,0	3,8	92,8
3 alto	4,0	0,4	1,5	74,8	12,6	3,7	92,5
Media (n = 24)	3,3	0,3	1,5	75,7	12,4	3,8	93,3
DS/DE (n = 24)	0,5	0,1	0,1	0,7	0,5	0,1	0,34
12 bajo	3,1	0,2	1,5	76,4	12,0	3,8	93,6
12 bajo	3,2	0,3	1,4	77,0	11,5	3,4	93,3
12 bajo	3,2	0,2	1,8	80,9	7,5	3,2	93,5
12 alto	3,6	0,3	1,5	76,8	11,7	3,1	93,2
12 alto	3,6	0,3	1,6	76,0	11,9	3,5	93,0
12 alto	3,6	0,3	1,8	77,3	10,8	3,0	92,9
Media (n = 24)	3,4	0,3	1,6	77,4	10,9	3,3	93,2
DS/DE (n = 24)	0,2	0,0	0,2	1,8	1,7	0,3	0,23
6 bajo	3,2	0,5	1,9	81,3	8,1	2,1	93,3
6 bajo	3,3	0,5	1,2	79,6	10,1	2,3	93,1
6 bajo	3,5	0,5	2,6	80,6	7,5	1,9	92,6
6 alto	4,4	0,5	2,4	77,7	10,7	1,5	92,2
6 alto	4,5	0,6	2,2	77,1	10,5	1,8	91,6

Eventos	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	Total C18
6 alto	4,6	0,7	1,3	77,4	11,0	1,7	91,4
Media (n = 24)	3,9	0,5	1,9	78,9	9,6	1,9	92,4
DS/DE (n = 24)	0,6	0,1	0,6	1,8	1,5	0,3	0,54
Media de Nex710 (n = 216)	3,8	0,4	2,3	78,1	9,5	2,5	92,5
DS/DE (n = 216)	0,2	0,0	0,6	1,8	1,6	0,7	0,45

5 Para la preparación de las muestras, semillas individuales aisladas se colocaron en tubos de racimo marcados sobre una placa de extracción de 96 pocillos que contenía una bola de acero de 0,3 cm (Small Parts, Miramar, FL). Los tubos se taparon y la semilla se molió en seco en un GenoGrinder (SPEX CertiPrep Group, Metuchen, NJ) durante 3,0 minutos a 1300 perforaciones/minuto. Las tapas se retiraron y se añadieron 0,6 ml de heptano a cada pocillo. Los pocillos se volvieron a tapar y se colocaron de nuevo en el GenoGrinder para una molienda adicional durante 2,0 minutos a 1200 perforaciones/min. Las muestras se retiraron a continuación y se centrifugaron a 3700 rpm durante 10,0 minutos a 6°C. Empleando un robot Beckman Coulter MC, el material sobrenadante se transfirió a una placa de 96 pocillos con insertos de vidrio (MicroLiter, Suwanee, GA). Se añadieron 40 µL de metóxido de sodio al 1% a la muestra. El metóxido de sodio, se diluyó a partir de una solución de reserva con 30% de metanol (Fluka/Sigma Aldrich, St. Louis, MO). Las placas se taparon con una estera recubierta de teflón y se dejaron incubar a temperatura ambiente durante 4 horas antes del análisis GC.

15 En las muestras se analizó el contenido en ácidos grasos en un Agilent 6890 GC-FID (Agilent Technologies, Santa Clara, CA) equipado con una columna J&W Scientific DB-23, ID de 15 metros x 0,25 mm y 0,25 µm de espesor de la película, (J&W Scientific, Folsom, CA). La temperatura inicial del horno era de 200°C y esta temperatura se mantuvo durante la duración de la ejecución. La entrada se fijó con una relación de separación de 1:100 y una temperatura de 280°C. Un caudal en rampa de 0,8 ml/min de helio se mantuvo durante los dos minutos iniciales. El caudal se incrementó después a una tasa de 1,0 ml/min a 2,5 ml/min y se mantuvo durante 1,5 minutos. El detector se fijó a 300°C con una reposición de gas portador constante y un flujo de la columna de 30 ml/min, flujo de hidrógeno combustible de 30 ml/min y flujo de oxidante de 400 ml/min. Un volumen de inyección de 2 µL se utilizó para todas las muestras.

20 Los picos de ésteres metílicos de ácidos grasos individuales se identificaron mediante una comparación con los tiempos de retención de patrones de referencia de éster metílico (GLC#428, Nu-Chek-Prep, Inc., Elysian, MN) utilizando Waters Corp. Empower Software. El porcentaje de áreas individuales se calculó para todos los analitos en el patrón de referencia, basado en el total de las áreas de los picos de cromatografía integrados. También se disparó un heptano vacío para identificar cualquier contaminación en la GC.

25 Una comparación de las semillas T1 con los tres niveles más bajos de C16:0 (que son lo más probable plantas positivas para ZFP TF) y los tres niveles más altos de C16:0 (que son lo más probable plantas nulas para ZFP TF) indicaba que los cambios en el contenido en C16:0 podría ser debido a la segregación del transgén ZFP TF (Tabla 5). Un cambio correspondiente en C18 total (C18:0 + C18:1 + C18:2 + C18:3) también se observó en todos los eventos; las semillas con niveles bajos de C16:0 tenían niveles totales de C18 incrementados y viceversa. Esto era a pesar de la variabilidad en los contenidos en ácidos grasos individuales de C18:0, C18:1, C18:2 y C18:3 debido a la segregación de los genes mutantes *fad2* y *fad3* (Hu, X., et al., Theor. Appl. Genet. 2006, 113: 497-507) en el genotipo Nex710 que se había transformado. Por lo tanto, se llevó a cabo el siguiente paso para identificar las plantas positivas para ZFP TF y hermanas nulas en una población segregante T1 de cada evento, para detectar el cambio real en el perfil de ácidos grasos en fondos genéticos similares.

#### 6.2 Análisis de la presencia de ZFP TF en plantas T1

30 Se escrutaron de 100-150 plántulas T1 de los tres eventos transgénicos 3, 6 y 12, para identificar plantas positivas para ZFP TF y hermanas nulas (Tabla 6). Las plantas se sometieron a ensayo para estudiar la presencia de al menos una casete de longitud completa del transgén ZFP TF. El ADN total genómico de las hojas de esas plantas se aisló según las recomendaciones del fabricante con el kit de extracción QIAGEN PLANT DNEasy (Qiagen, Valencia, CA). El protocolo se modificó mediante la adición de PVP-40 al tampón Qiagen AP1 hasta una concentración final de 1%. El ADNg purificado se cuantificó mediante un protocolo de cuantificación de ADN de PicoGreen® (Molecular Probes, Inc., Eugene, OR). Las muestras de ADN se analizaron en busca de la presencia de, como mínimo, una copia de longitud completa del transgén ZFP TF en un ensayo de PCR que contenía los siguientes reactivos: 5 µL de 10X tampón EX Taq (Takara Bio Inc., Otsu, Siga, Japón), 5 µL de 10% de polivinilpirrolidona-40, 1,0 µL de cebador directo ubi10 (SEQ ID NO: 35: 5'-GGTCAACGGATCAGGATATTCTTG -3') a partir de una solución de reserva 10

5  $\mu\text{M}$ , 1,0  $\mu\text{L}$  de cebador inverso *AtuORF23* (SEQ ID NO: 36: 5'-CCATGTTGGCAAAGGCAACC -3'), solución de reserva 10  $\mu\text{M}$ , 4  $\mu\text{L}$  de dNTPs (2,5 mM cada uno), 0,2  $\mu\text{L}$  de TaKaRa EX Taq<sup>®</sup> Hot Start, 31,8  $\mu\text{L}$  de agua y 2  $\mu\text{L}$  de ADN a una concentración de 10 ng/ $\mu\text{L}$ . Las condiciones de la ciclación de la PCR fueron las siguientes: 94°C durante 2 min, 10 ciclos de PCR "touch-down" a 98°C durante 10 s, 65°C durante 20 s con una disminución de temperatura de 0,5°C en cada ciclo hasta 60°C, seguido de 72°C durante 3 min. Esto fue seguido por 35 ciclos a 98°C durante 10 s, 60°C durante 20 s y 72°C durante 3 min y una extensión final a 72°C durante 10 min. Migraron 5  $\mu\text{L}$  de cada reacción en un gel de agarosa al 1% para detectar la presencia de la banda de tamaño esperado de 2662 pb de ZFP TF.

10 Un ensayo de PCR cuantitativa en tiempo real para el gen *pat*, que está ligado molecularmente a ZFP TF en la estructura artificial pDAB4695, se desarrolló y se aplicó para identificar plantas positivas para *pat* y nulas para confirmar su segregación genética con la casete de ZFP TF al que está ligado molecularmente.

Tabla 6: Identificación de plantas positivas para ZFP TF y hermanas nulas en una población de segregación T1 de los tres eventos

Eventos	Semillas T1 totales plantadas	Plantas T1 ZFP TF/nulas identificadas	Plantas T1 ZFP TF/nulas avanzadas a semilla T2
3	150	115/7	12/7
6	100	66/27	5/5
12	150	126/6	8/6

15 La mezcla de reacción para el ensayo TaqMan RT-PCR de *pat* estaba compuesta por lo siguiente: 3  $\mu\text{L}$  de 2X mezcla maestra de sondas Roche LIGHTCYCLER<sup>®</sup> Cycloer 480 II, 0,6  $\mu\text{L}$  de 10% de PVP-40, 0,2  $\mu\text{L}$  cada una de cebador directo *pat* (SEQ ID NO: 37: 5'-ACAAGAGTGGATTGATGATCTAGAGAGGT -3') de una solución de reserva 10  $\mu\text{M}$ , 0,2  $\mu\text{L}$  de cebador inverso PAT (SEQ ID NO: 38: 5'-CTTTGATGCCTATGTGACACGTAACAGT -3') de una solución de reserva 10  $\mu\text{M}$ , 0,2  $\mu\text{L}$  de la sonda (SEQ ID NO: 39: 5'-6FAM-CCAGCGTAAGCAATACCAGCCACAACACC-extintor-3') de una solución de reserva 5  $\mu\text{M}$ , y se multiplexó con los siguientes reactivos para un patrón de referencia interno de HMG (GenBank: AF127919), del modo siguiente: 0,2  $\mu\text{L}$  cada una de cebador directo de HMG (SEQ ID NO: 40: 5'-CCTCTCTACCACCGTCTCACATG -3') de una solución de reserva 10  $\mu\text{M}$ , 0,2  $\mu\text{L}$  de cebador inverso de HMG (SEQ ID NO: 41: 5'-GATCTGGCCGGACTGTTTCA -3') de una solución de reserva 10  $\mu\text{M}$ , 0,2  $\mu\text{L}$  de la sonda (SEQ ID NO: 42: 5'-6FAM-CGCTCCTCAGCTACCACCTCAACCA-extintor-3'), de una solución de reserva 5  $\mu\text{M}$ , 1  $\mu\text{L}$  de ADN y 0,2  $\mu\text{L}$  de agua. Las condiciones de ciclación de la PCR eran: 1 ciclo a 95°C durante 5 min, 40 ciclos a 95°C durante 10 s, 60°C durante 40 s y 72°C durante 1 s. Todas las muestras se amplificaron por triplicado en placas de 384 pocillos junto con 1-4 copias de patrones de ADN genómico transgénico en un aparato de PCR Roche LIGHTCYCLER<sup>®</sup> 480.

30 Basándose en las relaciones de segregación de las plantas positivas para ZFP TF y nulas dentro de cada evento, los eventos 3, 6 y 12 contenían aproximadamente 2, 1 y 2 inserciones respectivamente del transgén ZFP TF en el genoma de *B. napus* (Tabla 6, columna 3). Estos datos coincidían con los datos de segregación del gen *pat* en los tres eventos.

### 6.3 Regulación al alza del ARNm de la sintetasa II de $\beta$ -cetoacil-ACP en plantas T1

35 Todas las plantas T1 positivas para ZFP TF y hermanas nulas que se muestran en la Tabla 6, columna 3 fueron sometidas a un análisis de la expresión del ARNm para el gen endógeno de la sintetasa II de  $\beta$ -cetoacil-ACP, el transgén de ZFP TF y el gen endógeno de tubulina. El último sirvió como gen de referencia para normalizar las expresiones de los genes de la sintetasa II de  $\beta$ -cetoacil-ACP y ZFP TF. Seis perforaciones de hojas de cada una de las etapas de planta de 6 hojas se tomaron para muestras sobre hielo y el ARN total se extrajo usando el kit de Qiagen RNAEASY<sup>®</sup>. La síntesis de la hebra de ADNc y las diluciones se completaron como se ha descrito en el Ejemplo 4.1.

40 El análisis con qRT-PCR del ADNc sintetizado se completó para los tres genes. Se realizaron análisis de la expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP y de tubulina, como se describe en los Ejemplos 4.2 y el Ejemplo 4.3, respectivamente. El ensayo de ZFP TF se modificó. La reacción de PCR se estableció del modo siguiente: Las condiciones de reacción eran las siguientes: 7,5  $\mu\text{L}$  de 2X tampón maestro LC480 Probes Master (Roche Diagnostic, Indianapolis, IN), 0,3  $\mu\text{L}$  de cebador 1 directo específico del gen (SEQ ID NO: 43: 5'-TCGATCTTGATATGTTGGGAGA -3') (de una solución de reserva 10  $\mu\text{M}$ ), 0,3  $\mu\text{L}$  de cebador 2 inverso específico del gen (SEQ ID NO: 44: 5'-AGGTGCAGAATCATGTGGTG -3') (de una solución de reserva 10  $\mu\text{M}$ ), 0,15  $\mu\text{L}$  de sonda UPL n° 85, 1,5  $\mu\text{L}$  de 10% (p/v) de PVP-40 y 3,9  $\mu\text{L}$  de agua. Una microplaca de 384 pocillos se separó y se añadieron 13,5  $\mu\text{L}$  de la mezcla descrita anteriormente, por pocillo. Una lámina de sellado se fijó a la placa y la placa se centrifugó durante 1 minuto a 3000 rpm. La película se retiró y se añadieron 1,5  $\mu\text{L}$  por pocillo de las primeras hebras descongeladas, diluidas.

Además, se añadieron los patrones de ADNc a los pocillos de control separados. La lámina de sellado se selló sobre la placa y se repitió la etapa de centrifugación. Un programa de PCR se llevó a cabo con las siguientes condiciones; i. Activar a 95°C durante 5 minutos; ii. Desnaturalizar a 95°C durante 10 s (@ 4,8°C/s); iii. Reasociar/Extender a 60°C durante 25 s (@ 2,5°C/s); iv. Adquirir a 72°C durante 1 s (@ 4,8°C/s); v. Repetir las etapas ii - iv, 39-49 veces; vi. Enfriar a 38°C durante 5 s.

La expresión del ARNm de ZFP TF era la más baja en el evento 6 y la más alta en el evento 3 de las muestras positivas para ZFP TF (Figura 4). Las plantas nulas no expresaban el ARNm de ZFP TF, lo que indicaba que el ensayo de ZFP TF era específico de la presencia del transgén de ZFP TF. Los niveles de ARNm del gen de la sintetasa II de  $\beta$ -cetoacil-ACP, la diana para ZFP TF, se regulaban al alza en plantas positivas para ZFP TF de todos los eventos (Figura 5). Se observó hasta un aumento de 3-4 veces en la expresión del ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP mediante comparaciones por parejas de las plantas positivas para ZFP TF y las hermanas nulas correspondientes, dentro de cada evento. La regulación al alza de los niveles de ARNm de la sintetasa II de  $\beta$ -cetoacil-ACP era proporcional al aumento de los niveles de expresión del ARNm de ZFP TF dentro de cada evento (Figura 4), es decir, la más baja en el evento 6 y la más alta en el evento 3.

#### 6.4 Análisis de ácidos grasos en hojas de plantas T1

Un segundo conjunto de muestras de hojas T1 se recogió para el análisis de ácidos grasos junto con la recogida de muestras de análisis del ARNm descritas en el Ejemplo 6.3. Sin embargo, solamente se analizó un subconjunto de las muestras en relación con los ácidos grasos, solo para las plantas que habían avanzado hasta la madurez como se muestra en la Tabla 6, columna 4. El análisis FAME de material de las hojas se completó del modo siguiente. Se recogió material de hojas de 10-100 mg sobre hielo y luego se secó por congelación antes de una reacción de transesterificación con metóxido de sodio 0,25 M en metanol anhidro a 40°C durante 20 minutos. Después de la reacción de transesterificación, los ácidos grasos se extrajeron tres veces con una solución de heptano que contenía heptadecanoína como patrón sustituto. Las fracciones de hexano aisladas se secaron y se resuspendieron en un volumen conocido de heptano. El éster metílico de ácidos grasos resultante (FAME) se analizó mediante GC-FID utilizando una columna capilar BPX 70 de SGE (15 m x 0,25 mm x 0,25  $\mu$ m). Cada FAME fue identificado por su tiempo de retención y se cuantificó mediante la inyección de una mezcla de referencia de aceite de colza de Matreya LLC como patrón de calibración. La finalización de la reacción se verificó comprobando la presencia de FAMES endógenos en una cuarta extracción/derivación.

Los tres eventos mostraron una disminución del contenido en C16:0 total y un aumento concomitante del contenido en C18 total, en comparaciones por parejas de las plantas positivas para ZFP TF y las hermanas nulas correspondientes, dentro de cada evento (Tabla 7). Por ejemplo, las plantas positivas para ZFP TF del evento 12 mostraron una reducción del contenido en C16:0 en un 6,3% y un aumento del contenido total de C18 del 1,44%, en comparación con sus propias plantas hermanas nulas.

Tabla 7: Perfil de ácidos grasos de hojas T1

Nombre de la muestra	Plantas analizadas	C16:0		C18 total	
		Media	DS/DE	Media	DS/DE
12null	6	10,98	0,31	69,93	0,91
12ZFP	8	10,29	0,39	70,94	1,02
3null	7	10,79	0,56	69,84	1,18
3ZFP	12	10,34	0,29	71,25	0,74
6null	5	11,94	0,47	70,36	1,91
6ZFP	5	10,44	0,62	72,16	2,88

#### Ejemplo 7: Análisis de semillas transgénicas T2 de ZFP TF

Un subconjunto de plantas positivas para ZFP TF que mostraban las expresiones de ARNm más altas para la sintetasa II de  $\beta$ -cetoacil-ACP y plantas hermanas nulas dentro de cada evento, avanzaron hasta la madurez para obtener semillas T2 (Tabla 6, columna 4). La excepción era el evento 3, en donde unas pocas plantas positivas para ZFP TF con intervalos de expresión bajos, también habían avanzado. Estas podrían representar la otra inserción de se-

gregación del transgén pDAB4695. Cada planta T1 se cubrió con una bolsa de malla para facilitar la autopolinización dentro de una planta dada. Las semillas se recogieron y se sometieron a ensayo los niveles de ácido graso.

Los análisis de ácidos grasos en T2 se realizaron en un grupo de 24 semillas por planta, como se ha descrito para el análisis de semillas T1 en el Ejemplo 6.1. Los resultados mostraron una disminución del contenido en C16:0 en plantas positivas para ZFP TF de los tres eventos mediante una comparación por parejas con sus hermanas nulas correspondientes (Tablas 8A y 8B). Esta disminución del contenido en C16:0 era del 12,6%, 11,2% y 22,3% para los eventos 12, 3 y 6, respectivamente. Una disminución consistente también se observó en el contenido de C16:1. En correspondencia con la disminución del contenido en C16:0 y C16:1, un aumento concomitante en el contenido total en C18 (C18:0 + C18:1 + C18:2 + C18:3) se observó en plantas positivas para ZFP TF en una comparación por parejas con las plantas hermanas nulas correspondientes. Las tres comparaciones por parejas de las plantas positivas para ZFP TF y las hermanas nulas correspondientes, eran estadísticamente significativas ( $p = <0,001$ ) para los aumentos del contenido total en C16 (C16:0 + C16:1) y el contenido total en C18, en el análisis con el programa informático JMP statistical (Figura 6A y B). Aunque los tres eventos eran variables en su contenido en C16:0 basado en observaciones de plantas hermanas nulas (Tablas 8A y 8B), ZFP TF era eficaz en la disminución del contenido en C16:0 a niveles aproximadamente similares; entre ~ 2,93 a 3,05. Mientras que todas las plantas positivas para ZFP TF de los eventos 3 y 12 (Tablas 8 y 9) tenían semillas, solo una planta que contenía ZFP TF del evento 6, tenía semillas. Esta planta particular del evento 6 tenía el contenido más alto en C16:0 basándose en los datos de ácidos grasos de hojas T1. Las Tablas 8A y 8B muestran una comparación del perfil de ácidos grasos en T2 de tres eventos transgénicos de *B. napus* de plantas positivas para ZFP TF pDAB4695 y las hermanas nulas correspondientes. Estos datos se obtuvieron con el análisis FAME y se describen en la Sección 6.1. Los ácidos grasos de C12:0 a C24:1 se analizaron y los más importantes se muestran en estas tablas. Todos los números se representaron como un porcentaje de ácidos grasos totales presentes en semillas de *B. napus*.

Tabla 8A: Análisis de ácidos grasos C16:0, C16:1, C18:0, C18:1, C18:2 y C18:3

Muestras	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
Media de 12Null	3,49	0,26	1,29	76,59	11,45	3,73
DS/DE (N = 6)	0,06	0,01	0,05	0,25	0,21	0,04
Media de 12ZFP	3,05	0,20	1,33	76,53	11,63	3,87
DS/DE (n = 8)	0,11	0,02	0,04	0,39	0,33	0,06
Media de 3Null	3,30	0,25	1,31	74,34	13,18	3,94
DS/DE (N = 7)	0,17	0,02	0,10	1,04	0,86	0,25
Media de 3ZFP	2,93	0,18	1,30	74,10	13,62	4,14
DS/DE (n = 12)	0,12	0,01	0,08	1,80	1,45	0,25
Media de 6Null	3,77	0,38	2,02	73,46	12,42	3,66
DS/DE (N = 4)	0,15	0,08	0,14	2,80	2,44	0,66
6ZFP (n = 1)	2,93	0,17	1,07	77,89	10,65	3,24

ES 2 657 067 T3

Tabla 8B: Análisis de ácidos grasos C20:0, C20:1, C20:2, C22:0, C22:1, C24:0 y C24:1

Muestras	C20:0	C20:1	C20:2	C22:0	C22:1	C24:0	C24:1
Media de 12Null	0,53	1,26	0,05	0,33	0,02	0,14	0,14
DS/DE (N = 6)	0,03	0,07	0,01	0,04	0,01	0,02	0,02
Media de 12ZFP	0,55	1,37	0,06	0,36	0,02	0,15	0,16
DS/DE (n = 8)	0,02	0,05	0,01	0,02	0,01	0,01	0,02
Media de 3Null	0,59	1,49	0,06	0,43	0,02	0,15	0,21
DS/DE (N = 7)	0,07	0,11	0,01	0,07	0,01	0,03	0,04
Media de 3 ZFP	0,57	1,57	0,07	0,42	0,03	0,15	0,19
DS/DE (n = 12)	0,04	0,21	0,02	0,07	0,01	0,04	0,04
Media de 6Null	0,80	1,49	0,07	0,54	0,03	0,30	0,24
DS/DE (N = 4)	0,08	0,13	0,01	0,08	0,01	0,07	0,02
6ZFP (n = 1)	0,52	1,86	0,07	0,42	0,04	0,20	0,22

La Tabla 9 muestra los perfiles de ácidos grasos totales basándose en los ácidos grasos individuales mostrados en las Tablas 8A y 8B.

5 Tabla 9

Muestras	Total C18*	Total LC**	Total Sats***
Media de 12Null	93,06	95,52	5,83
DS/DE (n = 6)	0,13	0,05	0,09
Media de 12ZFP	93,36	96,01	5,48
DS/DE (n = 8)	0,11	0,13	0,09
Media de 3Null	92,78	95,74	5,82
DS/DE (n = 7)	0,39	0,17	0,32
Media de 3ZFP	93,16	96,14	5,41
DS/DE (n = 12)	0,41	0,13	0,19

Muestras	Total C18*	Total LC**	Total Sats***
Media de 6Null	91,55	95,01	7,47
DS/DE (n = 4)	0,36	0,29	0,40
6ZFP (n = 1)	92,85	96,18	5,17

\*Total C18 representa una suma de los carbonos siguientes: C18:0, C18:1, C18:2 y C18:3.

\*\*Total LC o cadena larga representa una suma de los siguientes carbonos: C18 total + C20 total (C20:0 + C20:1 + C20:2) + C22 total (C22:0 + C22:1) + C24 total (C24:0 + C24:1)

\*\*\*Total Sats representa todos los ácidos grasos saturados que se muestran en las Tablas 8A y 8B.

5 Combinando todos los ácidos grasos de cadena larga (C18 total, C20 total, C22 total y C24 total), se observó un aumento del 0,4% a 1,2% en plantas positivas para ZFP TF de diferentes eventos, en comparación con hermanas nulas. Entre los ácidos grasos de cadena más larga, un aumento de C20:1 en un 25% fue notable en plantas positivas para ZFP TF para el evento 6. Una disminución en el contenido total de ácidos grasos saturados también se observó en plantas que contenían ZFP TF en todos los eventos. Había un intervalo de reducción del 6% en el evento 12 hasta el 31% de reducción en el evento 6.

10 Resumiendo, se ejemplificó la regulación al alza de la transcripción mediada por ZFP TF del gen de la sintetasa II de  $\beta$ -cetoacil-ACP en plantas T0 y T1 y la disminución concomitante del contenido en ácidos grasos C16 total y el aumento en C18 en las semillas T2. Estos datos demuestran la aplicación con éxito de tecnologías novedosas con ZFP TF para dirigirse a y modificar genes específicos en la vía de biosíntesis de ácidos grasos, dando como resultado cambios previstos en los perfiles del aceite de las semillas, que son heredables a través de generaciones de la prole.

Ejemplos 8-15: Regulación al alza/regulación a la baja de FatB

15 Los plástidos son el principal sitio para una biosíntesis *de novo* de ácidos grasos en las plantas superiores (Ohlrogge et al., 1979, Proc Natl Acad Sci USA 76:1194-8; Thelen y Ohlrogge, 2002, Metabolic Engineering 4:12-21). Los ácidos grasos no utilizados en los plástidos se exportan al citoplasma a través de la actividad específica de tioesterasas de acil-ACP (FATs) por hidrólisis de proteínas portadoras de acil-acilo (acil-ACP) para liberar ácidos grasos libres. Hay dos clases de enzimas FAT en las plantas, FATA y FATB. La clase FATA prefiere 18:1-ACP *in vitro* mientras que la clase FATB prefiere sustratos acil-ACP saturados, tales como 16:0-ACP y 18:0-ACP, pero también muestra otra especificidad de sustrato heterogénea (Doermann et al., 1995, Arch. Biochem. Biophys. 316: 612-618; Voelker et al., 1997, Plant Physiology 114: 669-677; Salas y Ohlrogge, 2002, Archives of Biochemistry and Biophysics, 403:25-34). De forma similar a FATA, FATB también se considera que está presente en todos los tejidos vegetales, pero se expresa predominantemente en semillas en desarrollo (Jha et al., 2006, Plant Physiology and Biochemistry 44:645-655).

30 El genoma nuclear de *Arabidopsis* codifica dos genes *FatA* y un solo gen *FatB*. La inactivación de la actividad *FatB* puede disminuir drásticamente el contenido en ácidos grasos saturados en glicerolípidos. Por ejemplo, en un mutante *FatB* creado por la inserción de un T-ADN, el contenido en palmitato (16:0) y estearato (18:0) se redujo en un 42 - 56% y 30 - 50%, dependiendo respectivamente del tejido (Bonaventure et al., The Plant Cell, 2003 15:1020-1033). La tasa de crecimiento de las plantas se vio afectada en gran medida en este mutante, dando como resultado un 50% menos de peso fresco a las 4 semanas, en comparación con el tipo silvestre. Este hallazgo apoya la opinión de que *FatB* sirve como un factor importante que gobierna el destino de los ácidos grasos saturados en las plantas.

35 El principal cultivo de semillas oleaginosas *de Brassica napus* es una especie anfidiplóide estrechamente relacionada con la especie modelo *Arabidopsis*, pero tiene un tamaño de genoma que es aproximadamente ocho veces mayor que el de *Arabidopsis*. Como resultado, se han descrito seis genes *FatB* en esta especie (documento WO 2009/007091) en comparación con uno en *Arabidopsis*. Esta complejidad hace que sea difícil entender la función de los genes *FatB* individuales para su manipulación. A la inversa, la presencia de un mayor número de genes hace que sea más fácil manipular la expresión de múltiples genes hacia el objetivo deseado de ácidos grasos, en lugar de desactivar uno completamente. Los factores de transcripción de dedos de zinc se sabe que se unen a esencialmente cualquier secuencia de ADN y modulan la función génica de los genes diana (Van Eenennaam et al., Metab. Eng. 2004 6:101-108; Sánchez et al., 2006, Plant Biotechnology Journal 4:103-114). Como resultado, esta herramienta se puede aplicar a los genes *FatB* para comprender su función en *B. napus*. Este conocimiento podría ayudar en la

manipulación de la expresión de uno a múltiples genes *FatB* para obtener el aceite de semilla de canola más "sano" deseado que tiene menos ácidos grasos saturados.

Los Ejemplos 8-15 siguientes muestran una regulación transcripcional al alza de genes *FatB* en *Brassica napus* L. produciendo cambios en el perfil de ácidos grasos de las semillas.

5 Ejemplo 8: Identificación de la secuencia diana para genes *FatB* en *Brassica napus* L.

8.1 Identificación de la secuencia diana

Se empleó un protocolo de paseo genómico (del inglés, "Genome Walking") modificado para descubrir y caracterizar las secuencias promotoras para los genes *FatB* de *Brassica napus* variedad Nex710. Estas secuencias de nucleótidos se aislaron e identificaron para el diseño y la producción de ZFP TFs, que se pueden utilizar para modificar la expresión génica del gen *FatB*. Nueve genotecas de paseo genómico se construyeron utilizando el kit de Clontech Genome Walking (Palo Alto, CA). Estas genotecas se utilizaron para obtener ~1 kb de la secuencia aguas arriba del sitio de inicio de la transcripción para los genes *FatB*. Una pareja de cebadores directos anidados específicos del gen *FatB* se diseñaron y se sintetizaron basándose en secuencias *FatBEST* disponibles procedentes de bases de datos públicas. Una pareja de cebadores inversos anidados se diseñaron y se sintetizaron para hibridarse con las secuencias adaptadoras. Múltiples fragmentos de PCR fueron amplificados a partir de las nueve genotecas. Los fragmentos de PCR resultantes se clonaron en vectores TOPO (Invitrogen, Carlsbad, CA, EE.UU.) y se secuenciaron para identificar las secuencias promotoras de *FatB*. Se identificaron dos genes diferentes *FatB*, *FatB4* y *FatB5*, gracias a este enfoque. Se observó que, aunque las secuencias codificadoras de estos genes *FatB* estaban muy conservadas, las secuencias divergían en las regiones promotoras 5' UTR y de promotor.

20 La expresión de *FatB* y la cuantificación de los genes *FatB4* y *FatB5* mediante análisis con RT-PCR cuantitativa en el desarrollo de semillas de *B. napus*, sugería que los genes se sobreexpresaban mucho en las semillas de *B. napus*. Se aisló el ARN total, se cuantificó y los sitios de iniciación de la transcripción se cartografiaron como se ha descrito anteriormente en el Ejemplo 1. Las secuencias promotoras de estos dos genes se utilizaron para diseñar ZFP TFs para una modificación de la expresión transcripcional.

25 Ejemplo 9: Diseño de dominios de unión de ADN de ZFP específicos de genes *FatB*

Las proteínas con dedos de zinc se diseñaron contra varios sitios diana en el gen *FatB*. Los sitios diana y las hélices de reconocimiento para los diseños de ZFP representativos, se muestran a continuación en las Tablas 10A y 10B.

Tabla 10A: Sitios de unión diana para ZFP TFs de *FatB*

ZFP	Sitio diana de <i>FatB4</i> (5' a 3')	Sitio diana de <i>FatB5</i> (5' a 3')	Especificidad de <i>FatB</i>
13685*	aaCGAAAAGgAGATCGAGAGAGg agagag (SEQ ID NO:47)	Sin sitio de unión	<i>FatB4</i>
13714	Sin sitio de unión	cgAAAGGGAGATCGAGAGAG gcaccgca (SEQ ID NO:48)	<i>FatB5</i>
13722	aaGGAGAAcTTTAGGGTTTGGg agact (SEQ ID NO:49)	aaGGAGAA/TTTAGGGTTTGG ggagact (SEQ ID NO:50)	<i>FatB4</i> y <i>FatB5</i>
13743**	ctCCGAAGAGATTGGCGTAAcac ttcgt (SEQ ID NO:51)	ctCCGAAGAGATTGGCGTAAc cttcatt (SEQ ID NO:52)	<i>FatB4</i> y <i>FatB5</i>

Tabla 10B: Regiones de reconocimiento de hélice de ZFP

ZFP	1***	2	3	4	5	6
13685	RSDNLSA (SEQ ID NO:75)	QSAHRKT (SEQ ID NO:76)	RSDDLK (SEQ ID NO:21)	QSSHRKT (SEQ ID NO:77)	RSDHLSV (SEQ ID NO:78)	QNAHRIE (SEQ ID NO:79)
13714	RSDNLSA (SEQ ID NO:75)	QSAHRKT (SEQ ID NO:76)	RSDDLK (SEQ ID NO:21)	QSSHRKT (SEQ ID NO:77)	RSDHLSK (SEQ ID NO:80)	QANARIT (SEQ ID NO:81)
13722	RSDHLST (SEQ ID NO:82)	HSNTRKN (SEQ ID NO:83)	RSDHLSQ (SEQ ID NO:84)	NSASRKN (SEQ ID NO:85)	QSGNLAR (SEQ ID NO:16)	QSGHLR (SEQ ID NO:86)
13743	NSDSLTE (SEQ ID NO:87)	RRADLSR (SEQ ID NO:88)	RSDLSA (SEQ ID NO:89)	QNAHRKT (SEQ ID NO:90)	RSDHLSQ (SEQ ID NO:84)	RNADRIT (SEQ ID NO:91)

\* Las letras mayúsculas representan la secuencia de unión de ZFP, mientras que las letras minúsculas flanqueantes representan el contexto de la secuencia flanqueante para el sitio de unión de ZFP TF. Las letras minúsculas en cursiva entre las letras mayúsculas representan las bases omitidas en el diseño de ZFP.

\*\* La secuencia de unión de ZFP es la misma entre FatB4 y FatB5, pero las secuencias flanqueantes 3' son diferentes entre los dos genes.

\*\*\* Los números representan los dedos de zinc (1 es el dedo 1, 2 es el dedo 2, etc.)

Ejemplo 10: Regulación al alza mediada por ZFP TF de genes *FatB* naturales en *B. napus*

5 Para ejemplificar la regulación al alza mediada por ZFP TF de *FatB* dentro de células de *B. napus*, se construyeron estructuras artificiales que contenían cuatro diseños de genes *ZFP TF* (Tablas 10 y 11), estos genes se entregaron de manera estable en las células de *B. napus* a través de una transformación mediada por *Agrobacterium*.

10.1 Diseño de la estructura artificial

10 Los plásmidos binarios de ZFP TF diseñados y construidos, se indican en la Tabla 11 a continuación. La expresión de los genes de ZFP TF estaba dirigida por un promotor constitutivo relativamente fuerte, tal como un promotor obtenido a partir del promotor del virus del mosaico de nervaduras de mandioca (CsVMV). La transformación de las plantas mediada por *Agrobacterium* se llevó a cabo como se ha descrito en el Ejemplo 3.2 para incorporar de manera estable ZFP TFs en el genoma de *B. napus*.

Tabla 11: Descripción de estructuras artificiales de ZFP TFs dirigidas a los genes *FatB* de *B. napus*

S.N.	ZFP	Nº de la estructura artificial	Casete del gen
1	13685	pDAB4689	RB7 MAR/CsVMV/Op-2 NLS-ZFP 13685 VP16/AtuORF23 3'UTR/AtUbi10/Pat/AtuORF1 3' UTR
2	13714	pDAB4690	RB7 MAR/CsVMV/Op-2 NLS-ZFP 13714-VP16/AtuORF23 3'UTR/AtUbi10/Pat/AtuORF1 3' UTR
3	13722	pDAB4691	RB7 MAR/CsVMV/Op-2 NLS-ZFP 13722-VP16/AtuORF23 3'UTR/AtUbi10/Pat/AtuORF1 3' UTR
4	13743	pDAB4692	RB7 MAR/CsVMV/Op-2 NLS-ZFP 13743-VP16/AtuORF23 3' UTR/AtUbi10/Pat/AtuORF1 3' UTR

15 pDAB4689 es un plásmido binario que contiene la secuencia de localización nuclear *opaque-2* (Op-2) o las casetes de expresión génica NLS /13685 /VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; Fusión RB7 MAR v3 (región de fijación a la matriz (Thompson et al., 1997, documento WO9727207)) :: CsVMV Pro-

- motor v2 (Promotor del virus del mosaico de las nervaduras de mandioca (Verdaguer et al., 1996, Plant Molecular Biology 31: 1129-1139)) :: *opaque-2* NLS (Van Eenennaam et al., 2004, Metabolic Engineering 6:101-108; Holmes-Davis et al., 2005, Plant Molecular Biology 57:411-423)/13685 Dedo de zinc/VP16 (Jamieson et al., Biochem. Biophys. Res. Commun. 2006, 348:873-879) :: AtuORF23 3'UTR v1(marco de lectura abierto 23 de *Agrobacterium tumefaciens*, región no traducida 3' (Gelvin et al., 1987, documento EP222493)) :: AtUbi10 Promotor v2 (Promotor Ubicuitina-10 de *Arabidopsis thaliana* (Callis, et al., 1990, J. Biol. Chem. 265-12486-12493)) :: *pat* v5 (acetil transferasa de fosfotricina (Wohlleben et al., 1988, Gene 70:25-37)) :: AtuORF 1 3'UTR v4 (marco de lectura abierto 1 de *Agrobacterium tumefaciens*, región no traducida 3' (Huang et al., J. Bacteriol. 172:1814-1822)). El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13685 dedo de zinc/VP16 en pDAB3912 a través de los sitios de restricción *NcoI* - *SacI*. La estructura artificial resultante que se etiquetó como pDAB8215 contenía la fusión CsVMV Promotor v2 :: *opaque-2* NLS/13685 dedo de zinc/VP16 :: casete de expresión génica AtuORF23 3'UTR v1. pDAB8215 se clonó en el binario pDAB4668 a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4689 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.
- 15 pDAB4690 es un plásmido binario que contiene las casetes de expresión génica *opaque-2* NLS/13714/VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; Fusión RB7 MAR v3 :: CsVMV Promotor v2 :: *opaque-2* NLS/13714 dedo de zinc/VP16 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13714 dedo de zinc/VP16 en pDAB3912 a través de los sitios de restricción *NcoI*- *SacI*. La estructura artificial resultante que fue etiquetada como pDAB8216, contenía la fusión CsVMV Promotor v2 :: *opaque-2* NLS/13714 dedo de zinc/VP16 :: casete de expresión génica AtuORF23 3'UTR v1. pDAB8216 se clonó en el binario pDAB4668 a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4690 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.
- 25 pDAB4691 es un plásmido binario que contiene las casetes de expresión génica *opaque-2* NLS/13722/VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; Fusión RB7 MAR v3 :: CsVMV Promotor v2 :: *opaque-2* NLS/13722 dedo de zinc/VP16 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 (acetil transferasa de fosfotricina) :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13722 dedo de zinc/VP16 en pDAB3912 a través de los sitios de restricción *NcoI*- *SacI*. La estructura artificial resultante que fue etiquetada como pDAB8217, contenía la fusión CsVMV Promotor v2 :: *opaque-2* NLS/13722 dedo de zinc/VP16 :: casete de expresión génica AtuORF23 3'UTR v1. pDAB8217 se clonó en el binario pDAB4668 a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4691 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.
- 35 pDAB4692 es un plásmido binario que contiene las casetes de expresión génica *opaque-2* NLS/13743/VP 16 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; Fusión RB7 MAR v3 (región de fijación a la matriz) :: CsVMV Promotor v2 :: *opaque-2* NLS/13743 dedo de zinc/VP16 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 (Promotor de ubicuitina-10 de *Arabidopsis thaliana*) :: *pat* v5 :: AtuORF 1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13743 dedo de zinc/VP16 en pDAB3912 a través de los sitios de restricción *NcoI*- *SacI*. La estructura artificial resultante que fue etiquetada como pDAB8218, contenía la fusión CsVMV Promotor v2 :: *opaque-2* NLS/13743 dedo de zinc/VP16 :: casete de expresión génica AtuORF23 3'UTR v1. pDAB8218 se clonó en el binario pDAB4668 a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4692 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

#### Ejemplo 11: Análisis de muestras de callos transformados

- 45 Las líneas de callos transgénicos de *B. napus* fueron producidas mediante la transformación de cuatro estructuras artificiales de ZFP TF, pDAB4689 - pDAB4692 y una estructura artificial de control, pDAB8210, a través de una transformación de plantas mediada por *Agrobacterium*, como se ha descrito en el Ejemplo 3.2. La estructura artificial de control, pDAB8210, se componía de una casete de expresión génica *pat* (AtUbi10 Promotor/*pat*/ AtuORF1 3' UTR) y una casete de ZFN no diana (CsVMV Promotor/ZFN/AtuORF23 3' UTR). El ARN total se extrajo y se sintetizó el ADNc de todas las líneas, como se ha descrito en el Ejemplo 4.1. La única modificación frente al protocolo descrito anteriormente, era el uso de cebadores Oligo dT para el cebado de las reacciones de ADNc, en lugar de hexámeros aleatorios y la modificación correspondiente de la reacción de ADNc, a través de las especificaciones del fabricante.

#### 11.1 Análisis de la expresión del ARNm de *FatB4* y *FatB5*

##### 11.1.1 Ensayo con PCR cuantitativa en tiempo real (qRT-PCR) de *FatB4* para *B. napus*

- 55 Una mezcla de PCR se estableció de la siguiente manera para la amplificación del ADNc de *FatB4* : 1,5 µl de 10X tampón de PCR Hot Start (Qiagen, Valencia, EE.UU.), 1,2 µl de dNTPs 10 mM, 1 µl de MgCl<sub>2</sub> 25 mM, 0,15 µl de Taq de Qiagen Hot Start (5 U/µl), 0,5 µl de cebador zz143 FW de una solución de reserva 10 µM (SEQ ID NO: 53: CTTTGAACGCTTATCTTCCTC) (solución de reserva 10 µM), 0,5 µl de cebador FATB5\_R4 REV de una solución de reserva 10 µM (SEQ ID NO: 54: TTCCACAACATCTCCCAAG), 0,25 µl de sonda TaqMan MGB (Life Technologies,

Carlsbad, California) *FatB4\_MGB\_Probe\_4* de una solución de reserva 5  $\mu\text{M}$  (SEQ ID NO: 55: FAM-CTCAGGCTCCACCC), 1,5  $\mu\text{l}$  de 10% (p/v) de PVP-40 y  $\text{H}_2\text{O}$  hasta 13,5  $\mu\text{l}$  por reacción. El o los cuadrantes apropiados de una microplaca de 384 pocillos se separaron y se llenaron con 13,5  $\mu\text{L}$  de mezcla maestra por pocillo. A continuación, se fijó suavemente una lámina de sellado a la placa. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrífuga de microplacas de Qiagen. A continuación, se añadieron 1,5  $\mu\text{l}$  de ADNc de primeras hebras descongeladas, diluidas a los pocillos apropiados, seguido de la adición de 1,5  $\mu\text{l}$  de patrones de ADNc plasmídico de menor a mayor concentración de ADN, en los pocillos de control. Finalmente, se fijó firmemente una lámina de sellado a la placa y se centrifugó. El programa de PCR se ejecutó en un sistema de PCR en tiempo real LIGHTCYCLER<sup>®</sup> 480 (Roche Diagnostics, Indianapolis, IN, EE.UU.), utilizando el programa siguiente: 1) activar a 95°C durante 15 minutos; 2) desnaturalizar a 95°C durante 20-30 segundos (@ 4,8°C/s); 3) reasociar a 60°C durante 20-30 segundos (@ 2,5°C/s); 4) adquirir a 72°C durante 45-60 s (@ 4,8°C/s); repetir las etapas 2) - 4), 39-49 veces; 5) enfriar a 38°C durante 5 s para detener la reacción.

El amplicón *FatB4* tenía un tamaño de 678 pb. Esta secuencia incluye un intrón de 79 pb basándose en la secuencia genómica. Además, el cebador inverso se diseñó para incluir un intrón, favoreciendo de este modo la amplificación específica solamente de los ADNc de *FatB4*, eliminando de esta manera la amplificación del ADN genómico.

#### 11.1.2 Ensayo con qRT-PCR de *FatB5* para *B. napus*

Una mezcla de PCR se estableció de la siguiente manera para la amplificación del ADNc de *FatB5*: 1,5  $\mu\text{l}$  de 10X tampón de PCR Hot Start (Qiagen, Valencia, EE.UU.), 1,2  $\mu\text{l}$  de dNTPs 10 mM, 1  $\mu\text{l}$  de  $\text{MgCl}_2$  25 mM, 0,15  $\mu\text{l}$  de Taq de Qiagen Hot Start (5 U/ $\mu\text{l}$ ), 0,5  $\mu\text{l}$  de cebador zz145 FW de una solución de reserva 10  $\mu\text{M}$  (SEQ ID NO: 56: CTTTGAAAGCTCATCTTCCTC), 0,5  $\mu\text{l}$  de cebador *FATB5\_R4* REV de una solución de reserva 10  $\mu\text{M}$  (SEQ ID NO: 57: TTCCACAACATCTCCCAAG), 0,25  $\mu\text{l}$  de sonda TaqMan MGB (Life Technologies, Carlsbad, California) *FatB5\_MGB\_Probe\_4* de una solución de reserva 5  $\mu\text{M}$  (SEQ ID NO: 58: FAM-AACCTTCATCCTCCCA), 1,5  $\mu\text{l}$  de 10% (p/v) de PVP-40 y  $\text{H}_2\text{O}$  hasta 13,5  $\mu\text{l}$  por reacción. El ensayo restante y las especificaciones del ciclo de PCR eran como se ha descrito para el ensayo de qRT-PCR de *FatB4* en el Ejemplo 11.1.1.

El amplicón producido tenía 678 pb de tamaño. Esta secuencia de ADNc incluye un intrón de 76 pb basándose en la secuencia genómica. Además, el cebador inverso se diseñó para incluir un intrón favoreciendo de este modo la amplificación del ADNc de *FatB5*, eliminando de esta manera la amplificación del ADN genómico.

#### 11.2 Análisis de la expresión del ARNm de tubulina

Este ensayo se completó como se describe en el Ejemplo 4.3.

#### 11.3 Análisis de la expresión del ARNm de ZFP TF

El análisis del ARNm de *ZFP TF* de una muestra de callo  $T_0$  se completó según el Ejemplo 4.4. El tamaño del amplicón en este ensayo era menor de 1 Kb, dependiendo del tamaño de ZFP en cada uno de los diseños. La cuantificación de la expresión de *ZFP TF* de una planta  $T_0$  se realizó con un ensayo basado en el dominio de activación VP16 del modo siguiente. Se estableció una mezcla de *VP16* qRT-PCR del modo siguiente para la amplificación del ADNc de *ZFP TF*: 7,5  $\mu\text{l}$  de 2X tampón LIGHTCYCLER<sup>®</sup> 480 Probes Master (Roche Diagnostics, Indianapolis, IN, EE.UU.), 0,3  $\mu\text{l}$  del cebador *VP16\_UPL\_F1* de una solución de reserva 10  $\mu\text{M}$  (SEQ ID NO: 59: TCGATCTTGATA-TGTTGGGAGA), 0,3  $\mu\text{l}$  de *VP16\_UPL\_R1* de una solución de reserva 10  $\mu\text{M}$  (SEQ ID NO: 60: AGGTGCAGAAT-CATGTGGTG), 0,15  $\mu\text{l}$  de la sonda UPL n° 85 (Roche Diagnostics, Indianapolis, IN, EE.UU.), 1,5  $\mu\text{l}$  de 10% (p/v) de PVP-40) y  $\text{H}_2\text{O}$  hasta 13,65  $\mu\text{l}$ . El o los cuadrantes apropiados de una microplaca de 384 pocillos se separaron y se llenaron con 13,5  $\mu\text{L}$  de mezcla maestra por pocillo. A continuación, se fijó suavemente una lámina de sellado sobre la placa. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrífuga de microplacas de Qiagen. A continuación, se añadieron 1,5  $\mu\text{l}$  de ADNc de primeras hebras descongeladas, diluidas, seguido de la adición de 1,5  $\mu\text{l}$  de patrones de ADNc plasmídico de menor a mayor concentración de ADN, en los pocillos de control. Se fijó firmemente una lámina de sellado sobre la placa y se centrifugó. Un programa de PCR se ejecutó en un sistema de PCR en tiempo real LIGHTCYCLER<sup>®</sup> 480 (Roche Diagnostics, Indianapolis, IN, EE.UU.), utilizando el programa siguiente: 1) activar a 95°C durante 15 minutos; 2) desnaturalizar a 95°C durante 10 segundos (@ 4,8°C/s); 3) reasociar/extender a 60°C durante 25 segundos (@ 2,5°C/s); 4) adquirir a 72°C durante 1 s (@ 4,8°C/s); repetir las etapas 2 - 4), 39-49 veces. Finalmente la reacción se enfrió a 38°C durante 5 s para detener la reacción.

El tamaño del amplicón *VP16* era de 68 pb, este tamaño fragmento se corresponde con el fragmento *VP16* que se esperaba que fuera producido mediante la amplificación con PCR.

#### 11.4 Análisis de la expresión de las muestras de callos

Líneas de callos transgénicos que crecían sobre una selección de *Herbiace*<sup>®</sup> se analizaron para estudiar los niveles de expresión de *FatB4*, *FatB5*, *tubulina* y *ZFP TF* mediante qRT-PCR. El gen *tubulina* servía como gen de referencia para normalizar la expresión de los niveles de ARNm de *FatB4*, *FatB5* y *ZFP TF*. Las relaciones *FatB4/tubulina* y *FatB5/tubulina* se calcularon para normalizar las expresiones de ARNm. Los resultados mostraron una regulación al alza estadísticamente significativa del ARNm de *FatB4* y *FatB5* para el diseño de la estructura artificial pDAB4691 (p = 0,005; Fig. 9). Las líneas de callos pDAB4692 mostraron una tendencia a la regulación al alza en el ARNm de

*FatB4* y *FatB5*, pero la tendencia no era estadísticamente significativa. Las líneas de callos pDAB4689 también se analizaron en un experimento distinto, que mostraba una tendencia de regulación al alza, pero no era estadísticamente significativa (datos no mostrados). Las líneas celulares de control no amplificaban un amplicón específico de *ZFP TF*, confirmando que el ensayo de *ZFP TF* era específico de líneas que solamente expresaban *ZFP TF*.

## 5 Ejemplo 12: Análisis de la expresión de muestras de plantas T<sub>0</sub> transgénicas

### 12.1 Ensayo para la actina de *B. napus* para uso como un control interno para la qRT-PCR

El ARN total se extrajo y el ADNc se sintetizó a partir de todas las líneas como se describe en el Ejemplo 11. Se preparó una mezcla de PCR del modo siguiente para la amplificación del ADNc de *actina* (Bo Yang et al., 2007, Plant Science 173:156-171; n° de orden para el gen de actina en GENBANK, AF111812): 7,5 µl de 2X tampón LIGHTCYCLER® 480 Probes Master, 0,3 µl del cebador BN\_Actin\_F de una solución de reserva 10 µM (SEQ ID NO: 61: ACGAGCTACCTGACGGACAAG), 0,3 µl de BN\_Actin\_R de una solución de reserva 10 µM (SEQ ID NO: 62: GAGCGACGGCTGGAAGAGTA), 1,5 µl de 10% (p/v) de PVP-40) y q/s con H<sub>2</sub>O hasta 13,5 µl. El o los cuadrantes apropiados de una microplaca de 384 pocillos se separaron y se llenaron con 13,5 µL de mezcla maestra por pocillo. A continuación, se fijó suavemente una lámina de sellado sobre la placa. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrifuga de microplacas de Qiagen. A continuación, se añadieron 1,5 µl de ADNc de primeras hebras descongeladas, diluidas, seguido de la adición de 1,5 µl de patrones de ADNc de menor a mayor concentración de ADN, en los pocillos de control. Se fijó firmemente una lámina de sellado sobre la placa y se centrifugó. Un programa de PCR se ejecutó en un sistema de PCR en tiempo real LIGHTCYCLER® 480 (Roche Diagnostics, Indianapolis, IN, EE.UU.), utilizando las siguientes condiciones: 1) activar a 95°C durante 10 minutos; 2) desnaturalizar a 95°C durante 10 segundos (@ 4,8°C/s); 3) reasociar/extender a 60°C durante 20 segundos (@ 2,5°C/s); 4) adquirir a 72°C durante 20 s (@ 4,8°C/s); repetir las etapas 2 - 4, 39-49 veces. Finalmente la reacción se enfrió a 38°C durante 5 s para detener la reacción. El amplicón producido tenía un tamaño de 80 pb.

### 12.2 Análisis de la expresión del ARNm de *FatB4* y *FatB5*

Plantas *B. napus* que habían sido transformadas con pDAB4689-pDAB4691, se analizaron para determinar el aumento de la expresión del ARNm de *FatB4* y *FatB5*. Además, se utilizaron dos tipos de controles. Un control no transgénico que consistía en plantas Nex710 sirvió como control negativo. Un segundo control transgénico de plantas Nex710 de *B. napus* que habían sido transformadas con pDAB8210 también se analizó. El diseño de la estructura artificial pDAB8210 consistía en dos casetes de expresión génica. La casete de expresión génica *pat* (*promotor* AtUbi10:: *gen pat* :: *AtuORF1* 3' UTR) y una casete de expresión génica no diana de la nucleasa de dedos de zinc (ZFN) (*AtUbi10* promotor :: *gen ZFN* :: *AtuORF23* 3' UTR). La ZFN expresada en pDAB8210 no se esperaba que se uniera y escindiera el genoma de plantas Nex710 de *B. napus* y no debía alterar el fenotipo de estas plantas.

Se regeneraron callos putativos transgénicos que crecían sobre una selección de *Herbiace*® en plantas (Ejemplo 3.3). Las muestras de hojas se recogieron a partir de las plantas en etapa de 6 hojas de *B. napus* que crecían en el invernadero para el análisis de la expresión del ARNm. Seis perforaciones de hojas de cada planta se aislaron y se colocaron sobre hielo antes de la extracción del ARN. El ARN total se extrajo usando el kit RNeasy de Qiagen. La síntesis de ADNc y las diluciones posteriores se completaron como se ha descrito en el Ejemplo 11. El protocolo del análisis de la expresión para *FatB4* y *FatB5* a través de qRT-PCR se ha descrito anteriormente. El análisis de la expresión de *ZFP TF* era más/menos con el ensayo. El ensayo de qRT-PCR del gen de referencia *actina* se ha descrito anteriormente.

La expresión en hojas T<sub>0</sub> de los genes *FatB4* y *FatB5* variaba entre las estructuras artificiales (Fig. 10). La mayor regulación al alza del ARNm de *FatB4* y *FatB5* se observó en eventos de plantas transgénicas pDAB4691 que producían incrementos generales de 2,0 - 2,5 en las expresiones del ARNm, en comparación con controles no transgénicos Nex710 y transgénicos pDAB8210. La expresión de ambos genes era estadísticamente significativa (por debajo de p = 0,05) en comparación con los controles. Por lo tanto, una caracterización adicional de las estructuras artificiales continuó con pDAB4691.

## Ejemplo 13: Análisis de plantas transgénicas T<sub>1</sub>

### 13.1 Análisis de los ácidos grasos de semillas T<sub>1</sub>

Se observó una diferencia significativa en el contenido de C18:0 contenido entre semillas oleaginosas "bajo" y "alto" del evento pDAB4691-003-049.1 ("evento 49") (Tabla 12). El contenido de C18:0 de la categoría de semillas "baja" era similar al de las líneas Nex710 (control no transgénico) y pDAB8210 (control transgénico). El contenido de C18:0 en las semillas "alta" se correlacionaba con la expresión de *ZFP TF* que provocaba una regulación al alza del o de los genes *FatB*. Algunos aumentos de C20:0, C22:0 y C24:0 también se observaron basándose en el flujo de C18:0 a ácidos grasos de cadena más larga. El contenido de C16:0 en semillas individuales no parecía cambiar con ninguna de las categorías "baja" o "alta". Este resultado es indicativo del hecho de que *ZFP TF* de pDAB4691 se une y regula al alza el o los genes *FatB* específicos de las reacciones enzimáticas C18:0-ACP a C18:0, en lugar de las reacciones C16:0-ACP a C16:0 (véase el Ejemplo 1, Fig. 1).

Tabla 12: Perfil de ácidos grasos de semillas individuales T<sub>1</sub> medido con análisis FAME

Muestras*	Total Sats	C16:0	C16:1	C18:0	C20:0	C22:0	C24:0
Evento 49 bajo	6,35	4,11	0,48	1,45	0,46	0,21	0,07
Evento 49 bajo	6,35	3,70	0,35	1,62	0,58	0,30	0,12
Evento 49 bajo	6,49	3,64	0,36	1,66	0,62	0,32	0,18
Evento 49 alto	7,50	3,55	0,42	2,59	0,80	0,34	0,17
Evento 49 alto	7,74	3,63	0,45	2,59	0,88	0,38	0,22
Evento 49 alto	7,93	3,67	0,47	2,82	0,86	0,38	0,17
Media del evento 49 (n = 24)	6,90	3,70	0,40	2,03	0,67	0,31	0,15
DS/DE del evento 49 (n = 24)	0,51	0,14	0,05	0,37	0,10	0,04	0,03
Media de Nex710 (n = 24)	6,47	3,70	0,45	1,79	0,57	0,26	0,11
DS/DE de Nex710 (n = 24)	0,35	0,24	0,06	0,17	0,06	0,04	0,03
Media de 8210 (n = 150)	6,27	3,51	0,34	1,63	0,61	0,33	0,14
DS/DE de 8210 (n = 150)	0,40	0,24	0,07	0,25	0,08	0,05	0,04
* El análisis de ácidos grasos se realizó en 24 semillas T <sub>1</sub> individuales del evento 49 y Nex710 y 144 semillas de pDAB8210 que comprendía seis eventos. Las categorías "bajo" y "alto" representan el contenido en FA específico de la semilla individual cuando se ordena en función del contenido en C18:0.							

### 13.2 Análisis de la presencia de ZFP TF en plantas T<sub>1</sub>

- 5 Se plantaron cien semillas T<sub>1</sub> del evento 49 se plantaron en el invernadero. Noventa y siete plantas germinaron en plántulas. El número de copias de ZFP TF se estimó utilizando un ensayo de qRT-PCR de *pat* como se ha descrito en el Ejemplo 6.2 para obtener material vegetal aislado de la etapa de 2-4 hojas. El evento 49 contenía múltiples inserciones (~3 inserciones) y 0-7 copias de genes *pat*, y por lo tanto ZFP TF, que segregaban en la población T<sub>1</sub>. Solo se identificó una planta nula a partir de la población T<sub>1</sub>.

### 13.3 Regulación al alza del ARNm de *FatB4* y *FatB5* en plantas T<sub>1</sub>

- 10 El análisis de la expresión génica de *FatB* natural se realizó en las 97 plantas de la población T<sub>1</sub> que había segregado. Se incluyeron plantas no transgénicas Nex710 que se habían plantado al mismo tiempo en el estudio como controles. Seis perforaciones de hojas por planta se recogieron y se colocaron en hielo hasta que se pudo completar la extracción de ARN. El ARN total se extrajo usando el kit RNeasy de Qiagen. La síntesis del ADNc y las diluciones posteriores se completaron como se ha descrito en el Ejemplo 11. El análisis de la expresión de *FatB4*, *FatB5*, *tubulina* y *VP16* (ZFP TF) se realizó como se ha descrito anteriormente.

- 15 Un análisis estadístico de las relaciones entre *FatB4/FatB5* y *tubulina* mostraba tendencias lineales significativas con la expresión de *tubulina* (Fig. 11), lo que indicaba que aumentos en la relación *FatB4/FatB5*, no tenían una relación 1:1 con los aumentos en el control endógeno, *tubulina*. Como prueba, las relaciones no se utilizaron en este análisis y *tubulina* se incluyó como una covariable en los modelos de la expresión de *FatB4/FatB5*. Para eliminar cualquier colinealidad entre *tubulina* y *VP16* (ZFP TF), se ortogonalizaron las matrices de incidencia para las ecuaciones del modelo.

El siguiente modelo era adecuado para todos los conjuntos de datos en el estudio actual: en donde estado<sub>i</sub> es el estado transgénico de y<sub>ijk</sub> (transgénico o nulo/control);

$$y_{ijk} = \text{estado}_i + \text{VP16 transgénico}^* + e_{ijk}$$

5 *VP16 transgénico\** es una regresión lineal sobre *VP16* anidado dentro de las líneas transgénicas; y e<sub>ijk</sub> es el residuo aleatorio. Dado que había tres mediciones repetidas para cada muestra, se utilizó una estructura residual correlacionada:

$$R = \sigma_e^2 \left( \mathbb{I}_n \otimes \begin{bmatrix} 1 & \phi & \phi \\ \phi & 1 & \phi \\ \phi & \phi & 1 \end{bmatrix} \right)$$

en donde I<sub>n</sub> es la matriz de identidad con rango igual al número de muestras y φ es la correlación entre los residuos de mediciones repetidas.

10 Los resultados del análisis para el evento 49 detectaron una regulación al alza significativa en *FatB4* con regresiones anidadas de la expresión de *FatB4* en *VP16*, lo que mostraba pendientes positivas muy significativas (Fig. 12, Tabla 13). Para *FatB5*, las regresiones anidadas mostraban pendientes positivas para la regulación al alza y la pendiente era estadísticamente significativa. Para los eventos de regulación al alza hay que señalar que las plantas transformadas y de control no eran de la misma línea, por lo que podría haber algunos factores de confusión de la línea y efectos de transformación.

15

Tabla 13: Resultados para los eventos de regulación al alza

Evento	Gen	Efecto <sup>A</sup>	Solución	Valor de P
4691-3-049.001	FatB4	Control transformado	10331.11	<0,00001
		VP16 transformado*	0.764E-05	<0,00001
	FatB5	Control transformado	5084.09	<0,00001
		VP16 transformado*	0.308E-05	<0,00001

<sup>A</sup>Control transformado = la diferencia en la expresión de mínimos cuadrados promedio entre las líneas transformadas y las líneas que no eran portadoras de transgenes. VP16 transformado\* = regresión de la expresión de FatB4/FatB5 sobre la expresión de VP16 anidada dentro de las líneas transgénicas.

Plantas T<sub>1</sub> del evento 49 se clasificaron después en función de las relaciones de expresión de *FatB4/tubulina* y luego *FatB5/tubulina* para encontrar las plantas con mayor expresión para el avance a la generación T<sub>2</sub> en el invernadero. Ocho de las plantas con mayor expresión de *FatB4* se seleccionaron para el avance a la madurez. Seis de estas plantas eran también las que tenían mayor expresión de *FatB5*. El número de copias de *ZFP TF* en estas plantas de expresión elevada, variaba de 2-4 copias mientras que el número de copias que segregaban en toda la progenie T<sub>1</sub> eran 0-7 copias. Para las plantas de control, una planta nula del evento 49 y diez plantas de control Nex710 también avanzaron a la madurez para recoger las semillas T<sub>2</sub>.

25 13.4 Análisis de ácidos grasos de semillas T<sub>2</sub>

El análisis de ácidos grasos (FA) se realizó en un volumen de 24 semillas de todas las plantas, como se ha descrito en el Ejemplo 6.1. El perfil de FA de una planta nula se combinó con diez plantas de control Nex710 para los cálculos del perfil de FA de control. Como promedio, se observó un aumento del 12% en el contenido de C18:0 para el evento 49 en comparación con el de Nex710 (Tablas 14A y 14B). Este aumento era estadísticamente significativo con p = 0,05 (Fig. 13). Los FAs de larga cadena, tales como C20:0, C22:0 y C24:0, también mostraron algunos aumentos que condujeron a aumentar el contenido en FA total saturados en un 5%. Puesto que la enzima FatB cataliza la conversión de ACP C18:0 en FA C18:0 libre, se acumulará más C18:0 como resultado de la regulación al alza transcripcional de *FatB4* y *FatB5* (Fig. 1). El pDAB4691 (Ejemplo 6, Tablas 1 y 2) es específico para ambos genes *FatB4* y *FatB5*, lo que da como resultado un cambio significativo en el perfil de FA (Tablas 14A y 14B).

35

Tabla 14A: Perfil de ácidos grasos (C16:0, C16:1, C18:0, C18:1, C18:2 y C18:3) de muestras del evento 49 transgénico y de control Nex710

S.N.	ID de la planta	Total Sats	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
1	4691-3-049 (8)	6,42	3,47	0,24	1,61	76,32	11,20	3,52
2	Nexera710 (11)	6,11	3,40	0,23	1,44	76,26	11,25	3,76

Los números entre paréntesis en la columna 2 representan el número de plantas analizadas.

5 Tabla 14B: Perfil de ácidos grasos (C20:0, C20:1, C20:2, C22:0, C22:1, C24:0, C24:1) de muestras del evento 49 transgénico y de control Nex710

S.N.	ID de la planta ID	Total Sats	C20:0	C20:1	C20:2	C22:0	C22:1	C24:0	C24:1
1	4691-3-049 (8)	6,42	0,66	1,39	0,05	0,43	0,02	0,21	0,18
2	Nexera710 (11)	6,11	0,61	1,43	0,06	0,41	0,02	0,20	0,17

Los números entre paréntesis en la columna 2 representan el número de plantas analizadas.

10 No se observaron cambios significativos en el contenido total de C16 con la regulación al alza de los genes *FatB4* y *FatB5* naturales. Es probable que se produzca una catálisis de C16:0-ACP a C16:0 con otros genes *FatB* pero no con *FatB4* y *FatB5*, las dianas ejemplificadas para la regulación al alza usando el diseño de ZFP TF pDAB4691 (Fig. 1).

El contenido en C16:0 de una semilla individual no parece cambiar en las categorías de semillas T1 "bajo" y "alto" (Tabla 12), lo que es indicativo del hecho que el diseño de ZFP TF pDAB4691 se une más específicamente a el o los genes *FatB* para la reacción enzimática C18:0-ACP a C18:0, en lugar de para la reacción C16:0-ACP a C16:0 (Fig. 1).

15 Se debe observar que la regulación al alza de los genes de la *sintetasa II* de  $\beta$ -cetoacil-ACP naturales (Ejemplos 1-6) inducidos por ZFP TF, provocaba diferentes cambios en el perfil de FA en comparación con regulaciones al alza del gen *FatB4* y *FatB5* natural. Por ejemplo, una regulación al alza del gen de la *sintetasa II* de  $\beta$ -cetoacil-ACP causaba una reducción estadísticamente significativa en los contenidos de C16:0 y C16:1 y un aumento concomitante del contenido total de C18, en comparación con sus segregantes nulos (Ejemplo 6, Fig. 6, Tablas 5 y 8). Comparativamente, la regulación al alza del gen *FatB* mediada por ZFP-TF provocaba aumentos estadísticamente significativos en el contenido de C18:0, pero ningún cambio aparente en el contenido de C16:0 (Fig. 14). Una vez más, estos cambios en los perfiles de FA a través de una regulación al alza mediada por ZFP TF de los genes de la *sintetasa II* de  $\beta$ -cetoacil-ACP y *FatB* coinciden en tanto a la ruta de biosíntesis de FA (Fig. 1).

Ejemplo 14: Diseño de las estructuras artificiales de ZFP TF para la transformación de plantas

25 Cuatro estructuras artificiales fueron diseñadas y construidas para una regulación a la baja transcripcional de los genes *FatB* en *B. napus* L (Tabla 15). Los mejores diseños de ZFP para la regulación al alza de *FatB*, 13722 y 13714, se emplearon para la demostración de la regulación a la baja de *FatB* (Ejemplos 9 -13). Estos diseños de ZFP se fusionaron con una señal de localización nuclear *opaque-2* y un dominio de represión consistente en KRAB1 (Hanna-Rose y Hansen, 1996, Trends in Genetics, 12: 229-234) o NtERF3 (Ohta et al., The Plant Cell, 2001, 13:1959-1968), para la construcción de ZFP TFs funcionales. Todos los ZFP TFs se expresaban bajo promotores específicos de semillas; se utilizó el promotor de proteína de transferencia de lípidos 2 de *Arabidopsis thaliana* (Promotor AtLTP170) (Genbank ID: NC\_003076) o el promotor de la  $\beta$ -faseolina de *Phaseolus vulgaris* (promotor PvPhas) (documento de Patente de EE.UU. n° 5.591.605).

Tabla 15: Detalles de la estructura artificial de ZFP TF para una regulación a la baja dirigida de los genes *FatB*

S.N.	Diseño de ZFP	Nº de estructura artificial	Casetes génicas
1	13722	pDAS5203	AtLTP170/Op-2* NLS-ZFP-13722-KRAB1/ORF23 UTR//AtUbi10/Pat/AtuORF1 3' UTR 3'
2	13714	pDAS5204	AtLTP170/Op-2 NLS-ZFP-13714-KRAB1/ORF23 UTR//AtUbi10/Pat/AtuORF1 3' UTR 3'
3	13722	pDAS5212	PvPhas/Op-2-NLS-ZFP13722-KRAB1/Phas UTR//AtUbi10/Pat/AtuORF1 3' UTR 3'
4	13722	pDAS5227	AtLTP170/Op-2 NLS-ZFP-13722-NtERF3/ORF23 UTR//AtUbi10/Pat/AtuORF1 3' UTR 3'

\* Op-2 = Señal de localización nuclear *Opaque-2*

pDAS5203 es un plásmido binario que contenía las casetes de expresión génicas *opaque-2* NLS/13722/KRAB1 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; fusión de AtLTP Promotor 170 :: *opaque-2* NLS/13722 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 :: AtuORF1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13722 dedo de zinc/KRAB1 en un vector donante Gateway. La estructura artificial resultante contenía la casete de expresión génica, fusión de AtLTP Promotor 170 :: *opaque-2* NLS/13722 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1. Esta estructura artificial se clonó en un binario a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción producía pDAS5203 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5204 es un plásmido binario que contenía las casetes de expresión génicas *opaque-2* NLS/13714/KRAB1 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; fusión de AtLTP Promotor 170 :: *opaque-2* NLS/13714 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 :: AtuORF1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13714 dedo de zinc/KRAB1 en un vector donante Gateway. La estructura artificial resultante contenía la casete de expresión génica, fusión de AtLTP Promotor 170 :: *opaque-2* NLS/13714 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1. Esta estructura artificial se clonó en un binario a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción producía pDAS5204 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5212 es un plásmido binario que contiene las casetes de expresión génicas *opaque-2* NLS/13722/KRAB1 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; fusión de PvPhas Promotor :: *opaque-2* NLS/13722 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 :: AtuORF1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13722 dedo de zinc/KRAB1 en un vector donante Gateway. La estructura artificial resultante contenía la casete de expresión génica, fusión de PvPhas Promotor :: *opaque-2* NLS/13722 dedo de zinc/KRAB1 :: AtuORF23 3'UTR v1. Esta estructura artificial se clonó en un binario a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción producía pDAS5212 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5227 es un plásmido binario que contiene las casetes de expresión génicas *opaque-2* NLS/13722/NtERF3 y *pat*. Esta estructura artificial incluye los siguientes elementos génicos; fusión de RB7 MAR v3 :: AtLTP Promotor 170 :: *opaque-2* NLS/13722 dedo de zinc/NtERF3 :: AtuORF23 3'UTR v1 :: AtUbi10 Promotor v2 :: *pat* v5 :: AtuORF1 3'UTR v4. El binario se construyó clonando un fragmento de ADN que contenía la fusión *opaque-2* NLS/13722 dedo de zinc/NtERF3 en un vector donante Gateway. La estructura artificial resultante contenía la casete de expresión génica, fusión de AtLTP Promotor 170 :: *opaque-2* NLS/13722 dedo de zinc/NtERF3 :: AtuORF23 3'UTR v1. Esta estructura artificial se clonó en un binario a través de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción producía pDAS5227 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

Las estructuras artificiales se transformaron de manera estable en la variedad Nex710 de *B. napus* mediante transformación mediada por *Agrobacterium* de explantes de hipocótilo (Ejemplo 3.2 - 3.3). Las plantas T<sub>0</sub> que crecían en invernadero se autopolinizaron y las semillas T<sub>1</sub> se recogieron cuatro meses después del trasplante.

Ejemplo 15: Análisis de transgénicos para semillas T1 para 4 estructuras artificiales de ZFP TF

Las semillas T<sub>1</sub> se obtuvieron a partir de eventos transgénicos (líneas) 8, 20, 37 y 15 de las estructuras artificiales

pDAS5203, pDAS5204, pDAS5212 y pDAS5227, respectivamente. Se emplearon cinco semillas de plantas Nex710 como controles. El análisis de los FAs se realizó en 24 semillas individuales de cada uno de los eventos transgénicos T<sub>1</sub>, como se ha descrito anteriormente en el Ejemplo 6.1, para identificar primero la estructura artificial de ZFP TF que altera más eficazmente el perfil de los FA.

5 15.1 Análisis de los ácidos grasos (FA) de semillas T1

10 Las estructuras artificiales pDAS5203, pDAS5212 y pDAS5227 producían eventos transgénicos con los perfiles totales de FA saturados más bajos, en comparación con el control Nex710 (Tabla 16). Las diferencias por parejas entre cada uno de los dos eventos de estructuras artificiales, pDAS5203 y pDAS5212 y el control Nex710 eran significativas ( $p = <0,005$ ). Sin embargo, la diferencia por parejas de pDAS5227 mostraba una fuerte tendencia ( $p = 0,06$ ). Uno de los eventos de pDAS5203 tenía el contenido total más bajo de FA saturados, de un 5,3% (no se muestra), en comparación con el promedio de 6,38% en el control Nex710. Esta es una reducción del 17% en los FAs totales saturados.

Tabla 16A: Perfil de FA en semillas T1 de múltiples eventos transgénicos de estructuras artificiales

Nº de estructura artificial	Tipo de análisis	Total Sats	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
pDAS5212	Media (n=37)*	6,00	0,05	3,60	0,24	1,38	75,88	12,24	3,41
pDAS5212	DS/DE (n=37)	0,21	0,01	0,20	0,02	0,14	1,36	1,03	0,41
pDAS5212	Valor de p**	0,01	0,07	0,89	0,47	<0,0001	0,39	0,37	0,94
pDAS5204	Promedio (n=20)	6,26	0,05	3,52	0,25	1,65	76,33	11,52	3,58
pDAS5204	DS/DE (n=20)	0,29	0,01	0,15	0,02	0,16	0,78	0,66	0,30
pDAS5204	Valor de p	0,43	0,72	0,52	0,69	0,32	0,16	0,04	0,34
pDAS5203	Promedio (n=8)	5,98	0,05	3,68	0,26	1,33	75,32	12,74	3,55
pDAS5203	DS/DE (n=8)	0,36	0,01	0,25	0,05	0,13	1,76	1,31	0,47
pDAS5203	Valor de p	0,02	0,06	0,45	0,96	0,0001	0,96	0,98	0,49
pDAS5227	Promedio (n=15)	6,09	0,03	3,63	0,32	1,32	75,43	12,32	3,73
pDAS5227	DS/DE (n=15)	0,35	0,00	0,29	0,09	0,12	1,69	1,36	0,37
pDAS5227	Valor de p	0,07	0,002	0,69	0,012	<0,0001	0,84	0,48	0,1
Nex710	Promedio (n=5)	6,38	0,04	3,59	0,26	1,73	75,28	12,76	3,40
Nex710	DS/DE (n=5)	0,61	0,01	0,15	0,03	0,49	3,00	2,76	0,43

\* Los números entre paréntesis indican el número de eventos incluidos en el análisis.

\*\*Un valor de p igual o menor que 0,05 se consideraba estadísticamente significativo y se calculó con el programa informático JMP.

Tabla 16B: Perfil de FA en semillas T1 de múltiples eventos transgénicos de estructuras artificiales

Nº de estructura artificial	Tipo de análisis	Cargo total	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1

Nº de estructura artificial	Tipo de análisis	Cargo total	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
pDAS5212	Media (n=37)*	6,00	0,54	1,33	0,06	0,30	0,13	0,08
pDAS5212	DS/DE (n=37)	0,21	0,04	0,11	0,01	0,03	0,05	0,04
pDAS5212	Valor de p**	0,01	0,03	0,003	0,14	0,76	0,26	0,23
pDAS5204	Promedio (n=20)	6,26	0,59	1,20	0,05	0,32	0,13	0,11
pDAS5204	DS/DE (n=20)	0,29	0,06	0,17	0,01	0,05	0,05	0,03
pDAS5204	Valor de p	0,43	0,79	0,20	0,87	0,68	0,29	0,63
pDAS5203	Promedio (n=8)	5,98	0,51	1,28	0,05	0,29	0,12	0,11
pDAS5203	DS/DE (n=8)	0,36	0,03	0,15	0,01	0,02	0,01	0,05
pDAS5203	Valor de p	0,02	0,006	0,048	0,42	0,56	0,55	0,84
pDAS5227	Promedio (n=15)	6,09	0,56	1,39	0,06	0,39	0,15	0,16
pDAS5227	DS/DE (n=15)	0,35	0,07	0,23	0,01	0,10	0,08	0,05
pDAS5227	Valor de p	0,07	0,21	0,0005	0,066	0,008	0,081	0,004
Nex710	Promedio (n=5)	6,38	0,60	1,10	0,05	0,31	0,11	0,10
Nex710	DS/DE (n=5)	0,61	0,13	0,09	0,01	0,07	0,05	0,08

\* Los números entre paréntesis indican el número de eventos incluidos en el análisis.

\*\*Un valor de p igual o menor que 0,05 se consideraba estadísticamente significativo y se calculó con el programa informático JMP.

5 Los perfiles de FAs específicos de los eventos pDAS5203, pDAS5212 y pDAS5227 daban como resultado una reducción del 21-25% en C18:0, en comparación con Nex710 (Fig. 15A). Todas esas diferencias eran estadísticamente significativas con  $p = 0,001$  o menor (Tabla 16). Otros FAs de cadena larga, C20:0, C22:0 y C24:0 también mostraron disminuciones en la concentración. Sin embargo, no se observaron cambios en el contenido de C16:0, lo que daba como resultado un aumento significativo del contenido en C16:0/C18:0 en los tres eventos transgénicos de estructuras artificiales (Fig. 165B).

10 El perfil de FA de semillas transgénicas pDAS5227 mostraba una clara diferencia en comparación con pDAS5203 y pDAS5212. Una reducción del 23% en el contenido de C14:0 ( $p = 0,002$ ) y un aumento del 19% en C16:1 ( $p = 0,01$ ) se observaron en semillas pDAS5227, en comparación con Nex710 (Fig. 16A y B, Tabla 16). El diseño del ZFP TF pDAS5227 es idéntico al de pDAS5203 y pDAS5212, excepto por la presencia de un dominio de regulación a la baja ERF3 fusionado a ZFP en lugar de un dominio KRAB1 (Tabla 15).

15 Un diseño de ZFP diferente, 13714, presente en pDAS5204 no era tan eficaz en la reducción de C18:0, en comparación con el diseño 13722 (Tabla 16). A diferencia del diseño 13722, que se une tanto al gen *FatB4* como a *FatB5*, el diseño 13714 se une solo al gen *FatB5* (Ejemplo 9, Tabla 10B).

### 15.2 Identificación de plantas para el análisis del ARNm

Dos eventos transgénicos, 5212[1]-004 y 5227[4]-012, que se produjeron mediante una transformación con las estructuras artificiales pDAS5212 y pDAS5227, respectivamente, se seleccionaron para un análisis de regulación a la

5 baja transcripcional de *FatB5*. Se plantaron de cincuenta a cien semillas  $T_1$  en el invernadero y se cultivaron las plantas. Se recogieron cuatro perforaciones de hojas de las plantas en la etapa de 2-3 hojas. Este material de la planta se analizó para determinar el número de copias de ZFP TF usando un ensayo qRT-PCR de *pat* (Ejemplo 6.2). A continuación se seleccionaron las plantas aleatorias nulas para ZFP y positivas para ZFP para un avance a la madurez para recoger las semillas  $T_2$  (Tabla 17). Vainas inmaduras de las mismas plantas se recogieron 25 días después de la floración (DAF) para un análisis del ARNm de *FatB*. Ambos eventos segregaban como una sola copia. Estos eventos se marcaron como 5212-4 y 5227-12.

Tabla 17: Escrutinio de eventos transgénicos para la identificación de plantas positivas y nulas para ZFP

Nombre del evento	Semillas/plantas T1	Nº de plantas	Nº de copias	Nº de plantas en T2	
5212-004	semillas plantadas	50			
	semillas germinadas	49			
	Escrutinio del número de copias de <i>pat</i>		17	cero	5
			26	1 copia	3
			5	2 copias	4
			1	NT*	
	Prueba Chi	0,04			
5227-12	semillas plantadas	100			
	semillas germinadas	96			
	Escrutinio del número de copias de <i>pat</i>		19	cero	4
			32	1 copia	3
			27	2 copias	4
			18	NT *	
	Prueba Chi	0,034			

\* NT = no ensayado.

10

### 15.3 Extracción de ARN y desarrollo del ensayo qRT-PCR

15 Se aisló ARN a partir de semillas inmaduras de *Brassica* con el kit para plantas RNAeasy® de Qiagen (Valencia, CA). Las semillas se colocaron en tubos de racimo que contenían tampón RLT (Qiagen)/β-mercaptoetanol (BME) y un lecho de acero inoxidable. Las muestras se pusieron en una bandeja de tubos en racimo y se homogeneizaron a 500 golpes por minuto durante 5 minutos en un molino de perlas (Kleco, Visalia, CA). La bandeja de tubos se giró 180 grados y se homogeneizó durante 5 minutos adicionales. Las muestras se transfirieron a continuación a tubos de 1,5 ml Eppendorf y se centrifugaron durante 5 minutos a 20.000 x g y el material sobrenadante se transfirió a un tubo nuevo. El ARN se aisló como se describe en el protocolo de los fabricantes (Qiagen, Valencia, CA). El ARN se eluyó de las columnas con 50 uL de agua exenta de ARNasa gua y las muestras de ARN se cuantificaron con un NanoDrop 8000 de Thermo Scientific (Wilmington, DE).

20

De diez a veinte microgramos de ARN se trataron con ADNasa TURBO exenta de ADN (número de catálogo AM1907; Applied Biosystems, Foster City, CA) en tubos de 1,5 ml exentos de ARNasa/ADNasa para eliminar el ADN genómico. Para cada muestra de ARN se preparó una reacción de cincuenta microlitros que contenía ARN, 1X tampón de ADNasa y 2 unidades de ADNasa TURBO. Las reacciones se incubaron a 37°C durante 30 minutos. Se

añadieron cinco microlitros (5 µl) de reactivo de inactivación de ADNasa a cada tubo y se mezclaron. Las muestras se incubaron a temperatura ambiente durante tres minutos, se mezclaron de nuevo y se incubaron a temperatura ambiente durante dos minutos adicionales. Las muestras se centrifugaron a 20.000 x g durante cinco minutos. Cuarenta microlitros (40 µl) del material sobrenadante se eliminaron, dejando detrás la suspensión de inactivación de ADNasa. Cada muestra tratada con ADNasa se cuantificó con Nanodrop, de modo que se podían utilizar cantidades equivalentes de ARN para la síntesis de ADNc.

El ADN complementario (ADNc) se preparó para cada muestra de ARN utilizando el kit de ADNc High Capacity (Applied Biosystems, Foster City, CA). Brevemente, se utilizaron 1,5 µg de ARN como molde en una reacción de cincuenta microlitros que contenía 1X tampón de transcripción inversa, 1x dNTPs, 1X oligómeros aleatorios y 125 unidades de transcriptasa inversa Multiscribe. Además, se prepararon reacciones de control noRT para cada una, usando 1,5 µg de ARN tratado con ADNasa en las mismas condiciones que anteriormente sin la transcriptasa inversa Multiscribe. Todas las muestras se incubaron a 25°C durante 10 minutos, 37°C durante 2 horas, 85°C durante 5 minutos y durante la noche a 4°C.

Los cebadores y las sondas se diseñaron con el programa informático Primer Express 3 de Applied Biosystems (Foster City, CA) para FATB, KASII, ERF3, KRAB1 y tubulina. Los cebadores se sintetizaron mediante Integrated DNA Technologies (Coralville, IA). A continuación se muestran las secuencias de los cebadores:

SEQ ID NO: 63: FATB5 directo - 5' TCGCTGCCATAACAACCATTT 3';

SEQ ID NO: 64: FATB5 inverso - 5' CGCCTGGGTTTCCAGTCA 3';

SEQ ID NO: 65: KRAB1 directo - 5' AAGGATGTGTTTCGTGGATTTC 3';

SEQ ID NO: 66: KRAB1 inverso - 5' CACAATCTGCTGTGCAGTATCAAG 3';

SEQ ID NO: 67: ERF3 directo - 5' GCGGGCGGGAGTTGTTA 3';

SEQ ID NO: 68: ERF3 inverso - 5' CCCCATCGGCGTTACATG 3';

SEQ ID NO: 69: TUBULINA directo -5' GAAGCTGAGAACAGCGATTGC 3';

SEQ ID NO: 70: TUBULINA inverso - 5' GTTCCTCCTCCCAACGAATG 3'.

Las sondas se sintetizaron con 6FAM/MGB mediante Applied Biosystems (Foster City, CA):

SEQ ID NO: 71: sonda de FATB5 6FAM TTTCTCAGCCGCCA;

SEQ ID NO: 72: sonda de KRAB1 6FAM TAGGGAAGAGTGGAAGCT;

SEQ ID NO: 73: sonda de ERF3 6FAM CAGGCCTCAGCCTT; y

SEQ ID NO: 74: sonda de TUBULINA 6FAM TACAAGGTTTCCAAGTTT

La expresión génica de *FatB*, *ERF3*, *KRAB1* y *tubulina* se evaluó con PCR en tiempo real con un aparato Applied Biosystem 7900HT (Foster City, CA). Una curva estándar se preparó utilizando partes alícuotas de varias muestras de ADNc dentro de los diferentes grupos para establecer el rango dinámico y la eficiencia del ensayo. Se realizaron después diluciones en serie 1:5 y 1:4 de la dilución inicial. Cada muestra de ADNc se diluyó 1:50 con Tris 10 mM, pH 7,5. Las reacciones de PCR se prepararon en un volumen de veinte microlitros que contenía 1X la mezcla maestra Gene Expression Master Mix (Applied Biosystems; Foster City, CA), 0,8 µM de cebadores directos e inversos, 0,2 µM de la sonda 6FAM/MGB específica del gen y 4 microlitros de ADNc diluido. Todas las reacciones se realizaron con las siguientes condiciones en el Applied Biosystem 7900HT con la versión 2.4 del programa informático SDS: Etapa 1) 50°C durante 2 minutos; Etapa 2) 95°C durante 10 minutos; Etapa 3) a 95°C durante 15 segundos y; Etapa 4) a 60°C durante 1 minuto. Las etapas 3 y 4 se repitieron durante 39 ciclos adicionales. La expresión génica de *tubulina* se evaluó en cada placa mientras que *FatB*, *ERF3* y *KRAB1* se ejecutaron por separado. Todas las reacciones se llevaron a cabo por triplicado.

Las curvas de calibración se evaluaron para la regresión y la pendiente y los valores umbrales se ajustaron según fue necesario. Las muestras noRT (controles negativos para la transcriptasa inversa que contenían ARN con todos los componentes de la reacción de ADNc, excepto la transcriptasa inversa) se compararon con las muestras correspondientes para asegurar que una diferencia de Ct de 5 a 7 estaba presente. Los datos se importaron a Excel para su análisis. El valor promedio de la cantidad de *FatB*, *ERF3* y *KRAB1* se normalizó a la cantidad promedio de *tubulina* y se indicó como las relaciones de la expresión entre *FatB* y *tubulina*.

#### 15.4 Análisis de la regulación a la baja del ARNm de *FatB5* específica de semillas en semillas inmaduras

Los resultados mostraron que la regulación a la baja del ARNm de *FatB*, 25 días después de la floración (DAF) en semillas inmaduras, dependía del estado de la cigosidad de las plantas T<sub>1</sub>. Por ejemplo, no se observó una regula-

ción a la baja aparente del ARNm de *FatB5* en las plantas heterocigotas (1 copia) de semillas inmaduras 5227-12 (Fig. 17). Sin embargo, cuando la expresión de ZFP TF aumentaba a 2 veces en las plantas homocigotas (2 copias), se observó una reducción del 21% en la expresión del ARNm de *FatB5* ( $p = 0,025$ ). Se obtuvieron resultados similares en otro evento, 5212-4, de una estructura artificial diferente, pDAS5212. En este evento, de nuevo, no se observó una regulación a la baja aparente de *FatB5* en las plantas heterocigotas (Fig. 18). Sin embargo, una reducción del 15% en los transcritos de *FatB5* se observó en las plantas homocigotas cuando la expresión de ZFP TF aumentaba.

Las plantas nulas, heterocigotas y homocigotas de ambos eventos avanzaron hasta la madurez y el perfil de los FAs se determinó en una agrupación de 24 semillas de cada planta. Este proceso de separar por segregación los linajes nulos de los linajes que contenían ZFP TF, permitía diferencias en el perfil de los FAs, lo que reflejaba más de cerca la presencia del ZFP TF.

15.4 Análisis del perfil de los FAs en semillas maduras T<sub>2</sub>

Los perfiles de los FAs en semillas maduras T<sub>2</sub> variaban dependiendo de la cigosidad de ZFP TF de la planta madre T<sub>1</sub> para el evento 5227-12 (Tablas 18A y 18B). Aunque las plantas T<sub>1</sub> heterocigotas no mostraron ninguna regulación a la baja aparente del ARNm de *FatB* en 25 semillas inmaduras DAF, su perfil de los FAs en las semillas mostraba una reducción del 8% de contenido en C18:0. No se observó ninguna reducción significativa del contenido en C16:0. Además, se observaron aumentos menores en C18:1 (oleico) y los posteriores FAs corriente abajo, dando como resultado una reducción del 5,5% en total de los FAs totales saturados (Tablas 18A y 18B, Fig. 19). Cuando las plantas T<sub>1</sub> 5227-12 eran homocigotas para ZFP TF, el ARNm de *FatB* se reguló a la baja de manera significativa. El perfil de los FAs de esas semillas T<sub>2</sub> era diferente. Un aumento del 11% en el contenido de C16:1 se observó, produciendo una reducción del 2,7% en los FAs totales saturados (Tablas 18A y 18B). Sin embargo, no se observó ninguna reducción adicional de C18:0, en comparación con las plantas heterocigotas. Estos cambios estaban de acuerdo con el perfil de los FAs de semillas maduras T<sub>1</sub> (Ejemplo 15.1).

Tabla 18A: Perfil de los FAs indicados de semillas T2 maduras del evento 5227-12

ID de la planta	Nº de copias	Total Sats	C14:0	C16:0	C16:1	C18:0	C18:1 (Oleico)	C18:1 (vac-cénico)	Total C18:1	C18:2	C18:3
Media de 5227-12 (n=4)	0	7,21	0,05	4,00	0,28	2,06	71,47	5,27	76,74	10,92	3,22
DS/DE de 5227-12 (n=4)	0	0,21	0,01	0,09	0,02	0,10	0,56	0,29	0,62	0,29	0,34
Media de 5227-12 (n=3)	1	6,84	0,05	3,82	0,25	1,88	71,99	4,88	76,87	11,03	3,33
DS/DE de 5227-12 (n=3)	1	0,33	0,01	0,33	0,03	0,04	0,50	0,19	0,32	0,05	0,31
Media de 5227-12 (n=4)	2	7,00	0,05	3,96	0,30	1,91	70,99	5,09	76,08	11,62	3,30
DS/DE de 5227-12 (n=4)	2	0,29	0,00	0,47	0,03	0,40	1,50	0,25	1,44	1,21	0,26

Tabla 18B: Perfil de los FAs de semillas T2 maduras del evento 5227-12

ID de la planta	Nº de copias	Total Sats	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
Media de 5227-12 (n=4)	0	7,21	0,64	1,01	0,04	0,27	0,13	0,09
DS/DE de 5227-12 (n=4)	0	0,21	0,03	0,02	0,01	0,01	0,03	0,01
Media de 5227-12 (n=3)	1	6,84	0,62	1,07	0,04	0,28	0,12	0,09
DS/DE de 5227-12 (n=3)	1	0,33	0,02	0,11	0,00	0,03	0,02	0,01
Media de 5227-12 (n=4)	2	7,00	0,62	1,06	0,04	0,27	0,12	0,08
DS/DE de 5227-12 (n=4)	2	0,29	0,09	0,03	0,01	0,03	0,03	0,01

5 Se obtuvieron resultados similares en los perfiles de los FAs de semillas maduras T<sub>2</sub> del evento 5212-4. En las semillas de plantas heterocigotas, una reducción del 5,5% en el contenido de C18:0 se observó como resultado de una reducción de los FAs totales saturados del 3,2% (Tabla 18, Fig. 20). También se observó un aumento menor de C18:1; de lo contrario, las concentraciones de los FAs restantes mostraron disminuciones menores. Plantas homocigotas no mostraron una mayor reducción en el contenido de C18:0, en comparación con la de las plantas hemicigotas, a pesar de una regulación a la baja significativa del ARNm de *FatB*.

10 En resumen, la regulación transcripcional a la baja del gen *FatB5* con expresión específica de la semilla de ZFP TFs se ha demostrado en estos ejemplos. Los ZFP TFs eran eficaces en conferir una regulación transcripcional a la baja del gen *FatB5* diana en semillas homocigotas T2. Sin embargo, la reducción de C18:0 en semillas homocigotas T2 no era más que la de las semillas T2 segregantes. Estos resultados indican que las copias homocigotas de ZFP TF pueden estar desencadenando un mecanismo de retroalimentación para ajustar el perfil de los FAs, ya que los genes *FatB* son indispensables para el crecimiento y desarrollo de las plantas (Bonaventure et al., *The Plant Cell*, 2003 15: 1020-1033). Es de destacar que la regulación a la baja del ARNm de *FatB* mediada por ZFP TF en semillas 15 inmaduras, está inversamente relacionada con la regulación al alza de *FatB* para los cambios en el contenido de C18:0.

20 Estos ejemplos muestran que los efectos dirigidos de ZFP TF sobre los genes *FatB* naturales, alteraban específicamente la cantidad de ARNm de *FatB*, produciendo de este modo cambios específicos y heredables en el perfil de los FAs en *B. napus*.

Tabla 19A: Perfil de los FAs de semillas T2 del evento 5212-4 con diferentes niveles de cigosidad

Nº de planta	Nº de copias	% de Sats total	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
Media de 5212-4 (n=5)	0	7,02	0,05	3,94	0,32	1,88	76,03	11,17	3,50
DS/DE de 5212-4 (n=5)	0	0,29	0,01	0,18	0,02	0,12	1,44	1,06	0,56
Media de 5212-4 (n=5)	1	6,79	0,04	3,83	0,32	1,81	76,86	10,74	3,36
DS/DE de 5212-4 (n=5)	1	0,19	0,00	0,07	0,01	0,17	0,26	0,31	0,11
Media de 5212-4 (n=5)	2	6,96	0,04	3,89	0,32	1,89	76,38	11,00	3,38
DS/DE de 5212-4 (n=5)	2	0,24	0,00	0,08	0,01	0,16	0,81	0,78	0,30

## ES 2 657 067 T3

Tabla 19B: Perfil de los FAs de semillas T2 del evento 5212-4 con diferentes niveles de cigosidad

Nº de planta	Nº de copias	% de Sats total	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
Media de 5212-4 (n=5)	0	7,02	0,64	1,16	0,04	0,33	0,20	0,11
DS/DE de 5212-4 (n=5)	0	0,29	0,07	0,06	0,01	0,04	0,05	0,01
Media de 5212-4 (n=5)	1	6,79	0,62	1,11	0,04	0,31	0,18	0,10
DS/DE de 5212-4 (n=5)	1	0,19	0,05	0,03	0,00	0,03	0,02	0,00
Media de 5212-4 (n=5)	2	6,96	0,63	1,12	0,05	0,31	0,21	0,10
DS/DE de 5212-4 (n=5)	2	0,24	0,07	0,05	0,01	0,03	0,05	0,01

### Lista de secuencias

<110> DOW AGROSCIENCES LLC et al.

5 <120> PROTEÍNAS CON DEDOS DE ZINC MODIFICADAS GENÉTICAMENTE QUE SE DIRIGEN A GENES VEGETALES IMPLICADOS EN LA BIOSÍNTESIS DE ÁCIDOS GRASOS

<130> 8325-4006.40

<140> PCT/US2010/002817

<141> 2010-10-22

10 <150> 61/279,528

<151> 2009-10-22

<160> 96

<170> PatentIn version 3.5

<210> 1

15 <211> 38

<212> DNA

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador sintético

20 <400> 1

cgcgatccg aacactgctg ttgctggctt tgatgaaa 38

<210> 2

<211> 35

<212> DNA

25 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 2

ctcgagctgc tactgctagt tccggtggag gagcc 35

30 <210> 3

<211> 736

<212> DNA

<213> Brassica napus

<400> 3

ES 2 657 067 T3

gaattcgccc ttcgoggatc cgaacactgc gtttgctggc tttgatgaaa atcttttgac 60  
 gtctccacgc acgagtttac gcatcccaga cgcctcctta gagagaagag agagatcgag 120  
 atcgatttcc tccctcttaa acctctctct ctctcgtgaa tctcatttcc ctttgccgct 180  
 agattctctc ttcacccttt tctcgccctt ccttcctctc ctccattactt ttttgctcgc 240  
 ttctgctctc tctctctctc agcactcttc gccatggtgg gtgctgctgc gtcttcctgt 300  
 tacgcatctc cgtatgacac ctggttcgtc gccgcctgca tgtccgtctc ccacggcggc 360  
 ggagattccc gacaagccgt cgctctcaaa tctaccgggc ggagtcgctc aagcagacaa 420  
 cagctcacca aatgctctgg atccggtagc agcactactt cctttgggcc ttgcaatcac 480  
 tacaatgctc tgtcttctct cttcgatcgc aactctgttt ctctcaatcg aaaccagagg 540  
 aggttgactc gagctgctac tgctagtcc ggtggaggag ccatggctgt tgcgatggat 600  
 atggaaaagg aagccaaggt tgacaacaaa cctcctacgg agcagcgcgg ggttggtgtg 660  
 acaggcatgg gagttgaaac atcactaggt catgaccctg acaccttcta tgagaatctc 720  
 ctacaagggc gaattc 736

5 <210> 4  
 <211> 732  
 <212> DNA  
 <213> Brassica napus

<400> 4  
 gaattcgccc ttcgoggatc cgaacactgc gtttgctggc tttgatgaaa atgtatgtat 60  
 cttttgacgt ctccacgtac gagtttacgc atccagacgc ctcgtagag agaagagaga 120  
 gatcgagatc gagatcgaga togatttcct ccctctctct ctcgtgaatc tcatttcccc 180  
 ttaccgctag attctctctt cacccttttc tcgcccttcc ttctcctcat tatttttttg 240  
 tcgtctctct ctctctcaca gcactcttct ctttagctat ggtgggtggt gctgctctt 300  
 cttcctgtta cgcattctcc ctatgcacct ggttcgtcgc tgcttgcatg tccgtctccc 360  
 acggcggcgg agattcccga caagccgtct ccctcaaatc taccgggcgg agtcgtcgaa 420  
 gcagacgaca gctcaccaaa tgcttggtc tttctggatc cggtagcgtt caggaggctc 480  
 tcgtcaactac ttcccttggg ccttgcaatc actacaatgc cttgtcttct ctcttcggat 540  
 cgaactctgt ttctctcaat cgaaccaga ggaggtgaa tctggctgct gctagtccg 600  
 gtggaggagc catggctggt gcgatggata tgcaaaagga agccaaggtt gacaacaaac 660  
 cccctacgga gcagcgcctg gttgtggtga caggcatggg agttgaaaca tcactaggtc 720  
 aaggcgaat tc 732

10 <210> 5  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 5  
 Arg Ser Asp Asn Leu Ser Val  
 1 5

<210> 6

<211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 6  
 Gln Lys Ile Asn Leu Gln Val  
 1 5  
  
 <210> 7  
 <211> 7  
 10 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 7  
 Arg Ser Asp Thr Leu Ser Glu  
 15 1 5  
  
 <210> 8  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 8  
 Thr Arg Ser Ser Arg Ile Asn  
 1 5  
  
 <210> 9  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 30 <400> 9  
 Arg Ser Asp Ala Leu Ala Arg  
 1 5  
  
 <210> 10  
 <211> 7  
 <212> PRT  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 10  
 Arg Ser Asp His Leu Ser Ala  
 1 5  
  
 40 <210> 11  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 45 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 11  
 Thr Ser Ser Ser Arg Ile Asn  
 1 5

<210> 12  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

5 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 12  
**Arg Ser Asp Asn Leu Ala Arg**  
 1 5

10 <210> 13  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

15 <400> 13  
**Asp Arg Ser His Leu Ala Arg**  
 1 5

<210> 14  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 14  
**Arg Ser Asp Asn Leu Ser Glu**  
 1 5

25 <210> 15  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

30 <400> 15  
**Arg Asn Ala His Arg Thr Thr**  
 1 5

<210> 16  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 16  
**Gln Ser Gly Asn Leu Ala Arg**  
 1 5

40 <210> 17  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 17

**Arg Ser Asp His Leu Ser Glu**  
 1 5  
 <210> 18  
 <211> 7  
 <212> PRT  
 5 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
 <400> 18  
**Gln Lys Ala Asn Arg Thr Lys**  
 1 5  
 10 <210> 19  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Descripción de Secuencia Artificial: Péptido sintético  
 <400> 19  
**Arg Ser Asp Asp Leu Thr Arg**  
 1 5  
 <210> 20  
 <211> 7  
 20 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
 <400> 20  
**Thr Ser Ala Asn Leu Ser Arg**  
 25 1 5  
 <210> 21  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
 <400> 21  
**Arg Ser Asp Asp Leu Ser Lys**  
 1 5  
 <210> 22  
 <211> 7  
 35 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
 <400> 22  
**Arg Ser Ala Asn Leu Thr Arg**  
 40 1 5  
 <210> 23  
 <211> 7  
 <212> PRT  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 23  
**Asp Lys Ser Asn Arg Lys Lys**  
 1 5

<210> 24  
 <211> 18  
 5 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético

<400> 24  
 10 cgtggagacg tcaaaga 18

<210> 25  
 <211> 20  
 <212> DNA  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético

<400> 25  
 aaggaagggc gagaaaagg 20

20 <210> 26  
 <211> 17  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético

25 <400> 26  
 agatgcgtaa caggaag 17

<210> 27  
 <211> 17  
 <212> DNA  
 30 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético

<400> 27  
 ctaccggcg gagtcgt 17

35 <210> 28  
 <211> 20  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 40 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 28  
 ttgactcgag ctgctactgc 20

<210> 29  
 <211> 20  
 45 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 29  
 50 ttccatattc catcgcaaca 20

ES 2 657 067 T3

<210> 30  
 <211> 447  
 <212> DNA  
 <213> Brassica napus

5 <400> 30  
 actcagcaact cttcgccatg gtgggtgctg ctgctcttc ctgttacgca tctccgctat 60  
 gcacctgggtt cgtcgccgcc tgcattgctg tctcccacgg cggcggagat tcccgacaag 120  
 ccgtcgtctt caaatctacc gggcggagtc gtcgaagcag acaacagctc accaaatgct 180  
 ctggatccgg tagcagcaact acttcctttg ggccttgcaa tcaactacaat gccttgtctt 240  
 ctctcttogg atcgaactct gtttctctca atcgaaacca gaggagggtg actcgagctg 300  
 ctactgctag ttcgggtgga ggagccatgg ctgttgcat ggatatgaa aaggaagcca 360  
 aggttgacaa caaacctcct acggagcagc gccgggttgt tgtgacaggc atgggagttg 420  
 aaacatcact aggtcaagg cgaattc 447

<210> 31  
 <211> 19  
 <212> DNA  
 10 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 31  
 acagcgattg cctacaagg 19

15 <210> 32  
 <211> 22  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 20 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 32  
 agatggttaa gatcaccaa gg 22

25 <210> 33  
 <211> 24  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

30 <400> 33  
 aaggaagagg aaggagtcta acag 24

<210> 34  
 <211> 18  
 <212> DNA  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 34  
 ctctgctct ccaccgta 18

40 <210> 35  
 <211> 24  
 <212> DNA

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 35

5 ggtcaacgga tcaggatatt ctg 24

<210> 36

<211> 20

<212> DNA

<213> Secuencia Artificial

10 <220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 36

ccatgttggc aaagcaacc 20

<210> 37

<211> 29

<212> DNA

<213> Secuencia Artificial

15 <220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 37

20 acaagagtgg attgatgac tagagaggt 29

<210> 38

<211> 29

<212> DNA

<213> Secuencia Artificial

25 <220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 38

cttggatgcc tatgtgacac gtaaacagt 29

<210> 39

<211> 29

<212> DNA

<213> Secuencia Artificial

30 <220>

<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 39

ccagcgtaag caataccagc cacaacacc 29

<210> 40

<211> 23

<212> DNA

<213> Secuencia Artificial

40 <220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 40

45 cctctctacc accgtctcac atg 23

<210> 41

<211> 20

<212> DNA

<213> Secuencia Artificial

50 <220>

<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 41  
gatctggccg gactgttca 20

5 <210> 42  
<211> 25  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Sonda sintética

10 <400> 42  
cgctcctcag ctaccacctc aacca 25

<210> 43  
<211> 22  
<212> DNA  
15 <213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Cebador sintético

<400> 43  
tcgatcttga tatgttgga ga 22

20 <210> 44  
<211> 20  
<212> DNA  
<213> Secuencia Artificial

<220>  
25 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 44  
aggtgcagaa tcatgtggtg 20

<210> 45  
<211> 16  
30 <212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 45  
35 ttgatgtga agatgt 16

<210> 46  
<211> 718  
<212> DNA  
<213> Brassica napus

40 <400> 46  
gaattcgccc ttcgoggatc cgaacactgc gtttgctggc tttgatgaaa gcttgtatgt 60  
atgtatcttt tgacgtctcc acgtacgagt ttacgcatcc agacgcctcg ttagagagaa 120  
gagrgagatc gagatcgaga tcgagatcga tttcctccct ctctctctcg tgaatctcat 180  
ttccccttac cgctagattc tctcttcacc cttttctcgc ccttccttct cctcattatt 240  
tttttgtcgt cttctgctct ctcwcakwc tcwkcwctct tmgyatggt gggtgstgct 300  
gcgtcttctt cctgttacgc atctccgcta tgcacctggc tcgtcgcygc ytgcatgtcc 360  
gtctcccacg gcggcggaga ttcccgacaa gccgtckcyc tcaaactctac cgggcggagt 420

ES 2 657 067 T3

cgctcgaagca gacaacagct caccaaagtc tctggatccg gtagcagcac tacttccttt 480  
 gggccttgca atcactacaa tgccttgtct tctctcttcg gatcgaactc tgtttctctc 540  
 aatcgaacc agaggagggtt gamtcgwgt gctrctgcta gttccggtgg aggagccatg 600  
 gctgttgca tggatagca aaaggaagcc aaggttgaca acaaaccctc tacggagcar 660  
 cgccgtgttg tggtgacagg catgggagtt gaaacatcac taggtcaagg gcgaattc 718  
 <210> 47  
 <211> 28  
 <212> DNA  
 5 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético  
 <400> 47  
 aacgaaagga gatcgagaga ggagagag 28  
 10 <210> 48  
 <211> 28  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético  
 <400> 48  
 cgaaaggag atcgagagag gcaccgca 28  
 <210> 49  
 <211> 28  
 20 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético  
 <400> 49  
 25 aaggagaact ttagggttg gggagact 28  
 <210> 50  
 <211> 28  
 <212> DNA  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético  
 <400> 50  
 aaggagaatt ttagggttg gggagact 28  
 <210> 51  
 <211> 28  
 35 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético  
 <400> 51  
 40 ctccgaagag attgcgtaa cacttcgt 28  
 <210> 52  
 <211> 28  
 <212> DNA  
 45 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Oligonucleótido sintético

<400> 52  
 ctccgaagag attggcgtaa ccttcatt        28

5    <210> 53  
      <211> 21  
      <212> DNA  
      <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 53  
 ctttgaacgc ttatcttct c        21

10   <210> 54  
      <211> 20  
      <212> DNA  
      <213> Secuencia Artificial

15   <220>  
      <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 54  
 ttccacaaca tctccccaag        20

20   <210> 55  
      <211> 14  
      <212> DNA  
      <213> Secuencia Artificial

25   <220>  
      <223> Descripción de Secuencia Artificial: Sonda sintética

<400> 55  
 ctcaggctcc accc        14

30   <210> 56  
      <211> 21  
      <212> DNA  
      <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

35   <400> 56  
 ctttgaagc tcattcttct c        21

40   <210> 57  
      <211> 20  
      <212> DNA  
      <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 57  
 ttccacaaca tctccccaag        20

45   <210> 58  
      <211> 16  
      <212> DNA  
      <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Sonda sintética

50   <220>  
      <223> Descripción de Secuencia Artificial: Sonda sintética

<400> 58  
 aaccttcac tcacca        16  
  
 <210> 59  
 <211> 22  
 5 <212> DNA  
    <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 <400> 59  
 10 tcgatcttga tatgttgga ga        22  
  
 <210> 60  
 <211> 20  
 <212> DNA  
 <213> Secuencia Artificial  
  
 15 <220>  
    <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 <400> 60  
 aggtgcagaa tcatgtggtg        20  
  
 <210> 61  
 20 <211> 21  
    <212> DNA  
    <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 25 <400> 61  
    acgagctacc tgacggacaa g        21  
  
 <210> 62  
 <211> 20  
 <212> DNA  
 30 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 <400> 62  
 gagcgacggc tgaagagta        20  
  
 35 <210> 63  
    <211> 21  
    <212> DNA  
    <213> Secuencia Artificial  
  
 <220>  
 40 <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 <400> 63  
 tcgctgcat aacaaccatt t        21  
  
 <210> 64  
 <211> 18  
 45 <212> DNA  
    <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético  
  
 <400> 64  
 50 cgctctgggtt tccagtca        18

<210> 65  
 <211> 22  
 <212> DNA  
 <213> Secuencia Artificial

5 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 65  
 aaggatgtgt tcgtggatt ca 22

10 <210> 66  
 <211> 24  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

15 <400> 66  
 cacaatctgc tgtgcagat caag 24

<210> 67  
 <211> 17  
 <212> DNA  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 67  
 gcgggcggga gttgta 17

25 <210> 68  
 <211> 18  
 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

30 <400> 68  
 ccccatcggc gttacatg 8

<210> 69  
 <211> 21  
 <212> DNA  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 69  
 gaagctgaga acagcgattg c 21

<210> 70  
 <211> 20  
 <212> DNA  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de Secuencia Artificial: Cebador sintético

<400> 70  
 gttcctctc ccaacgaatg 20

50 <210> 71  
 <211> 14  
 <212> DNA

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 71

5 tttctcagcc gccca 14

<210> 72

<211> 18

<212> DNA

<213> Secuencia Artificial

10 <220>

<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 72

tagggaagag tggaagct 18

<210> 73

<211> 14

<212> DNA

<213> Secuencia Artificial

15 <220>

<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 73

caggcctcag cctt 14

<210> 74

<211> 18

<212> DNA

<213> Secuencia Artificial

20 <220>

<223> Descripción de Secuencia Artificial: Sonda sintética

<400> 74

tacaagggtt ccaagttt 18

<210> 75

<211> 7

<212> PRT

<213> Secuencia Artificial

25 <220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 75

**Arg Ser Asp Asn Leu Ser Ala**

1 5

<210> 76

<211> 7

<212> PRT

<213> Secuencia Artificial

30 <220>

<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 76

**Gln Ser Ala His Arg Lys Thr**

1 5

<210> 77

<211> 7

<212> PRT

<213> Secuencia Artificial

35

40

45

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 77  
 Gln Ser Ser His Arg Lys Thr  
 1 5

5 <210> 78  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 78  
 Arg Ser Asp His Leu Ser Val  
 1 5

15 <210> 79  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 79  
 Gln Asn Ala His Arg Ile Glu  
 1 5

20 <210> 80  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 80  
 Arg Ser Asp His Leu Ser Lys  
 1 5

30 <210> 81  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

35 <400> 81  
 Gln Asn Ala Asn Arg Ile Thr  
 1 5

<210> 82  
 <211> 7  
 <212> PRT

40 <213> Secuencia Artificial

<220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 82  
 Arg Ser Asp His Leu Ser Thr  
 1 5

45 <210> 83  
 <211> 7  
 <212> PRT

<213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 83  
His Ser Asn Thr Arg Lys Asn  
5 1 5

<210> 84  
<211> 7  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 84  
Arg Ser Asp His Leu Ser Gln  
1 5

15 <210> 85  
<211> 7  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Péptido sintético

20 <400> 85  
Asn Ser Ala Ser Arg Lys Asn  
1 5

<210> 86  
<211> 7  
<212> PRT  
25 <213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 86  
Gln Ser Gly His Leu Ser Arg  
1 5

30 <210> 87  
<211> 7  
<212> PRT  
<213> Secuencia Artificial

<220>  
35 <223> Descripción de Secuencia Artificial: Péptido sintético

<400> 87  
Asn Ser Asp Ser Leu Thr Glu  
1 5

<210> 88  
<211> 7  
40 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de Secuencia Artificial: Péptido sintético

<400> 88  
Arg Arg Ala Asp Leu Ser Arg  
45 1 5

<210> 89

<211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 89  
**Arg Ser Asp Ser Leu Ser Ala**  
 1 5  
  
 <210> 90  
 <211> 7  
 10 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 90  
**Gln Asn Ala His Arg Lys Thr**  
 15 1 5  
  
 <210> 91  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
  
 20 <220>  
 <223> Descripción de Secuencia Artificial: Péptido sintético  
  
 <400> 91  
**Arg Asn Ala Asp Arg Ile Thr**  
 1 5  
  
 <210> 92  
 <211> 200  
 25 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de Secuencia Artificial: Polipéptido sintético  
  
 30 <220>  
 <221> caract\_misc  
 <222> (1)..(200)  
 <223> Esta región puede abarcar 5 a 200 restos "Gly"  
  
 <400> 92  
 35 Gly Gly



ES 2 657 067 T3

gcaacagtgg tattagccag attgagaatt ttgattgttc tgcttttcct acgagaattg 240  
ctggagagat caagtcattc tcgactgaag ggtgggttgc tccaaaactc tcaaagagga 300  
tggacaagtt catgctctat cttctcactg ctggcaagaa ggctttggtt gatggtgggg 360  
taaccgaaga agtcatggca gagtttgaca aagcceaagc cggagtcttg attggtctg 420  
caatgggagg catgaaggtc tttcaagatg ctattgaagc tatgaagatc tcttacaaga 480  
agatgaatcc tttctgtgtg cttctcgcga caacgaacat gggttctgct atgcttgctt 540  
tggatctggg atggatgggg ccaactatt ctatctcaac tgcttgtgca acaagcaact 600  
tttgattctt caattcagca aaccacatta tcaaaggaga agctgatgta atgctctgtg 660  
gtggctcgga ttcagttatt attccaatag ggttgggagg ttttgttgca tgccgtgctc 720  
tttctcaaag gaataatgat cccacaaaag cttcacgccc ttgggatagc aaccgagatg 780  
gtttcgtgat gggagagggg gctggagttt tgcttttggg agagcttgaa catgctaaga 840  
aaagaggagc aacaatctat gcagagttcc ttggtgggag tttcacatgt gatgcctatc 900  
acatgaccga gcctcgccct gatggtgctg gtgtgattct gtgtattgag agagcattgg 960  
ctgatgctgg gatttcctaaa gaacagataa actatataaa tgcacatgca acctctacac 1020  
cagctggaga ccttaaggag taccaagccc ttgctcactg ctttggccaa aatcctgaga 1080  
taaaagttaa ttccacaaaa tctatgattg gacacttgct gggagctgct ggggccgttg 1140  
aagctgtgoc aactgtgcag gccataagga ccgatgggt tcatccaaat atcaaccttg 1200  
agagtccaga caatggagtg gatacaaatt tgctgggtgg tcctgagaag gagagattgg 1260  
acattaaagc agccttgtca aattcattcg ggttgggtgg ccacaactcc agcatcattt 1320  
ttgctcctta caagtgaag cactcattgc ctgtactcca aacctggttg tgtaacttgc 1380  
tgtaagtgtt tacaagaact ccccatgtta tgttgttgcg ggaatcaaca cagtttgta 1440  
aactaccaag agctaagcta agtttcctta ggatcaagat ccgtttgtgc cagagaactt 1500  
ggacaaagag caaacgtagc agagtttggg ttagcttcc gtgtgatacc ttttgagtgg 1560  
aatctttgta gccttttctt tttttagtg tttcatttct atttgttaat cattacaatc 1620  
tgaaaattgc caaaccaatt ctccgttaaa ttagtaact ctacacaaaa aaaaaaaaaa 1680  
aaaaaaaaaa aaa 1693

<210> 94

<211> 1767

<212> DNA

5 <213> Arabidopsis thaliana

<400> 94

attctctott cttctcttca cccatttctc gctttctcct ttgttctctc atctgggttc 60

ttctcaaagc ctcttccttt ttttgccatg gtgggtgcgt cttcctctta cgcctctccg 120

ES 2 657 067 T3

ttatgtacct ggtttgttgc tgcttgcacg tccgtctctc acggtggagg agatagccgt 180  
 caggctgttg ctcttcaatc tgggtggcgg agtcggcgaa ggaggcagct tagcaaatgc 240  
 tctgtcgctt ctggatccgc tagcattcag gctctcgtca cttcttgttt ggattttggt 300  
 ccttgtactc actacaacaa caacaatgca ttgtcttctc tctttggatc gaatagtgtt 360  
 tctttgaatc gaaaccagag gagattgaat cgtgctgcta gctccggtgg agccatggca 420  
 gtgatggaga tggaaaagga agctgcbggt aacaagaaac cacctacgga gcagcgtcga 480  
 gttgtagtga caggcatggg agttgaaaca tcattgggtc atgaccaca taccttctat 540  
 gagaatttgc tacaaggcaa cagtggattt agccagattg aaaattttga ttgttctgaa 600  
 tttcctacgc gaattgcbgg agagatcaaa agcttctcga ctgaaggatg ggttgctcca 660  
 aaactttcta aaaggatgga caaattcatg ctctatcttc tcacagctgg taagaaagct 720  
 ttggctgatg gtggggttac tgatgaagta atggcagagt ttgacaaaac caaatgtgga 780  
 gttttgattg gctcggcaat gggaggaatg aaggtctttt acgatgctat tgaagctctg 840  
 agaatctctt acaagaagat gaatcctttt tgtgtacctt ttgcgacaac aaacatgggt 900  
 tctgctatgc ttgccatgga tctgggatgg atggggccaa actattctat ttcaactgct 960  
 tgtgccacaa gcaacttttg cattctgaat tcagcaaacc acattattaa aggtgaagct 1020  
 gatgtaatgc tctgtggtgg ctcagatgca gttattattc caatagggtt gggaggtttt 1080  
 gttgcatgcc gggctctttc acaaggaat aatgatocca caaaagcttc acgtccttgg 1140  
 gataccaatc gagatggttt cgtgatggga gagggagctg gagttctact tttggaagaa 1200  
 ctcgagcatg ctaagaaaag aggtgcaact atctacgcag agttcctcgg tgggagtttc 1260  
 acatgtgatg cctatcacat gaccgagcct caccctgatg gggctggtgt tattctctgt 1320  
 attgagagag cgtagctag tgctgggatt tccaaggaac aaataaatta cataaatgca 1380  
 catgcaacct caacgcatgc tggagatatt aaggaatacc aagcccttgc tcactgtttt 1440  
 ggccaaaatc ctgagcttaa ggtaaattcc acaaaatcta tgattggaca cttgctggga 1500  
 gctgctgggg ccgtggaggc tgttgcaact gtgcaggcga tacggaccgg atgggttcat 1560  
 ccaaatatca acctogagaa tccagacagt ggagtggata caaagctgct ggtgggtcct 1620  
 aagaaggaga gactggacat taaagcagcc ttgtcaaatt cattcgggtt tgggtgtcat 1680  
 aactccagca tcatttttgc tccttacaag tgaaagcgaa agcagttgct tgtactccaa 1740  
 acctgattgt ataacttgct gtaaggt 1767

<210> 95  
 <211> 893  
 <212> DNA  
 <213> Brassica rapa  
 <400> 95

5

ES 2 657 067 T3

gacatggata ttttctgaga tttggaagtg tggatttgat ataagcaaag caaagcaagt 60  
gctcgttgt atgtatcttt tgacgtctcc acgcacgagt ttacgcatcc agacgcctcg 120  
ttagagagaa gagagagatc gagatcgatt cctccctctt aaacctctct ctctctcgtg 180  
aatctcattt ccccttacgg ctagattctc tcttcaccct tttctcgccc ttccttcctt 240  
ctcctcatta cttttttgtc gtcttctgct ctctctctct ctcagcactc ttcgctttag 300  
ctatggtggg tgctgctgcg tcttctgct acgcatctcc gctatgcacc tggttcgtcg 360  
ctgcctgcat gtcctgctcc cacggcggcg gagattcccg acaagccgtc gctctcaaat 420  
ctagcgggcg gagtcgtcga agcagacaac aactcaccaa atgctctgga tccggtagca 480  
gcgttcagca ggctctcgtc actacttctt ttgggccttg caatcactac aatgccttgt 540  
cttctctctt cggatcgaac tctgtttctc tcaatcgaaa ccagaggagg ttgactcgag 600  
ctgctactgc tagttccggt aatttaacat ttcaccaatc tgggtttttg atttgtgggt 660  
attgtttgca gagttgacaa gctttgtttt tgtagtaaa gtttctcttt ttatgtgttt 720  
atgttatttg gtcaaatgat tcaacttggg gatcaagttt aacgggaaag tatgctgatt 780  
tggaatctca gcttaattgt caagtttcat tctttattgt gtgtgtaata tgcaaagcac 840  
atctctgtga gtttgattca atattggact ctagataagg agtatgatga tac 893

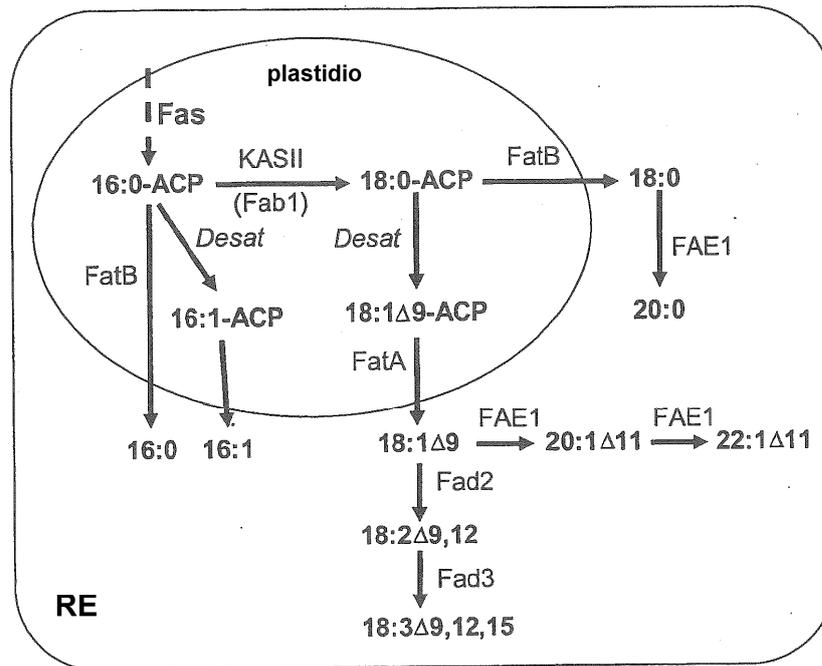
<210> 96  
<211> 812  
<212> DNA

5 <213> Brassica oleracea

<400> 96  
acggcacttt acgcaccaat atataccata tggtagccag ttaagggtag gaagtgattt 60  
tccccttttc ttatgtatac agactccaat tccccaaatt tactctggtc atcgaacatc 120  
gatacgtttg tatgtatgta cgtcttatct ctctctctgg ctatctctgt gtgttttggt 180  
gttaataaca ggaggttttg ttttgttttc tttgacgtct ccaogcacga ttttaagcct 240  
ccgttacgct ctgcacgcct cctgaaagag agagagagag agatcgaatc atcttaatta 300  
aacctctctc gtgaatccgc tagattctct cttcacggtt ttctcgcctc tcctctcctc 360  
tccattttgt tttcttcttc ttcgtcttct tcttcttgcc attgccatgg tgggtgctgc 420  
tgcgtcttcc tgttacgcat ctcccttatg caccttcttc gttgctgctt gcatgcctc 480  
ctctcacggc ggcggtgata cccgtcaagc ctttgggagg agccgctgaa ggagacaaca 540  
gctcggcaaa tgctctggat ccggtagcat tctcgtgtct tcttgtttgg agtttaagcc 600  
ttgcagtcac tacaacaaga acaacaagg caatgccttt cctctcctcg gatcgaatag 660  
cctttctctg aatcgaaagc agaggaaact gaatcgagca acttcttctt cttccggtat 720  
aaacattctt ctcccaaat ctcaagtctt ggtttctggg tatcgtttgc aagctgacaa 780  
gctttgtttt ttgtgtataa agtttctcct tt 812

**REIVINDICACIONES**

1. Una proteína de fusión que comprende un dominio funcional que es un dominio regulador de la transcripción o un dominio de escisión y una proteína con dedos de zinc de origen no natural que modula la expresión de al menos un gen endógeno vegetal de la tioesterasa B de ácidos grasos (FatB), en donde la proteína con dedos de zinc comprende las regiones de reconocimiento de hélices ordenadas como se muestra en una fila simple de la Tabla 10B.
2. La proteína de fusión según la reivindicación 1, en donde el gen vegetal está contenido dentro de una planta *Brassica*.
3. La proteína de fusión según la reivindicación 2, en donde la proteína con dedos de zinc se une a un sitio diana como se muestra en la Tabla 2.
4. Un polinucleótido que codifica una o varias proteínas con dedos de zinc o proteínas de fusión según cualquiera de las reivindicaciones 1 a 3.
5. Un método de modificación de la expresión de uno o varios genes implicados en la biosíntesis de ácidos grasos en una célula vegetal, comprendiendo el método:
  - introducir en la célula vegetal, uno o varios vectores de expresión que comprenden al menos un polinucleótido que codifica una proteína de fusión según cualquiera de las reivindicación 1 a 3, en condiciones tales que una o varias de las proteínas de fusión se expresan y la expresión de uno o varios genes FatB se modifica.
6. El método según la reivindicación 5, en el que la modificación comprende modular la expresión de al menos un gen FatB.
7. El método según la reivindicación 6, en el que se activa la expresión de al menos un gen FatB.
8. El método según la reivindicación 6, en el que se inhibe la expresión de al menos un gen FatB.
9. El método según la reivindicación 5, en el que el polinucleótido codifica una nucleasa de dedos de zinc y al menos un gen FatB se escinde.
10. El método según la reivindicación 9, que comprende adicionalmente la etapa de introducir un ácido nucleico donante de modo que el vector donante se introduce en el sitio de escisión mediante recombinación homóloga.
11. Una célula vegetal que comprende una o varias proteínas de fusión según cualquiera de las reivindicaciones 1 a 3 o uno o varios polinucleótidos según la reivindicación 4 y que comprende al menos un gen que ha sido modificado mediante cualquiera de los métodos según cualquiera de las reivindicaciones 5 a 10.
12. La célula vegetal según la reivindicación 11, en donde la célula está dentro de una semilla y el contenido en ácidos grasos en la semilla se modifica.
13. Una planta que comprende al menos una célula según cualquiera de las reivindicaciones 11 a 12.
14. La semilla o la progenie de una planta según la reivindicación 13, que comprende una o varias proteínas de fusión según cualquiera de las reivindicaciones 1 a 3 o uno o varios polinucleótidos según la reivindicación 4.
15. Una planta que comprende al menos una célula que comprende la proteína de fusión según las reivindicaciones 1 a 3 o un polinucleótido según la reivindicación 4.
16. La semilla o la progenie de una planta según la reivindicación 15, que comprende una o varias proteínas de fusión según cualquiera de las reivindicaciones 1 a 3 o uno o varios polinucleótidos según la reivindicación 4.



Ruta de biosíntesis de ácidos grasos

FIG. 1

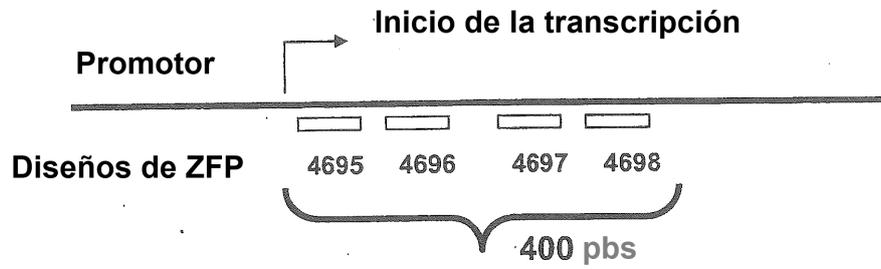


FIG. 2

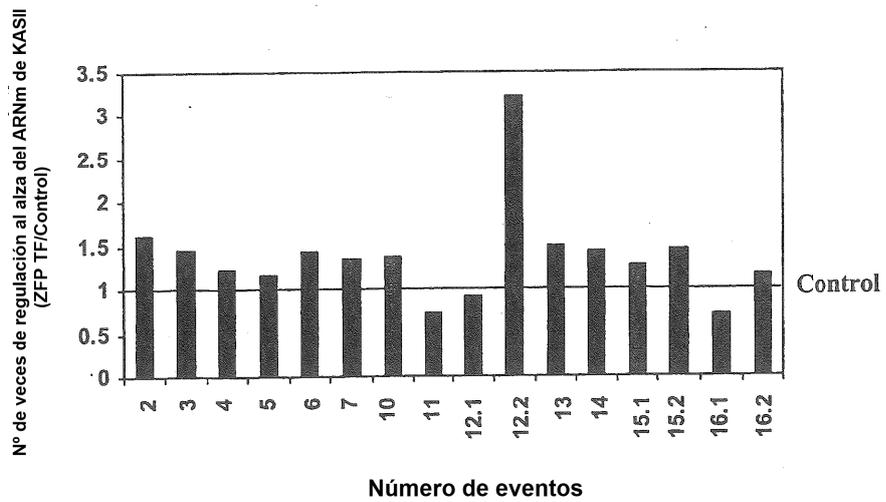


FIG. 3

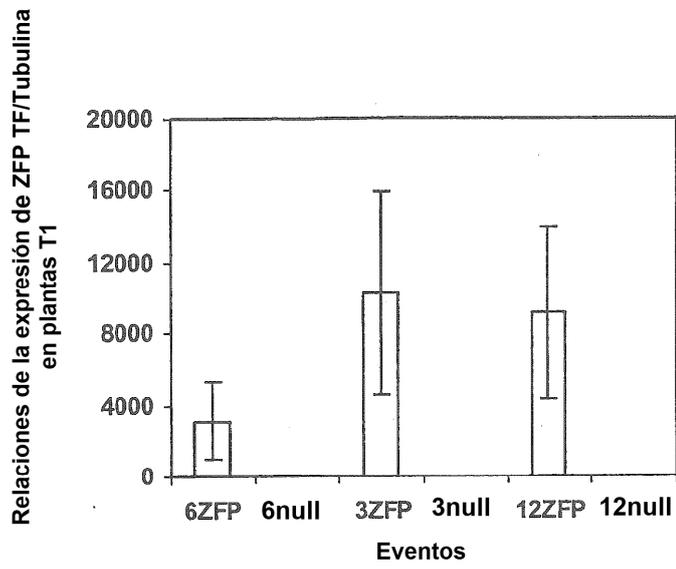


FIG. 4

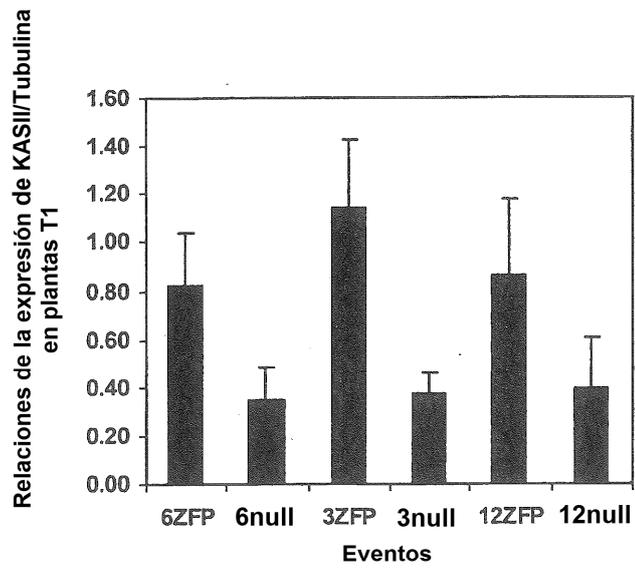


FIG. 5

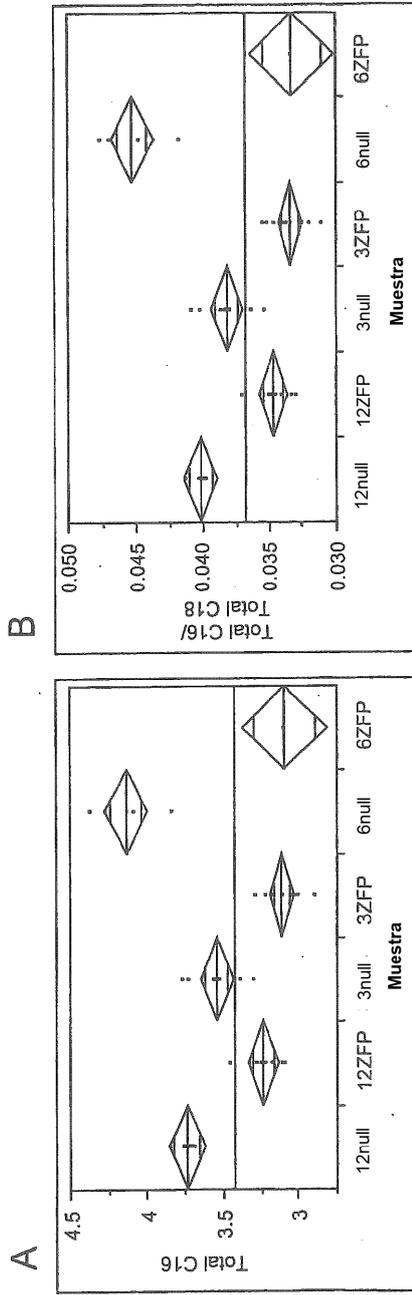


FIG. 6

ES 2 657 067 T3

		1		50
AF244520	(1)	-----		
AF318307	(1)	ATTCCTCTCTCTTCTCTTCA	CCATTCTCGCTTTCTCCTT	TGTTCTCTC
Consenso	(1)			
		51		100
AF244520	(1)	-----		
AF318307	(51)	ATCTGGGTTCTTCTCAAAGC	TCTTCTTTTGGCCATGGTGG	TGCGT
Consenso	(51)			
		101		150
AF244520	(1)	-----		
AF318307	(101)	CTTCCTCTTACGCATCTCCG	TATGTACCTGGTTTGTGTG	CTGCTGCATG
Consenso	(101)			
		151		200
AF244520	(1)	-----		
AF318307	(151)	TCCGTCTCTCACGGTGGAGG	AATAGCCGTCAGGCTGTGCT	TCTCAATC
Consenso	(151)			
		201		250
AF244520	(1)	-----		
AF318307	(201)	TGGTGGGCGGAGTCGGCGA	AGGAGGCAGCTTAGCAAATG	CTCTGTCGCTT
Consenso	(201)			
		251		300
AF244520	(1)	-----		
AF318307	(251)	CTGGATCCGCTAGCATTCA	GGCTCTCGTCACTTCTTGT	TGGATTTGGT
Consenso	(251)			
		301		350
AF244520	(1)	-----		
AF318307	(301)	CCTTGTA	CTCACTACAACAACAATGC	ATTGTCTTCTCTTTGGATC
Consenso	(301)			
		351		400
AF244520	(1)	-----		
AF318307	(351)	GAATAGTGTCTTCTTGAAT	CGAAACCAGAGAGATTGAAT	CGTCTGCTT
Consenso	(351)			
		401		450
AF244520	(18)	AGTTC	CGGTGGAGGAGCCCATGGCTGT	TGCATGGATTATGGAAAAGGAAGC
AF318307	(400)	AGCTCCGGTGGAGCCCATGGCAGT	---GATGGAGATGGAAAAGGAAGC	
Consenso	(401)	AG TCCGGTGGAGCCATGGC GT	GATGGAGATGGAAAAGGAAGC	
		451		500
AF244520	(68)	CAAGGTTGACAAACAACCT	CCTACGGAGCAGCGCCGGGTT	TGTGTGACAG
AF318307	(444)	TGCGGTTAACAAGAAACCA	CCTACGGAGCAGCGTCGAGTT	TGTAGTGACAG
Consenso	(451)	GGTT ACAA AAACC CCTACGGAGCAGCG	CG GTTGT GTGACAG	
		501		550
AF244520	(118)	GCATGGGAGTTGAAACATCA	CTAGGTCATGACCCGACACCTT	TATATGAG
AF318307	(494)	GCATGGGAGTTGAAACATCA	CTAGGTCATGACCCGACACCTT	TATATGAG
Consenso	(501)	GCATGGGAGTTGAAACATCA T	GGTCATGACCC A ACCTT TATGAG	
		551		600
AF244520	(168)	AATCTCTACAAGGCAACAGT	GGTATTAGCCAGATTGAAATTTT	GATTG
AF318307	(544)	AATCTCTACAAGGCAACAGT	GGTATTAGCCAGATTGAAATTTT	GATTG
Consenso	(551)	AAT T CTACAAGGCAACAGTGGTATTAGCCAGATTGA	AATTTTGATTG	
		651		700
AF244520	(218)	TTCTGCTTTCTCCTACCGA	GAATTGCTGGAGAGATCAA	GTCATTTCTCGACTG
AF318307	(594)	TTCTGCAATTTCTCCTACCGA	GAATTGCTGGAGAGATCAA	AAGCTTTCTCGACTG
Consenso	(601)	TTCTG TTTCTCCTACG GAATTGC	GGAGAGATCAA TTCTCGACTG	

FIG. 7A

AF244520	(268)	AAGGCTGGGTTGCTCCAAAAC	CTCAAAGAGGATGGACAA	ATTCATGCTC	
AF318307	(644)	AAGGATGGGTTGCTCCAAAAC	TTCTAAAGGATGGACAA	ATTCATGCTC	
Consenso	(651)	AAGG TGGGTTGCTCCAAAAC	TC AA AGGATGGACAA	TTCATGCTC	
		701		750	
AF244520	(318)	TATCTTCTCAC	TCTGG AAGAA GCTTTGG	TGATGGTGGGGT	AACCGA
AF318307	(694)	TATCTTCTCAC	TCTGG AAGAA GCTTTGG	TGATGGTGGGGT	TACTGA
Consenso	(701)	TATCTTCTCAC	GCTGG AAGAA GCTTTGG	TGATGGTGGGGT	AC GA
		751		800	
AF244520	(368)	AGAAGT	CATGGCAGAGTTTGACAAA	CCAAATG	GGAGTCTTGATTGGCT
AF318307	(744)	TGAAGT	AATGGCAGAGTTTGACAAA	CCAAATG	TGGAGTCTTGATTGGCT
Consenso	(751)	GAAGT	ATGGCAGAGTTTGACAAA	CCAAATG	GGAGT TTGATTGGCT
		801		850	
AF244520	(418)	CTGCAATGGGAGG	CATGAAGGTCTTT	CAAGATGCTATTGAAGCT	ATGAAG
AF318307	(794)	CGCAATGGGAGG	AATGAAGGTCTTT	CAAGATGCTATTGAAGCT	CTGAGA
Consenso	(801)	C GCAATGGGAGG	ATGAAGGTCTTT	A GATGCTATTGAAGCT	TGA
		851		900	
AF244520	(468)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	CCCTTT	GCGACAACGAA
AF318307	(844)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	ACCTTT	GCGACAACAAA
Consenso	(851)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	CCCTTT	GCGACAAC AA
		901		950	
AF244520	(518)	CATGGGTTCTGCTATGCTTGCT	TTGGATCTGGGATGGATGGGGCCAAACT		
AF318307	(894)	CATGGGTTCTGCTATGCTTGCT	CAATGGATCTGGGATGGATGGGGCCAAACT		
Consenso	(901)	CATGGGTTCTGCTATGCTTGCT	TGGATCTGGGATGGATGGGGCCAAACT		
		951		1000	
AF244520	(568)	ATTCTAT	TCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	CAATTCA
AF318307	(944)	ATTCTAT	TCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	CAATTCA
Consenso	(951)	ATTCTAT	TCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	AATTCA
		1001		1050	
AF244520	(618)	GCAAACCACATTAT	AAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC	
AF318307	(994)	GCAAACCACATTAT	AAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC	
Consenso	(1001)	GCAAACCACATTAT	AAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC	
		1051		1100	
AF244520	(668)	GGAT	CAGTTATTATTCCAATAGGGTTGGGAGGTTT	TGTTGCATGCCGG	IG
AF318307	(1044)	AGAT	CAGTTATTATTCCAATAGGGTTGGGAGGTTT	TGTTGCATGCCGG	GG
Consenso	(1051)	GAT	CAGTTATTATTCCAATAGGGTTGGGAGGTTT	TGTTGCATGCCG	G
		1101		1150	
AF244520	(718)	CTCTTTT	CAAAGGAATAATGATCCCACAAAAGCTTCACG	CCCTTGGGAT	
AF318307	(1094)	CTCTTTT	CAAAGGAATAATGATCCCACAAAAGCTTCACG	CCCTTGGGAT	
Consenso	(1101)	CTCTTTT	CAAAGGAATAATGATCCCACAAAAGCTTCACG	CCCTTGGGAT	
		1151		1200	
AF244520	(768)	ACCAAC	CGAGATGGTTTCGTGATGGGAGAGGGAGCTGGAGTT	CTCTTTT	
AF318307	(1144)	ACCAAT	CGAGATGGTTTCGTGATGGGAGAGGGAGCTGGAGTT	CTACTTTT	
Consenso	(1151)	A CAA	CGAGATGGTTTCGTGATGGGAGAGGGAGCTGGAGTT	T CTTTT	
		1201		1250	
AF244520	(818)	GGAAGAGCT	GAACATGCTAAGAAAAGAGG	GCAAC	ATCTATGCAGAGT
AF318307	(1194)	GGAAGACT	CGACATGCTAAGAAAAGAGG	GCAACT	ATCTACGCAGAGT
Consenso	(1201)	GGAAGA	CT GA CATGCTAAGAAAAGAGG	GCAAC	ATCTA GCAGAGT
		1251		1300	
AF244520	(868)	TCCT	GGTGGGAGTTTCACATGTGATGCCTATCACATGACCGAGCCT	CGC	
AF318307	(1244)	TCCT	GGTGGGAGTTTCACATGTGATGCCTATCACATGACCGAGCCT	CAC	
Consenso	(1251)	TCCT	GGTGGGAGTTTCACATGTGATGCCTATCACATGACCGAGCCTC	C	
		1301		1350	

FIG. 7B

ES 2 657 067 T3

AF244520	(918)	CCTGATGGGCTGGTGTATTCTGTGATTGAGAGAGCAATTGGCTGATGC	
AF318307	(1294)	CCTGATGGGCTGGTGTATTCTGTGATTGAGAGAGCGTTAGCTAGTGC	
Consenso	(1301)	CCTGATGG GCTGGTGT ATTCT TGTATTGAGAGAGC TT GCT TGC	1351 1400
AF244520	(968)	TGGGATTTCCAAAGAACAATAAACTAATAAATGCACATGCAACCTCA	
AF318307	(1344)	TGGGATTTCCAAAGAACAATAAACTAATAAATGCACATGCAACCTCAA	
Consenso	(1351)	TGGGATTTCCAA GAACA ATAAA TA ATAAATGCACATGCAACCTC A	1401 1450
AF244520	(1018)	CACCACTGGAGACCTTAAGGATACCAAGCCCTTGCTCACTGTTTGGC	
AF318307	(1394)	CCATGCTGGAGATTAAGGATACCAAGCCCTTGCTCACTGTTTGGC	
Consenso	(1401)	C C GCTGGAGA TTAAGGA TACCAAGCCCTTGCTCACTG TTTGGC	1451 1500
AF244520	(1068)	CAAAATCCTGAGATAAAGTAAATTCCACAAAATCTATGATTGGACACTT	
AF318307	(1444)	CAAAATCCTGAGATAAAGTAAATTCCACAAAATCTATGATTGGACACTT	
Consenso	(1451)	CAAAATCCTGAG T AA GT AATTCCACAAAATCTATGATTGGACACTT	1501 1550
AF244520	(1118)	GCTGGGAGCTGCTGGGGCCGTGAGGCTGTGCAACTGTGCAGGCATATA	
AF318307	(1494)	GCTGGGAGCTGCTGGGGCCGTGAGGCTGTGCAACTGTGCAGGCATATA	
Consenso	(1501)	GCTGGGAGCTGCTGGGGCCGT GA GCTGT GCAACTGTGCAGGC ATA	1551 1600
AF244520	(1168)	GGACCGGATGGGTTCAATCAAAATCAACCTGAGAGTCCAGACAAATGGA	
AF318307	(1544)	GGACCGGATGGGTTCAATCAAAATCAACCTGAGAGTCCAGACAAATGGA	
Consenso	(1551)	GGACCGGATGGGTTCAATCAAAATCAACCT GAGAATCCAGACAAATGGA	1601 1650
AF244520	(1218)	GTGGATACAAATTTGCTGGTGGTCCAGAGAAGGAGAGATTTGGACATTAA	
AF318307	(1594)	GTGGATACAAAGCTGCTGGTGGTCCAGAGAAGGAGAGATTTGGACATTAA	
Consenso	(1601)	GTGGATACAAATGCTGGTGGTCCAGAGAAGGAGAGATTTGGACATTAA	1651 1700
AF244520	(1268)	AGCAGCCTTGTCAAATTCATTCCGGGTTTGGTGGCCAACTCCAGCATCA	
AF318307	(1644)	AGCAGCCTTGTCAAATTCATTCCGGGTTTGGTGGCCAACTCCAGCATCA	
Consenso	(1651)	AGCAGCCTTGTCAAATTCATTCCGGGTTTGGTGG CA AACTCCAGCATCA	1701 1750
AF244520	(1318)	TTTTTGCTCCTTACAAGTGAAGC- -ACTCA- TTGCCTGTACTCCAACC	
AF318307	(1694)	TTTTTGCTCCTTACAAGTGAAGC- -AAGCAGTTGCTGTACTCCAACC	
Consenso	(1701)	TTTTTGCTCCTTACAAGTGAAGC A CA TTGC TGTACTCCAACC	1751 1800
AF244520	(1365)	TGGTTGTGTAAGTAACTTGCTGTAAGTGTTTACAAGAACTCCCATGTTATGTT	
AF318307	(1744)	TGATTGTATAAGTAACTTGCTGTAAG- GT-----	
Consenso	(1751)	TG TTGT TAACTTGCTGTAAG GT	1801 1850
AF244520	(1415)	GTTGCGGGAATCAACACAGTTTGTAAACTACCAAGAGCTAAGCTAAGTT	
AF318307	(1768)	-----	
Consenso	(1801)	-----	1851 1900
AF244520	(1465)	TCCTTAGGATCAAGATCCGTTTGTGCCAGAGAACTTGGACAAAGAGCAA	
AF318307	(1768)	-----	
Consenso	(1851)	-----	1901 1950
AF244520	(1515)	CGTAGCAGAGTTGGATTTAGCTTCCGTGTGATACCTTTTGAGTGAATC	
AF318307	(1768)	-----	
Consenso	(1901)	-----	1951 2000

FIG. 7C

```
AF244520 (1565) TTTGTAGCCTTTTCTTTTTTGTAGTGTTCATTCTATTTGTTAATCATT
AF318307 (1768) -----
Consenso (1951)
                2001                                     2050
AF244520 (1615) ACAATCTGAAAATTGCCAAACCAATTCTCCGTTAAATTTAGTAACTCTAC
AF318307 (1768) -----
Consenso (2001)
                2051                                     2079
AF244520 (1665) ACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AF318307 (1768) -----
Consenso (2051)
```

**FIG. 7D**

ES 2 657 067 T3

```

1                               50
Seq ID 3 (1) -----
Seq ID 4 (1) -----
Seq ID 46 (1) -----
Seq ID 47 (1) -----
AC189461 (1) -----
BH723504 (1) ACGGCACCTTTACGCACCAATATATACCATATGGTAGCCAGTTAAGGGTAC
Consenso (1) -----

51                               100
Seq ID 3 (1) -----
Seq ID 4 (1) -----
Seq ID 46 (1) -----
Seq ID 47 (1) -----
AC189461 (1) -----
BH723504 (51) GAAGTGATTTTCCCCTTTTCTTATGTATACAGACTCCAATTCGCCAAATT
Consenso (51) -----

101                               150
Seq ID 3 (1) ----- GAATTCGCCCTT
Seq ID 4 (1) ----- GAATTCGCCCTT
Seq ID 46 (1) ----- GAATTCGCCCTT
Seq ID 47 (1) -----
AC189461 (1) ----- GACATCGATAATT
BH723504 (101) TACTCTGGTCATCGAACATCGATACGTTTGTATGTATGTAACGTCATTATCT
Consenso (101) ----- GAATTCGCCCTT

151                               200
Seq ID 3 (13) CGCGGA--TCCGAACACTGCGTTTGCTGGCTTTGATGAAA-----
Seq ID 4 (13) CGCGGA--TCCGAACACTGCGTTTGCTGGCTTTGATGAAA-----
Seq ID 46 (13) CGCGGA--TCCGAACACTGCGTTTGCTGGCTTTGATGAAA-----C
Seq ID 47 (1) -----
AC189461 (13) TTCTGAGATTTGGAAGTCTGGATTTGATATAAGCAAAGCAAAGCAAGTGC
BH723504 (151) CTCTCTC-TGGCTATCTCTGTTGTTTGTGTTAAATACAGGAGG-----
Consenso (151) CGCGGA TCCGAACACTGCGTTTGCTGGCTTTGATGAAA

201                               250
Seq ID 3 (51) -----ATCTTTTGACGTCTCCACGCACGAGTTTACGCATCCCA
Seq ID 4 (51) -----ATGTATGTATCTTTTGACGTCTCCACGTTACGAGTTTACGCATCCA-
Seq ID 46 (53) TTGTATGTATGTATCTTTTGACGTCTCCACGTTACGAGTTTACGCATCCA-
Seq ID 47 (1) -----
AC189461 (63) TCGCTTGATGTATCTTTTGACGTCTCCACGCACGAGTTTACGCATCCA-
BH723504 (196) -TTTGTGTTTGTGTTTCTTTGACGTCTCCACGCACGATTTTAAAGCTTCCGT
Consenso (201) TGTATGTATCTTTTGACGTCTCCACGCACGAGTTTACGCATCCA

251                               300
Seq ID 3 (89) GACGCCTCGTTAGAGAGAAGAGAGAGATCGAGATCGAATTCCTC-TCCTC
Seq ID 4 (96) GACGCCTCGTTAGAGAGAAGAGAGAGATCGAGATCGAGATCGAGATCGATCGAT
Seq ID 46 (102) GACGCCTCGTTAGAGAGAAGAGAGAGATCGAGATCGAGATCGAGATCGATCGAT
Seq ID 47 (1) -----
AC189461 (112) GACGCCTCGTTAGAGAGAAGAGAGAGATCGAGATCGAATTCCTC-TCCTC
BH723504 (245) TACGCTCTGCACGCTCTCTGAAAGAGAGAGAGAGAGATCGAATTCATC
Consenso (251) GACGCCTCGTTAGAGAGAAGAGAGAGATCGAGATCGAGATCGA TC TC

301                               350

```

FIG. 8A



Seq ID 3	(424)	CTCACCAAATGCT-----CTGGATCCGGTAGCAG-----
Seq ID 4	(432)	CTCACCAAATGCTTGGCTCTTTCTGGATCCGGTAGCSTT---CAGGAGGC
Seq ID 46	(439)	CTCACCAAATGCT-----CTGGATCCGGTAGCAG-----
Seq ID 47	(168)	CTCACCAAATGCT-----CTGGATCCGGTAGCAG-----
AC189461	(453)	CTCACCAAATGCT-----CTGGATCCGGTAGCAGCGTTCAAGCAGGC
BH723504	(542)	CTCAGCAAATGCT-----CTGGATCCGGTAGCATTCTCGTGTCTTC
Consenso	(601)	CTCACCAAATGCT-----CTGGATCCGGTAGCAG-----G-----C
		651-----700
Seq ID 3	(453)	-----CACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
Seq ID 4	(479)	TCTCGTCACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
Seq ID 46	(468)	-----CACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
Seq ID 47	(197)	-----CACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
AC189461	(494)	TCTCGTCACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
BH723504	(583)	-----TTGTTGGAGTTTAAACCTTGCAATCACTACAACAAGAACAACA
Consenso	(651)	CACTACTTCCTTTGGCCTTGCAATCACTACAAT-----
		701-----750
Seq ID 3	(487)	--GCCTTGTCTT--CTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
Seq ID 4	(519)	--GCCTTGTCTT--CTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
Seq ID 46	(502)	--GCCTTGTCTT--CTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
Seq ID 47	(231)	--GCCTTGTCTT--CTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
AC189461	(534)	--GCCTTGTCTT--CTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
BH723504	(627)	AAGGCAATGCTTTCCTCTCTTCGGATCGAATAGCCTTTTCTCTCAATCGA
Consenso	(701)	GCCTTGTCTTCTCTCTTCGGATCGAACTCTGTTTCTCTCAATCGA
		751-----800
Seq ID 3	(532)	AACCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTTCCGGTGGAGGAGC
Seq ID 4	(564)	AACCAGAGGAGGTTGAAATCTGGCTG---CTGCTAGTTCCGGTGGAGGAGC
Seq ID 46	(547)	AACCAGAGGAGGTTGAMTCGGWCTGCTRCTGCTAGTTCCGGTGGAGGAGC
Seq ID 47	(276)	AACCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTTCCGGTGGAGGAGC
AC189461	(579)	AACCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTTCCGGTAAATTTAAC
BH723504	(677)	AACCAGAGGAAACTGAAATCGAGCAACTTCTTCTTCCGGTATAAACAT
Consenso	(751)	AACCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTTCCGGTGGAGGAGC
		801-----REV-----850
Seq ID 3	(582)	-----C--ATGGCTGTTGCGATGGATA-TGAAAAGGAAGCCAAGGTT
Seq ID 4	(611)	-----C--ATGGCTGTTGCGATGGATA-TGAAAAGGAAGCCAAGGTT
Seq ID 46	(597)	-----C--ATGGCTGTTGCGATGGATA-TGAAAAGGAAGCCAAGGTT
Seq ID 47	(326)	-----C--ATGGCTGTTGCGATGGATA-TGAAAAGGAAGCCAAGGTT
AC189461	(629)	ATTTAC---CAATCTGGGTTTTGATTTGTGGTATTGTTGAGAGGTT
BH723504	(727)	TCTTCTCCAAATCTCAAGTCTTGGTTTCTGGTATCGTTTGCAG-GCT
Consenso	(801)	CATGGCTGTTGCGATGGATA-TGAAAAGGAAGCCAAGGTT
		851-----900
Seq ID 3	(622)	GACAACAAACCTCCTACG-GAGCAGCG---CCGTGTTGTTGTGACAGGCA
Seq ID 4	(651)	GACAACAAACCCCTACG-GAGCAGCG---CCGTGTTGTTGTGACAGGCA
Seq ID 46	(637)	GACAACAAACCCCTACG-GAGCAGCG---CCGTGTTGTTGTGACAGGCA
Seq ID 47	(366)	GACAACAAACCTCCTACG-GAGCAGCG---CCGGTGTGTTGTGACAGGCA
AC189461	(676)	GACAAGCTTTGTTTGT-TAGTAAAGTTTCTCTTTTATGTTTATGT
BH723504	(776)	GACAAGCTTTGTTTGTGATAAAGTTTCTCTTT-----
Consenso	(851)	GACAACAAACCTCCTACG-GAGCAGCG---CCGTGTTGTTGTGACAGGCA
		901-----950

FIG. 8C

```

Seq ID 3 (668) TGGGAGTTGAAACATCACTAGGTCAATGACCCTGACACCTTTTATGAGAAT
Seq ID 4 (697) TGGGAGTTGAAACATCACTAGGTCAAGGGCGAATTC-----
Seq ID 46 (683) TGGGAGTTGAAACATCACTAGGTCAAGGGCGAATTC-----
Seq ID 47 (412) TGGGAGTTGAAACATCACTAGGTCAAGGGCGAATTC-----
AC189461 (725) TATTTCGTCGAAATGATTCAACTTGGAGATCAAGTTTAAACGGGAAAGTATG
BH723504 (813) -----
Consenso (901) TGGGAGTTGAAACATCACTAGGTCAAGGGCGAATTC
951 1000

Seq ID 3 (718) CTCCTACAAGGGCGAATTC-----
Seq ID 4 (733) -----
Seq ID 46 (719) -----
Seq ID 47 (448) -----
AC189461 (775) CTGATTTGGAATCTCAGCTTAATTGTCAAGTTTCATTCTTTATTGTGTGT
BH723504 (813) -----
Consenso (951)
1001 1050

Seq ID 3 (737) -----
Seq ID 4 (733) -----
Seq ID 46 (719) -----
Seq ID 47 (448) -----
AC189461 (825) GTAATATGCAAAGCACATTTCTGTGAGTTTGATTCAATATTGGACTCTAG
BH723504 (813) -----
Consenso (1001)
1051 1069

Seq ID 3 (737) -----
Seq ID 4 (733) -----
Seq ID 46 (719) -----
Seq ID 47 (448) -----
AC189461 (875) ATAAGGACTATGATGATAC
BH723504 (813) -----
Consenso (1051)

```

FIG. 8D

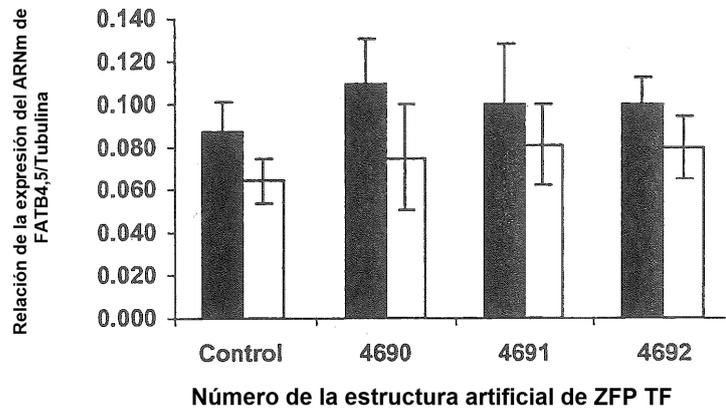


FIG. 9

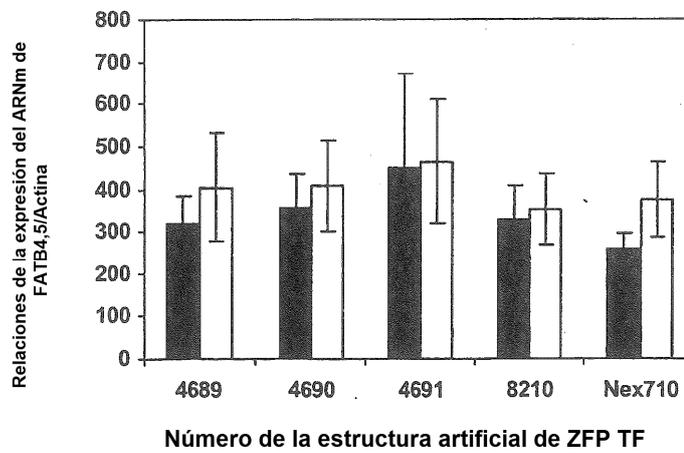


FIG. 10

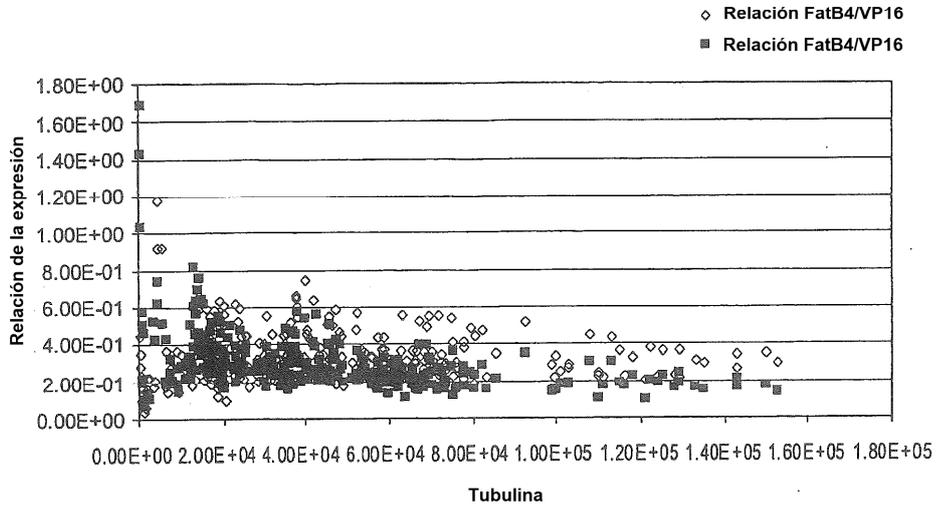


FIG. 11

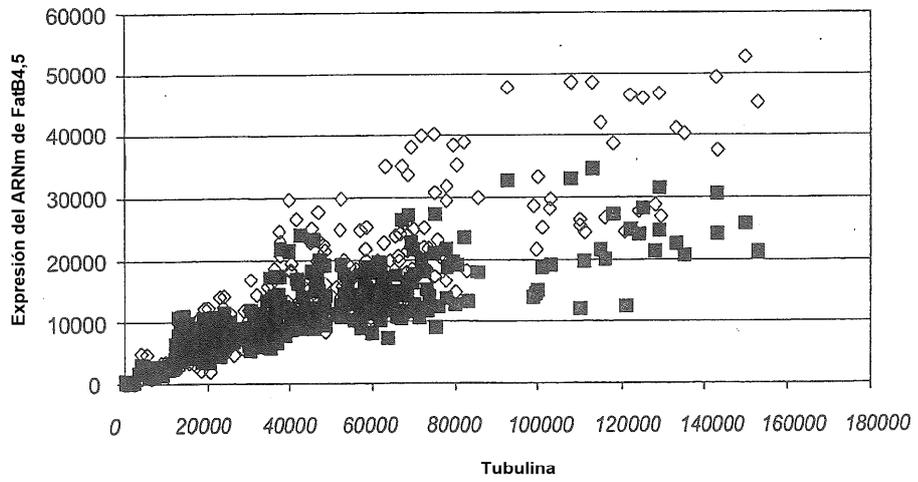


FIG. 12

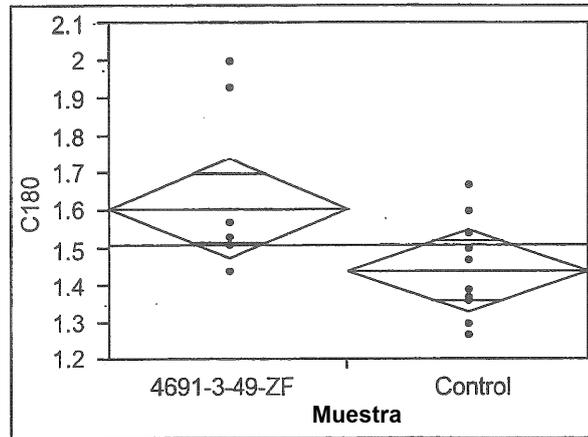


FIG. 13

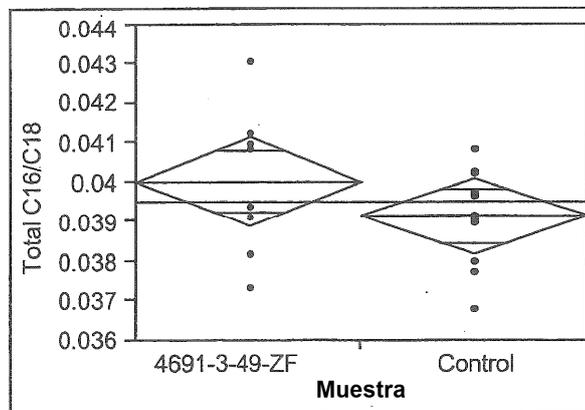


FIG. 14

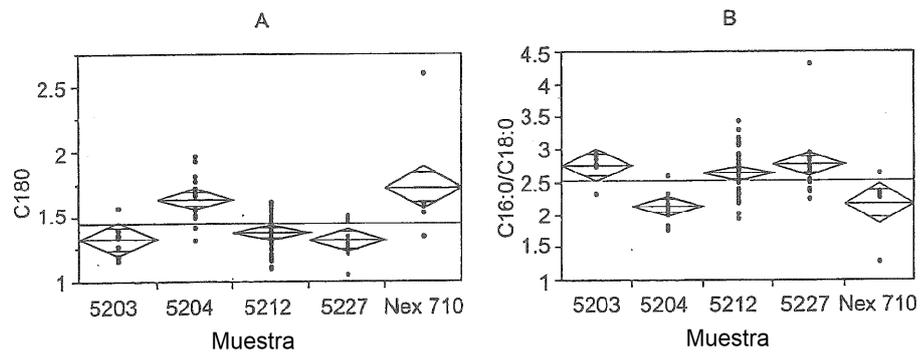


FIG. 15

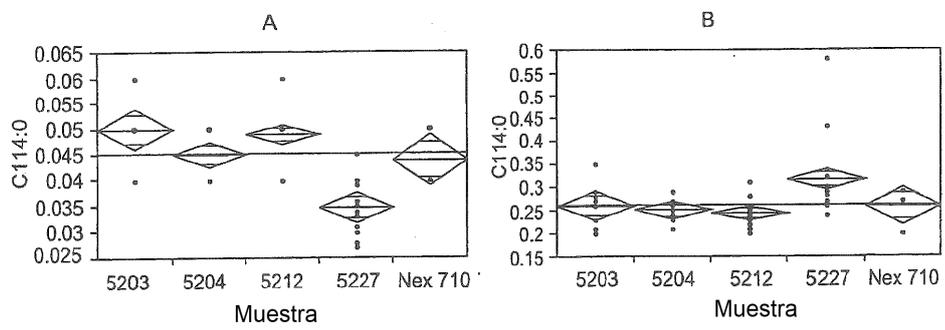
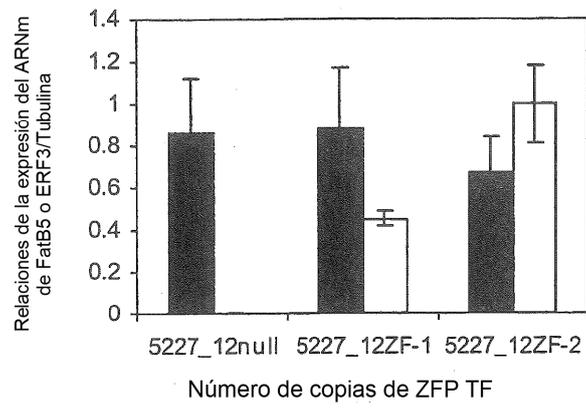
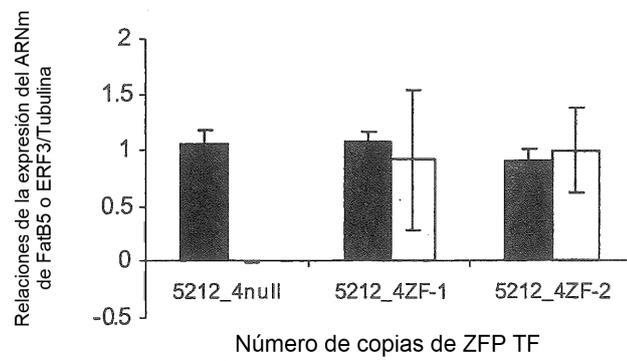


FIG. 16



**FIG. 17**



**FIG. 18**

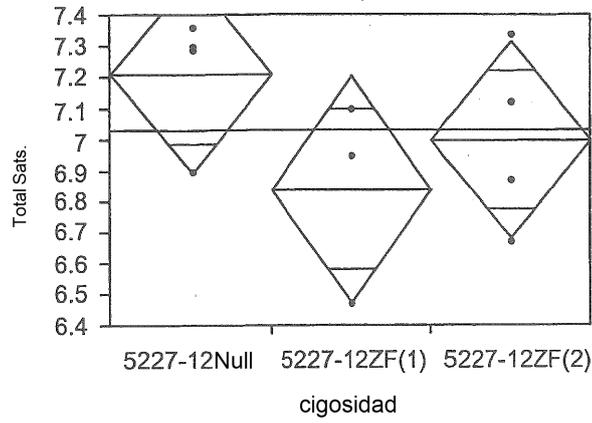


FIG. 19

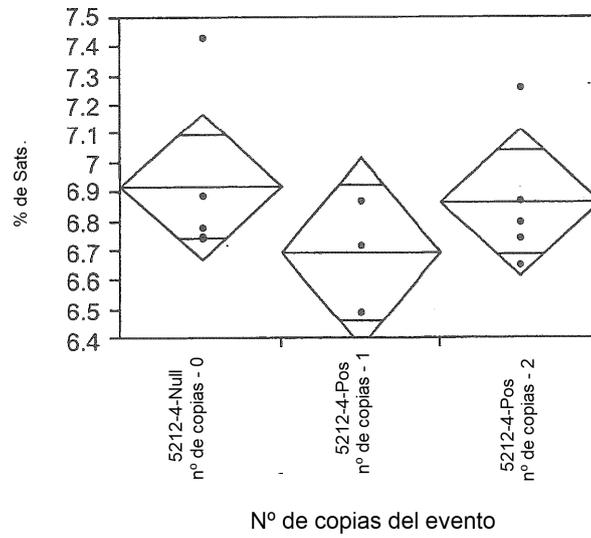


FIG. 20