

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 646 138**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 5/04 (2006.01)

A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.10.2010 PCT/US2010/002817**

87 Fecha y número de publicación internacional: **28.04.2011 WO11049627**

96 Fecha de presentación y número de la solicitud europea: **22.10.2010 E 10825333 (7)**

97 Fecha y número de publicación de la concesión europea: **20.09.2017 EP 2491127**

54 Título: **Proteínas de dedos de cinc manipuladas genéticamente dirigidas a los genes de planta involucrados 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:

12.12.2017

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 646 138 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Proteínas de dedos de cinc manipuladas genéticamente dirigidas a los genes de planta involucrados en la biosíntesis de ácidos grasos

Campo técnico

5 La descripción se refiere generalmente a los campos de la ingeniería de genoma y expresión de proteína en las plantas. En particular, la presente descripción se refiere a los dominios de unión de ADN manipulados genéticamente, por ejemplo, proteínas de dedos de cinc, que dirigen los genes involucrados en la síntesis de ácidos grasos y procedimientos de uso de tales proteínas de dedos de cinc en la modulación de la expresión génica, inactivación génica, y modificación del gen dirigida para generar plantas con perfiles de ácidos grasos alterados.

10 Antecedentes

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

20 A la luz del efecto de los ácidos grasos en la dieta y la salud, se han realizado intentos para alterar el perfil de ácidos grasos de las plantas utilizadas con fines comestibles e industriales. Sin embargo, los procedimientos convencionales de alteración de plantas para mejorar el perfil de ácidos grasos se basan en la mutagénesis (por ejemplo, química, radiación, etc.) y/o en mejoramiento genético y requieren mucho tiempo, son laboriosos y no se dirigen específicamente a genes seleccionados. Véase, por ejemplo, Patente U.S. N° 5.861.187.

25 Por ejemplo, las tres referencias siguientes describen la regulación de la expresión de KAS en plantas mediante procedimientos de direccionamiento no específicos.

30 Pidkowich et al. ((2007) *PNAS* 104 (11):4742-4747) describe el papel de la proteína transportadora beta-cetoacil-acilo (ACP) sintasa II (KAS II) e informa que la modulación de la actividad de KAS II puede producir cambios sustanciales en la composición de ácidos grasos de almacenamiento de la semilla. Este documento demuestra además que la manipulación de la actividad KAS II es un potente regulador de la composición de aceite de la semilla y suficiente para convertir su composición de semilla oleaginosa templada a la de un aceite tropical similar a la palma (página 4743, columna izquierda, final del segundo párrafo).

35 La patente U.S. N° 6.323.395 se refiere a procedimientos y composiciones para obtener aceites vegetales que tienen niveles saturados alterados mediante la modificación de los niveles de expresión y actividad de KAS II dentro de la célula.

La Solicitud de patente internacional N° WO 93/10240 A1 describe un fragmento de ácido nucleico que codifica una beta-cetoacil-ACP sintetiza II y que es útil para controlar la composición de ácidos grasos en cultivos de semillas oleaginosas.

40 Recientemente, los dominios de unión a ADN manipulados genéticamente tales como dominios de unión a ADN de meganucleasa y proteínas de dedos de cinc (ZFP) se han usado ventajosamente para modular selectivamente la expresión génica y para la alteración dirigida de secuencias génicas en las plantas (véanse, por ejemplo, las Patentes U.S. Nros. 7.262.054; 7.235.354, 7.220.719, 7.001.768 y 6.534.261; las Publicaciones de patente U.S. Nros. 2008/0182332 y U.S. Serie N° 12/284.888). Por ejemplo, la Patente U.S. N° 7.220.719 describe un procedimiento para modular la expresión de un gen celular endógeno en una célula. Desde la página 24 hasta la 26 bajo el capítulo llamado "Regulación de la expresión génica en plantas", se describe que las ZFP se pueden usar para manipular genéticamente plantas y, en particular, para mejorar la producción de aceite mediante la modificación de los ácidos grasos. Más particularmente, las ZFP se usan para regular la expresión del gen FAD2-1 en los porotos de soja. Las proteínas de dedos de cinc (ZFP) son proteínas que se unen al ADN, ARN y/o proteína, de una manera específica de la secuencia, en virtud de un dominio estabilizado de metal conocido como dedos de cinc. Ver, 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 ZFP se encuentran comúnmente en los factores de transcripción, y hasta la fecha, se han identificado más de 10.000 secuencias de dedos de cinc en varios miles de factores de transcripción conocidos o putativos.

55 Los dominios de unión a ADN también se pueden usar con dominios nucleasa para obtener nucleasas modificadas genéticamente. Por ejemplo, el dominio de unión a ADN de una endonucleasa de asentamiento se puede alterar para generar nuevas endonucleasas de asentamiento. De forma similar, los dominios de dedos de cinc también se

han combinado con los dominios de escisión de nucleasa para producir nucleasa de dedos de cinc (ZFN) para el direccionamiento específico de una ruptura de cadena doble en la región de un genoma donde se desea la modificación (por ejemplo, supresión, mutación, recombinación homóloga, o inserción de una secuencia exógena) (véase, por ejemplo, las Publicaciones de la solicitud de patente U.S. Nros. 2007/0134796; 2005/0064474; 2008/0182332). Las ZFP manipuladas genéticamente facilitan enormemente la inserción de secuencias exógenas o la modificación de secuencias endógenas en sitios blanco específicos en las plantas y proporcionan la alteración dirigida de genomas de plantas con mayores eficiencias que los procedimientos convencionales (véanse, por ejemplo, las patentes U.S. Nros 7.262.054, 7.235.354, 7.220.719, 7.001.768 y 6.534.261).

No obstante, aún existe la necesidad de composiciones y procedimientos para la alteración dirigida de genes implicados en la síntesis de ácidos grasos con el fin de producir plantas y productos de planta (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 aceite de semilla, se pueden producir productos alimenticios a base de aceite que contienen menos grasas saturadas. Tales productos beneficiarán la salud pública mediante la reducción de la incidencia de aterosclerosis y enfermedad cardíaca coronaria.

Compendio

La presente descripción describe composiciones y procedimientos para modular la expresión y para alteración dirigida en plantas enteras o células de planta de uno o más genes vegetales involucrados en la biosíntesis de ácidos grasos, de este modo se altera la composición de ácidos grasos en la planta entera o células de planta. Las plantas enteras o células de planta pueden ser de especies de plantas monocotiledóneas (monocotiledóneas) o dicotiledóneas (dicotiledóneas), que incluyen en algunas realizaciones particulares plantas productoras de aceite, y también incluyen células cultivadas, células en una planta en cualquier etapa de desarrollo y células de planta que se han retirado de una planta entera y estas células (o sus descendientes) se regenerarán en plantas. Las células de planta pueden contener una o más secuencias génicas homeólogas o parálogas, cualquier número de las cuales o todas las cuales pueden ser el blanco de modificación mediante los procedimientos descritos en la presente memoria.

En un aspecto, se describe en la presente memoria un dominio de unión a ADN (por ejemplo, proteína de dedos de cinc (ZFP)) que se une específicamente a un gen involucrado en una vía de biosíntesis de ácidos grasos de plantas. En algunas realizaciones, el gen es un gen de *Brassica napus*. En algunos aspectos particulares, el gen de *Brassica napus* puede codificar acetil-COA carboxilasa (ACCasa), β -cetoacil-ACP sintetasas (KAS, por ejemplo, KAS I-KAS IV), ácido graso tioesterasa B (FATB, por ejemplo, FATB1-FATB, u otras tioesterasas plastídicas) ácido graso sintasa (FAS), ácido graso elongasa (FAE, por ejemplo, FAE1), ácido graso tioesterasa A (FatA), ácido graso desaturasa (Fad2, Fad3), G-3-P deshidrogenasa plastídica (GPDH), gliceriquinasa (GK), desaturasa de proteína transportadora de estearoli-acilo (S-ACP-DES) y oleoil-ACP hidrolasa. 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 de dedos de cinc natural, que comprende las regiones de la hélice de reconocimiento mostrada en una fila única de la Tabla 1, en la que la proteína de fusión modula la expresión de al menos un gen de β -cetoacil-ACP sintetasa (KAS) endógeno. Más preferentemente el gen KAS está contenido dentro de una planta de *Brassica*. En una realización aún preferida, la proteína de dedos de cinc se une a un sitio blanco como se muestra en la Tabla 2.

En otro aspecto adicional, también se proporcionan proteínas de fusión que comprenden cualquiera de los dominios de unión a ADN (por ejemplo, ZFP) descritos en la presente memoria. En ciertas realizaciones, la proteína de fusión comprende una proteína de dedos de cinc y un dominio regulador transcripcional (por ejemplo, dominio de activación o represión), también conocido como ZFP TF. En otras realizaciones, la proteína de fusión comprende una ZFP y un dominio de nucleasa para formar una nucleasa de dedos de cinc (ZFN) que se escinde en una región genómica de interés con el gen blanco. En ciertos aspectos, la ZFN comprende un polipéptido de fusión que comprende un dominio de unión de dedos de cinc manipulado genéticamente que tiene especificidad por un gen involucrado en una vía de biosíntesis de ácidos grasos de plantas (por ejemplo, un gen que codifica ACCase, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, FATB, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK o S-ACP-DES) y un dominio de escisión, y/o uno o más polipéptidos de fusión que comprenden un dominio de unión de dedos de cinc manipulado genéticamente y un semidominio de escisión. En ciertos aspectos, los dominios de unión del dedos de cinc se unen a un sitio blanco como se muestra en la Tabla 2 o en la Tabla 10A. En ciertos aspectos, los dominios de unión del dedos de cinc comprenden una secuencia seleccionada del grupo que consiste en proteínas de dedos de cinc que comprenden los dominios de reconocimiento (por ejemplo, una fila única) mostrados en la Tabla 1 o la Tabla 10B. Se pueden obtener dominios de escisión y semidominios de escisión, por ejemplo, a partir de diversas endonucleasas de restricción y/o endonucleasas de asentamiento. En un aspecto, los semidominios de escisión se derivan de una endonucleasa de restricción de tipo IIS (por ejemplo, Fok I).

En una realización preferida la presente invención proporciona un polinucleótido que codifica una o más proteínas de dedos de cinc que comprenden las regiones de la hélice de reconocimiento mostradas en una fila única de la Tabla 1 o proteínas de fusión que comprenden un dominio funcional, que es un dominio regulador transcripcional o un

dominio de escisión y una proteína de dedos de cinc natural que comprende las regiones de la hélice de reconocimiento que se muestran en una fila única de la Tabla 1, en la que dicha proteína de fusión modula la expresión de al menos un gen de β -cetoacil-ACP sintetasa (KAS) endógeno.

5 En otros aspectos, en la presente se describen polinucleótidos que codifican cualquiera de los dominios de unión a ADN (*por ejemplo*, proteínas de dedos de cinc) y/o proteínas de fusión descritas en la presente memoria. En ciertos aspectos, la presente se describe un vector de expresión de ZFP que comprende a polinucleótido, que codifica una o más ZFP descritas en la presente memoria, unidas operativamente a un promotor. En una realización, una o más de las ZFP son ZFN.

10 Las ZFP y las proteínas de fusión que comprenden estas ZFP se pueden unir y/o escindir un gen involucrado en la síntesis de ácidos grasos dentro de la región codificadora del gen o en una secuencia no codificadora dentro o adyacente al gen, tal como, por ejemplo, una secuencia líder, secuencia de remolque o intrón, o dentro de una región no transcripta, corriente arriba o corriente abajo de la región codificadora. En ciertos aspectos, las ZFP o ZFN se unen y/o escinden una secuencia codificadora o una secuencia reguladora de un gen involucrado en la biosíntesis de ácidos grasos.

15 En otro aspecto, se describen en la presente memoria composiciones que comprenden una o más proteínas, proteínas de fusión o polinucleótidos descritos en la presente memoria. Las células de plantas pueden contener un blanco génico único o copias parálogas múltiples del mismo gen. Por lo tanto, las composiciones pueden comprender una o más proteínas que contienen ZFP (y polinucleótidos que las codifican) que se dirigen a uno o más genes involucrados en la síntesis de ácidos grasos en una célula de planta. Las ZFP se pueden dirigir a todos los genes parálogos u homeólogos y a los genes parálogos u homeólogos particulares seleccionados en una célula de planta o una combinación de los mismos.

20 En otro aspecto, se proporciona en la presente memoria una célula huésped de planta que comprende una o más proteínas o polinucleótidos (*por ejemplo*, vector de expresión de ZFP) como se describe en la presente memoria. Para los polinucleótidos, la célula huésped de planta se pueden transformar de forma estable o transfectar de forma transitoria o una combinación de las mismas con uno o más vectores de expresión de ZFP. En una realización, el uno o más vectores de expresión de ZFP expresa una o más ZFN en la célula huésped de planta. En otra realización, el uno o más vectores de expresión de ZFP expresa una o más ZFP TF en la célula huésped de planta.

25 En otro aspecto, se proporciona en la presente memoria un procedimiento de modificación de uno o más genes involucrados en la biosíntesis de ácidos grasos en la célula de planta, el procedimiento que comprende:(a) introducir, en la célula de planta, uno o más vectores de expresión que codifican una o más ZFP TF que se unen a un sitio blanco en los uno o más genes involucrados en la biosíntesis de ácidos grasos en condiciones tales que las ZFP TF se expresan y la expresión de los uno o más genes se modifica, y más particularmente la una o más proteínas de dedos de cinc, que comprenden las regiones de la hélice de reconocimiento que se muestran en una fila única de la Tabla 1, se expresan y la expresión de uno o más genes KAS se modifica. La modificación comprende modular la expresión de al menos un gen KAS. La modulación pueden ser activación o represión de la expresión de al menos un gen KAS. En ciertos aspectos, al menos un sitio blanco está en un gen de ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, *FATB*, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK, y/o S-ACP-DES. En otros aspectos, se modula más de un gen involucrado en la biosíntesis de ácidos grasos. En cualquiera de los procedimientos de modificación de uno o más genes involucrados en la biosíntesis de ácidos grasos descritos en la presente memoria, el procedimientos produce células de plantas con contenido de ácido graso modificado, por ejemplo, una reducción de la cantidad de grasas saturadas en las células de planta. En algunas realizaciones, el contenido de ácido graso modificado en las células de planta produce un contenido de ácido graso modificado en las semillas de la planta, por ejemplo, una reducción de la cantidad de grasas saturadas en las semillas de la planta. En algunas realizaciones la planta es una planta productora de aceite con contenido de ácido graso modificado en las semillas de la planta productora de aceite (*por ejemplo*, grasas saturadas reducidas). En alguna realización particular, la planta es una planta de *Brassica napus* con contenido de ácido graso modificado en las semillas de la planta de *Brassica napus* (*por ejemplo*, grasas saturadas reducidas).

30 En otro aspecto, se proporciona en la presente memoria un procedimiento para escindir uno o más genes involucrados en la biosíntesis de ácidos grasos en una célula de planta, el procedimiento que comprende:(a) introducir, en la célula de planta, uno o más vectores de expresión que codifican una o más nucleasas (*por ejemplo*, ZFN) que se unen a un sitio blanco en los uno o más genes involucrados en la biosíntesis de ácidos grasos en condiciones tales que las nucleasas (*por ejemplo*, ZFN) se expresan y los uno o más genes se escinden, y más particularmente, al menos se escinde un gen KAS. En ciertos aspectos, al menos un sitio blanco está en un gen que codifica ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, *FATB*, FAS, FAE1, FatA, Fad2, Fad3, GPDH, GK, y/o S-ACP-DES. En otros aspectos, se escinde más de un gen involucrado en la biosíntesis de ácidos grasos. Además, en cualquiera de los procedimientos descritos en la presente memoria, la escisión de los uno o más genes puede producir supresión, adición y/o sustitución de nucleótidos en la región escindida.

35 También se describe en la presente memoria un procedimiento para introducir una secuencia exógena en el genoma de una célula de planta, el procedimiento comprende las etapas de:(a) poner en contacto la célula con una secuencia exógena (vector donante); y (b) expresar una o más nucleasas (*por ejemplo*, nucleasas de dedos de cinc)

como se describe en la presente memoria en la célula, en que la una o más nucleasas escinden ADN cromosómico; de 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, la una o más nucleasas son fusiones entre el dominio de escisión de una endonucleasa de restricción de tipo II y un dominio de unión de dedos de cinc manipulado genéticamente. En otros aspectos, la nucleasa comprende una endonucleasa de asentamiento, por ejemplo, una endonucleasa de asentamiento con un dominio de unión a ADN modificado. En cualquiera de los procedimientos descritos en la presente memoria, la secuencia exógena puede codificar un producto de proteína.

En aún otro aspecto, también se proporciona una célula de planta transgénica o no transgénica obtenida de acuerdo con cualquiera de los procedimientos descritos en la presente memoria.

En otro aspecto, se proporciona en la presente memoria una planta que comprende una célula de planta transgénica o no transgénica obtenida como se describe en la presente memoria.

En otro aspecto, se proporciona en la presente memoria una semilla de planta que comprende una célula de planta transgénica o no transgénica obtenida como se describe en la presente memoria.

En otro aspecto, en la presente memoria se describe un aceite proveniente de una semilla extraída de una planta que comprende una célula de planta transgénica o no transgénica obtenida como se describe en la presente memoria.

Breve descripción de los dibujos

La **Figura 1** es un esquema que representa la vías de biosíntesis de ácidos grasos en canola (*B. napus*). Esto fue adaptado por John Shanklin, Brookhaven National Laboratory, Upton, NY (Thelen and Ohlrogge, 2002, Metabolic engineering 4:12-21).

La **Figura 2** es un esquema que describe las ubicaciones del sitio blanco dentro del gen KASII para varios ejemplos de ZFP TF dirigidos a KASII. Los números indican los sitios blanco para las ZFP contenidas dentro de los constructos que se muestran 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 de ARNm de CASI en las hojas de planta T₀ transformadas con el constructo de ZFP TF de activación KASII 4695, analizado por qRT-PCR. Se usó un promedio de 27 plantas de control para calcular la regulación por aumento que se muestra en el gráfico. Se observó una regulación por aumento de más de tres veces de ARNm de KASII en ciertos eventos.

La **Figura 4** es un gráfico que representa las relaciones de expresión de KASII-ZFP TF/tubulina en hojas de plantas T₁ detectadas mediante qRT-PCR. Se compararon tres eventos junto con los valores nulos correspondientes.

La **Figura 5** es un gráfico que representa las relaciones de expresión de ARNm de KASII/tubulina promedio en las plantas T₁ que contienen ZFP TF y nulos segregantes de cada uno de los tres eventos determinado por qRT-PCR. Se observaron 2-3 veces la regulación por aumento de ARNm de KASII en estas hojas de plantas T₁.

Las **Figuras 6A y 6B** son gráficos, extraídos con el software estadístico JMP para el análisis de una vía de ácidos grasos blanco, que representan una disminución constante y significativa ($p = <0.001$) en el C16 total (6A) y una disminución en las relaciones de C16 total/C18 total (6B) en las plantas ZFP TF positivas y nulas hermanas dentro de cada evento. El C16 total estaba compuesto por los contenidos de C16:0 y C16:1 y C18 total estaba compuesto por los contenidos de C18:0, 18:1, 18:2 y 18:3.

La **Figura 7** representa un alineamiento de secuencia de AF318307 y AF244520. El sombreado que las regiones de homología exacta.

La **Figura 8** representa un alineamiento de secuencia de SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:46, SEQ ID NO:30, AC189461, y BH723504. Las secuencias de cebador directo e inverso (SEQ ID NOs 28 y 29) para la amplificación del ADNc de β -cetoacil-ACP sintetasa II se resaltan mediante líneas discontinuas por encima de la secuencia correspondiente. El sombreado indica regiones de homología exacta.

La **Figura 9** representa las expresiones del gen FatB4 y *FatB* en líneas de callo de *B. napus* transgénicas para diferentes diseños de TF ZFP presentes en los constructos pDAB4690 - pDAB4692. Se analizaron 17 líneas para control y 25 líneas para cada uno de los constructos de ZFP. Barras negras = expresión de ARNm de FatB4; barras grises = expresión de ARNm de *FatB*.

La **Figura 10** muestra las expresiones de FatB4 y *FatB* en hojas de plantas transgénicas *B. napus* T₀ analizadas con qRT-PCR. Se analizaron tres constructos que comprenden ZFP TF, pDAB4689 - pDAB4691 en plantas transgénicas. El número total de plantas transgénicas T₀ independientes analizadas para este experimento fue de 40, 62, 41 y 22 para pDAB 4689 - pDAB4691 y pDAB8210, respectivamente. El control Nex710 estaba compuesto por 10 plantas. El noventa y siete por ciento de los eventos transgénicos para 3 constructos de ZFP TF fueron positivos para la expresión de ZFP TF, según se determinó mediante el ensayo de expresión de ZFP TF (Ejemplo

8.3). Se obtuvieron resultados similares cuando se usó la expresión de ARNm de tubulina nativa como referencia para normalizar las expresiones génicas de *FatB*. Barra negra = expresión de *FatB4*; barra gris = expresión de *FatB*.

La **Figura 11** es un diagrama que muestra una relación lineal entre los genes *FatB* y la expresión de tubulina cuando se analiza a partir de plantas T_1 . Cuadrados negros = expresión de *FatB4*; diamantes grises = expresión de *FatB*.

- 5 La **Figura 12** muestra qRT-PCR para la expresión de *FatB4* y *FatB* en hojas de planta T_1 transgénicas para el constructo de TF ZFP pDAB4691. Cuadrados negros = expresión de *FatB4*; diamantes grises = expresión de *FatB*.

La **Figura 13** es un análisis de una vía del contenido de C18:0 por muestra utilizando el software estadístico JMP.

La **Figura 14** muestra un análisis de una vía del contenido de C16/C18 por muestra utilizando el software estadístico JMP.

- 10 La **Figura 15** representa un análisis de los perfiles T_1 FA de semillas maduras que comprende 4 eventos de constructos ZFP TF se muestra con el software estadístico JMP. Los paneles A y B representan un análisis de una vía por muestra (constructo) de contenido de C18:0 y contenido de C16:0/C18:0, respectivamente.

- 15 La **Figura 16** representa un análisis de los perfiles T_1 FA de semillas maduras que comprende 4 eventos de constructo ZFP TF se muestra con el software estadístico JMP. Los paneles A y B muestran el análisis de una vía por muestra (constructo) de contenido C14:0 y C16:1 respectivamente, se resalta el atributo diferenciador del constructo pDAS5227. El análisis de los perfiles de FA de T_1 de semilla madura que comprende 4 eventos de constructos TF ZFP se muestra con el software estadístico JMP. Los paneles A y B muestran el análisis de una vía por muestra (constructo) del contenido C14:0 y C16:1 respectivamente, se resalta el atributo diferenciador del constructo pDAS5227.

- 20 La **Figura 17** muestra la regulación por disminución del ARNm de *FatB* (barras negras) en la semilla inmadura T_2 transformada con el constructo pDAS5227 ZFP TF. La expresión de ZFP TF está representada por la expresión de ERF3 (barras grises). Se analizaron 25 semillas DAF inmaduras de cuatro plantas T_1 nulas, tres heterocigotas (5227_12ZF - 1) y cuatro homocigotas (5227_12ZF - 2) del evento 5227-12.

- 25 La **Figura 18** muestra la regulación por disminución del ARNm de *FatB* (barras negras) en la semilla inmadura T_2 transformada con el constructo pDAS5212 ZFP TF. La expresión de ARNm de ZFP TF está representada por la expresión de KRAB1 (barras grises). Se analizaron 25 semillas DAF inmaduras de cinco plantas nulas, tres heterocigotas y cuatro homocigotas del evento 5212-4.

- 30 La **Figura 19** muestra un análisis de una vía de los ácidos grasos saturados totales (sat) por cigosidad en la semilla T_2 del evento ZFP TF 5227-12. Las semillas T_2 obtenidas de las plantas T_1 originales nulas, heterocigotas y homocigotas de T_1 se marcan como 5227-12ZF nulo, 5227-12ZF (1) y 5227-12ZF (2), respectivamente.

La **Figura 20** muestra un análisis de una vía de los ácidos grasos saturados totales (sat) por cigosidad en la semilla T_2 del evento ZFP TF 5212-4. Las semillas T_2 obtenidas de las plantas T_1 originales nulas, heterocigotas y homocigotas se marcan como 5227-12ZF nulo, 5227-12ZF (1) y 5227-12ZF (2), respectivamente.

Descripción detallada

- 35 En la presente memoria se describen composiciones y procedimientos útiles para la modulación de la expresión y la escisión dirigida y la alteración de uno o más genes involucrados en la síntesis de ácidos grasos en plantas. La regulación de tales genes se puede modular, por ejemplo, usando factores de transcripción de ZFP manipulados genéticamente o modificación de regiones reguladoras de genes. Los genes se pueden alterar, por ejemplo, mediante escisión dirigida seguida de recombinación homóloga intracromosómica o mediante escisión dirigida seguida de recombinación homóloga entre un polinucleótido exógeno (que comprende una o más 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 organelas (*por ejemplo*, mitocondrias, plástidos), cromosomas artificiales y cualquier otro tipo de ácido nucleico presente en una célula tal como, por ejemplo, secuencias amplificadas, cromosomas de doble minuto y genomas bacterias y virus endógenos o infecciosos. Las secuencias genómicas pueden ser normales (es decir, de tipo salvaje) o mutantes; las secuencias mutantes pueden comprender, por ejemplo, inserciones, supresiones, translocaciones, reordenamientos y/o mutaciones puntuales. Una secuencia genómica también puede comprender uno de numerosos alelos diferentes.

- 50 Las composiciones descritas en la presente memoria comprenden una o más ZFP que comprenden dominios de unión por dedos de cinc manipulados genéticamente, polinucleótidos que codifican estos polipéptidos y combinaciones de ZFP y polinucleótidos que codifican ZFP. Un dominio de unión del dedos de cinc puede comprender uno o más dedos de cinc (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de cinc), y se puede manipular genéticamente para unirse a cualquier secuencia genómica de un gen, que incluyen secuencias reguladoras unidas operativamente al gen, involucrado en la biosíntesis de ácidos grasos.

Las ZFP como se describen en la presente memoria se pueden usar para regular la expresión génica, ya sea a través de la activación o represión de la transcripción génica. Las ZFP que comprenden fusiones de dominios de dedos de cinc unidos a dominios reguladores se pueden construir para crear factores de transcripción quiméricos que activan o reprimen la transcripción. Las ZFP también se pueden usar para la escisión dirigida de una región genómica de interés mediante la unión de dominios de dedos de cinc con dominios de escisión de nucleasas (o semidominios de escisión) para producir nucleasas de dedos de cinc. Por lo tanto, mediante la identificación de una región genómica blanco de interés en la que se desea la regulación, escisión o recombinación génica, se puede construir, de acuerdo con los procedimientos descritos en la presente invención, una proteína de dedos de cinc que comprende una o más proteínas de fusión que comprenden uno o más dominios reguladores y/o dominios de escisión (o semidominios de escisión) unidos a un dominio de dedos de cinc manipulado genéticamente para reconocer una secuencia del gen en esa región genómica. La presencia de tal ZFP que comprende una proteína de fusión (o proteínas) en una célula producirá la unión de la proteína de fusión a su sitio de unión y alteración de la regulación o escisión 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, la recombinación homóloga se produce a una tasa alta entre la región genómica y el polinucleótido exógeno.

Como se muestra en la Figura 1, existen varios genes que están involucrados en la biosíntesis de ácidos grasos. En consecuencia, las composiciones descritas en la presente memoria pueden dirigir uno o más de estos genes en una célula de planta, que incluye, pero sin limitación genes de ACCasa, KAS I, KAS II, KAS III, KAS IV, FATB1, FATB2, FATB3, FATB4, *FATB*, FAS, FAE1, Fata, Fad2, Fad3, GPDH, GK, o S-ACP-DES y ortólogos, parálogos y homeólogos de estos genes. Por ejemplo, 1, 2, 3, 4, 5, o más genes involucrados en la biosíntesis de ácidos grasos pueden ser dirigidos por proteínas (*por ejemplo*, ZFP) descritos en la presente memoria. En consecuencia, una o más ZFP o vectores de expresión que codifican ZFP de diferentes especificidades se pueden combinar para dirigir los genes deseados de interés en una planta.

General

La puesta en práctica de los procedimientos, así como la preparación y el uso de las composiciones descritas en la presente memoria emplean, a menos que se indique otra cosa, técnicas convencionales de biología molecular, bioquímica, estructura y análisis de cromatina, química de computación, cultivos celulares, ADN recombinante y campos relacionados que se encuentran en la pericia de la técnica. Estas técnicas se explican totalmente en la bibliografía. Véase, por ejemplo, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edición, Cold Spring Harbor Laboratory Press, 1989 y Tercera edición, 2001; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 y actualizaciones periódicas; la serie METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Third edition, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; and METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definiciones

Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en forma de cadena simple o cadena doble. A los fines de la presente descripción, estos términos no se deben interpretar como limitantes respecto de la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos modificados en la base, residuos de azúcar y/o fosfato (por ejemplo, esqueletos 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 apareará con T.

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

"Unión" se refiere a una interacción específica de secuencia, no covalente, entre macromoléculas (*por ejemplo*, entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión necesitan ser específicos de secuencia (*por ejemplo*, los contactos con residuos fosfato en un esqueleto de ADN), siempre que la interacción como un todo sea específica de secuencia. Dichas interacciones generalmente se caracterizan por tener una constante de disociación (K_d) de $10^{-6} M^{-1}$ o inferior. "Afinidad" se refiere a la fuerza de unión, el aumento de la afinidad de unión se correlaciona con una menor K_d .

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

Una "proteína de unión a ADN de dedos de cinc" (o dominio de unión) es una proteína o un dominio dentro de una proteína más grande, que se une al ADN en forma específica de secuencia a través de uno o más dedos de cinc, los cuales son regiones de la secuencia de aminoácidos dentro del dominio de unión cuya estructura se estabiliza mediante la coordinación de un ión cinc. El término proteína de unión a ADN de dedos de cinc a menudo se abrevia como proteína de dedos de cinc o ZFP.

Los dominios de unión de dedos de cinc se pueden "manipular genéticamente" para unirse a una secuencia de nucleótidos predeterminada (*por ejemplo*, una secuencia blanco en cualquier gen involucrado en biosíntesis de ácidos grasos). Los ejemplos no limitantes de procedimientos para manipular genéticamente las proteínas de dedos de cinc son el diseño y la selección. Una proteína de dedos de cinc diseñada es una proteína que no se encuentra en la naturaleza, cuyo diseño/composición es principalmente el resultado de criterios racionales. Los criterios racionales de diseño incluyen la aplicación de normas de sustitución y algoritmos computados para procesar información en una base de datos que almacena información sobre los diseños de ZFP y los datos de unión existentes. Ver, por ejemplo, las patentes U.S. 6.140.081, 6.453.242, 6.534.261 y 6.785.613; ver, también WO 98/53058, WO 98/53059, WO 98/53060, WO 02/016536 y WO 03/016496; y las patentes U.S. 6.746.838; 6.866.997; y 7.030.215. En consecuencia, una proteína de dedos de cinc "manipulada genéticamente" o proteína de dedos de cinc "no natural" es una en la que uno o más de los dominios de unión a ADN de dedos de cinc componente (hélices de reconocimiento) no son naturales y se han manipulado genéticamente para unirse a un sitio blanco preseleccionado.

Una proteína de dedos de cinc "seleccionada" es una proteína que no se encuentra en la naturaleza, cuya producción es principalmente el resultado de un proceso empírico tal como despliegue en fago, trampa de interacción o selección de híbridos. Véase, *por ejemplo*, 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 01/60970WO 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 de cadena simple o de cadena doble. El término "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede ser de cualquier longitud, por ejemplo, entre 2 y 250.000 nucleótidos de longitud (o cualquier valor entero entre ellos o por encima de este), preferentemente entre aproximadamente 100 y 5000 nucleótidos de longitud (o cualquier número entero entre ellos), con más preferencia entre aproximadamente 200 y 2500 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 homóloga no idéntica" se refiere a una primera secuencia que comparta cierto 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 salvaje 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 la recombinación homóloga entre ellas, mediante la utilización de mecanismos celulares normales. Dos secuencias homólogas no idénticas pueden tener cualquier longitud y su grado de no homología puede ser tan pequeño como de un nucleótido único (por ejemplo, para la corrección de una mutación genómica puntual mediante recombinación homóloga dirigida) o tan grande 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 homólogas no idénticas no necesariamente tienen la misma longitud. Por ejemplo, se puede usar un polinucleótido exógeno (es decir, polinucleótido donante) de entre 20 y 10.000 nucleótidos o pares de nucleótidos.

Las técnicas para determinar la identidad de las secuencias de ácidos nucleicos y aminoácidos son conocidas en la técnica. Generalmente, dichas técnicas incluyen determinar la secuencia de nucleótido del ARNm de un gen y/o determinar la secuencia de aminoácidos así codificada, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se puede determinar y comparar de esta manera. En general, la identidad se refiere a una correspondencia exacta entre nucleótido a nucleótido o aminoácido a aminoácido en dos secuencias de polinucleótidos o polipéptidos, respectivamente. Dos o más secuencias (polinucleótido o aminoácido) se pueden comparar mediante la determinación de su porcentaje de identidad. El porcentaje de identidad de dos secuencias, sean secuencias de ácidos nucleicos o aminoácidos, es la cantidad de coincidencias exactas entre dos secuencias alineadas, dividida por la longitud de las secuencias más cortas y multiplicado por 100. Un alineamiento aproximado de las secuencias de ácidos nucleicos es provisto por el algoritmo de homología local de Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). Este algoritmo se puede aplicar a las secuencias de aminoácidos mediante el uso de la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, y normalizado por Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). Un ejemplo de implementación de este algoritmo para determinar el porcentaje de identidad de una secuencia es provisto por Genetics Computer Group (Madison, WI) en la aplicación de utilidad "BestFit". Los parámetros predeterminados para este procedimiento se describen en el Wisconsin Sequence Analysis Package Program Manual, Versión 8 (1995) (disponible en Genetics Computer Group, Madison, WI). Un procedimiento preferido para establecer el porcentaje de identidad en el contexto de la presente descripción es usar el paquete de programas MPSRCH con derechos de autor de la Universidad de Edimburgo, desarrollado por John F. Collins y Shane S.

Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). A partir de este conjunto de paquetes, se puede emplear el algoritmo de Smith-Waterman donde se usan los parámetros predeterminados para la tabla de puntuación (por ejemplo, penalidad de brecha abierta de 12, penalidad de extensión de brecha de uno y una brecha de seis). A partir de los datos generados, el valor "Coincidencia" refleja la identidad de secuencia. Otros programas adecuados para calcular el porcentaje de identidad o semejanza entre las secuencias son conocidos en general en la técnica, por ejemplo, otro programa de alineamiento es BLAST, usado con parámetros predeterminados. Por ejemplo, se pueden usar BLASTN y BLASTP mediante los siguientes parámetros predeterminados: código genético = estándar; filtro = ninguno; cadena = ambas; corte = 60; esperado = 10; matriz = BLOSUM62 (para BLASTP); Descripción = 50 secuencias; calificados por = HIGH SCORE; base de datos = no redundante, GenBank + EMBL + DDBJ + PDB + traducciones GenBank CDS + Swiss Protein + Spupdate + PIR. Los detalles de estos programas se pueden hallar en Internet. Con respecto a las secuencias descritas en la presente, el rango de grados deseados de identidad de secuencia es de aproximadamente 35% a 100% y cualquier valor entero entre ellos. Generalmente, los porcentajes de identidad entre secuencias son de al menos 35%-40%; 40%-45%; 45%-50%; 50%-60%; 60%-70%; 70-75%, preferentemente 80-82%, con más preferencia 85-90%, aun con más preferencia 92%, aún con más preferencia 95%, y con máxima preferencia 98% de identidad de secuencia.

Alternativamente, el grado de semejanza de secuencia entre polinucleótidos se puede determinar por hibridación de polinucleótidos en condiciones que permiten la formación de dúplex estables entre regiones homólogas, seguida de digestión con nucleasas específicas de cadenas únicas y determinación del tamaño de los fragmentos digeridos. Dos secuencias de ácidos nucleicos o de dos polipéptidos son sustancialmente homólogas entre sí cuando las secuencias exhiben al menos aproximadamente 70%-75%, preferentemente 80%-82%, más preferentemente 85%-90%, más preferentemente aún 92%, con aún mayor preferencia 95%, y con máxima preferencia 98% de identidad de secuencia para una longitud definida de las moléculas, tal como se determina mediante los procedimientos anteriores. Tal como se usa en la presente memoria, sustancialmente homólogo también se refiere a secuencias que muestran completa identidad con una secuencia específica de ADN o polipéptido. Las secuencias de ADN que son sustancialmente homólogas se pueden identificar en un experimento de hibridación Southern, por ejemplo, en condiciones rigurosas, tal como se define para dicho sistema particular. La definición de las condiciones de hibridación adecuadas está dentro de la pericia en la técnica. Véase, *por ejemplo*, Sambrook et al., *supra*; Nucleic Acid Hybridization: A Practical Approach, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

La hibridación selectiva de dos fragmentos de ácido nucleico se puede determinar de la siguiente manera. El grado de identidad de secuencia entre dos moléculas de ácido nucleico afecta la eficiencia y la fuerza de los eventos de hibridación entre dichas moléculas. Una secuencia parcialmente idéntica de ácido nucleico inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica a una molécula blanco. La inhibición de la hibridación de la secuencia completamente idéntica se puede evaluar mediante ensayos de hibridación 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.). Tales ensayos se pueden realizar mediante el uso de grados variables de selectividad, por ejemplo, mediante condiciones que varían de baja a alta rigurosidad. Si se emplean condiciones de baja rigurosidad, se puede evaluar la ausencia de unión no específica mediante una sonda secundaria que carece incluso de un grado parcial de identidad de secuencia (por ejemplo, una sonda con menos de aproximadamente 30% de identidad de secuencia con la molécula blanco), de manera tal que en ausencia de eventos de unión no específicos, la sonda secundaria no se hibrida con el blanco.

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 de referencia, y luego, por selección de las condiciones adecuadas, la sonda y la secuencia de referencia se hibridan selectivamente, o unen entre sí para formar una molécula doble. Una molécula de ácido nucleico que es capaz de hibridarse selectivamente a una secuencia referencia en condiciones moderadamente rigurosas de hibridación generalmente se hibrida en condiciones que permiten la detección de una secuencia de ácido nucleico blanco de al menos aproximadamente 10-14 nucleótidos de longitud que tiene al menos aproximadamente 70% de identidad de secuencia con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación rigurosas generalmente permiten la detección de secuencias de ácido nucleico blanco de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencia superior a aproximadamente 90-95% con la secuencia de la sonda seleccionada de ácido nucleico. Las condiciones de hibridación útiles para hibridar la secuencia de sonda/referencia, donde la sonda y la secuencia de referencia tienen un grado específico de identidad de secuencia, se pueden determinar tal como se conoce en la técnica (véase, *por ejemplo*, Nucleic Acid Hybridization: A Practical Approach, editores B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

Las condiciones de hibridación son bien conocidas por los expertos en la técnica. La rigurosidad de hibridación se refiere al grado hasta el cual las condiciones de hibridación desfavorecen la formación de los híbridos que contienen nucleótidos no coincidentes, la mayor rigurosidad se correlaciona con una menor tolerancia para híbridos no coincidentes. Los factores que afectan la rigurosidad de la hibridación son bien conocidos por los expertos en la técnica e incluyen, sin limitaciones, temperatura, pH, fuerza iónica, y concentración de disolventes orgánicos tales como, por ejemplo, formamida y dimetilsulfóxido. Como es bien sabido por los expertos en la técnica, la rigurosidad

de la hibridación aumenta con temperaturas más altas, fuerza iónica menor y concentraciones de disolventes más bajas.

Con respecto a las condiciones de rigurosidad para la hibridación, es bien conocido en la técnica que se pueden emplear numerosas condiciones equivalentes para establecer una rigurosidad particular al variar, 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 hibridación (por ejemplo, sulfato de dextrano y polietilenglicol), parámetros de temperatura y tiempo de la reacción de hibridación, así como variadas condiciones de lavado. La elección de un conjunto particular de condiciones de hibridación se selecciona según procedimientos estándar en la técnica (véase, por ejemplo, Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, (1989) Cold Spring Harbor, N.Y.).

"Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. A los fines de la presente descripción, "recombinación homóloga (HR)" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de las rupturas de la cadena doble en las células. Este proceso requiere la homología de secuencia de nucleótidos, que usa una molécula "donante" para reparar el molde de una molécula "blanco" (es decir, la que experimentó la ruptura de la cadena doble), y es conocida de manera diversa como "conversión génica sin entrecruzamiento" o "conversión génica de tracto corto", porque lleva a la transferencia de información genética de la donante al blanco. Sin pretender estar ligado por cualquier teoría particular, dicha transferencia puede implicar la corrección de la falta de coincidencia de ADN heterodúplex que se forma entre el blanco roto y donante, y/o "apareamiento de cadenas dependiente de la síntesis" en que se usa la donante para resintetizar la información genética que pasará a formar parte del blanco, y/o procesos relacionados. Tal HR especializada a menudo produce una alteración de la secuencia de la molécula blanco de manera tal que parte o la totalidad de la secuencia de polinucleótidos donante se incorpora al polinucleótido blanco.

"Escisión" se refiere a la ruptura del esqueleto covalente de una molécula de ADN. La escisión puede ser iniciada por diversos procedimientos que incluyen, sin limitaciones, la hidrólisis enzimática o química de un enlace fosfodiéster. La escisión de la cadena simple y escisión de la cadena doble son posibles, y la escisión de la cadena doble puede ocurrir como resultado de dos eventos de escisión de cadena simple bien diferenciados. La escisión del ADN puede generar la producción de extremos romos o extremos escalonados. En ciertas formas de realización, se usan polipéptidos de fusión para la escisión dirigida del ADN de cadena doble.

Un "dominio de escisión" comprende una o más secuencias de polipéptidos que posee actividad catalítica para la escisión del ADN. Un dominio de escisión puede estar contenido en una cadena de polipéptido única o la actividad de escisión puede ser el resultado de la asociación de dos (o más) polipéptidos.

Un "semidominio de escisión" es una secuencia de polipéptido que, en conjunción con un segundo polipéptido (sea idéntico o diferente) forma un complejo que tiene actividad de escisión (por ejemplo, actividad de escisión de cadena doble).

Cromatina" es la estructura de nucleoproteína que comprende el genoma celular. La cromatina celular comprende el ácido nucleico, principalmente ADN, y proteína, que incluye las histonas y proteínas cromosómicas no histona. La mayor parte de la cromatina celular eucariota existe en la forma de nucleosomas, en el que 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; y ADN ligador (de longitud variable según el organismo) se extiende entre los núcleos del nucleosoma. Una molécula de histona H1 generalmente se asocia con el ADN ligador. A los fines de la presente descripción, el término "cromatina" significa que abarca todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye tanto la cromatina cromosómica como episómica.

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

Un "episoma" es un ácido nucleico en replicación, complejo de nucleoproteína u otra estructura que comprende un ácido nucleico que no forma parte del cariotipo cromosómico de una célula. Los ejemplos de episomas incluyen plásmidos y ciertos genomas virales.

Una "región accesible" es un sitio de la cromatina celular en la cual un sitio blanco presente en el ácido nucleico se puede unir a una molécula exógena que reconoce el sitio blanco. Sin pretender estar limitado por cualquier teoría particular, se cree que una región accesible es la que no está empaquetada en una estructura de nucleosoma. La estructura diferenciada de una región accesible a menudo se puede detectar por su sensibilidad a sondas químicas y enzimáticas, por ejemplo, nucleasas.

Un "sitio blanco" o "secuencia blanco" es una secuencia de ácido nucleico que define una porción de un ácido nucleico al que se unirá la molécula de unión, siempre que existan suficientes condiciones de unión. Por ejemplo, la secuencia 5'–GAATTC–3' es un sitio blanco para la endonucleasa *Eco RI* de restricción.

5 Una molécula "exógena" es una molécula que normalmente no está presente en una célula, pero que se puede introducir en la célula mediante uno o más procedimientos genéticos, bioquímicos o de otro tipo. La "presencia normal en la célula" se determina con respecto a la particular etapa de desarrollo y las condiciones ambientales de la célula. En consecuencia, por ejemplo, una molécula que está presente en las células solo durante las primeras etapas de desarrollo embrionario del músculo es una molécula exógena con respecto a una célula muscular adulta. De modo similar, una molécula inducida por shock de calor es una molécula exógena con respecto a una célula no sometida al shock de calor. Una molécula exógena puede comprender, por ejemplo, una versión funcional de una molécula endógena que funciona mal o una versión que funciona mal de una molécula endógena que funciona normalmente.

15 Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como la generada por un proceso de química combinatorio, o una macromolécula tal como una proteína, ácido nucleico, hidrato de carbono, lípido, glucoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprende una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser de cadena simple o doble; pueden ser lineales, ramificados o circulares; y pueden tener cualquier longitud. Los ácidos nucleicos incluyen los que tienen capacidad para formar dúplex, así como ácidos nucleicos formadores de tríplex. Véase por ejemplo, las patentes U.S. Nros. 5.176.996 y 5.422.251. Las proteínas incluyen, pero sin limitación, proteínas de unión a ADN, factores de transcripción, factores de remodelación de cromatina, proteínas de unión a ADN metiladas, polimerasas, metilasas, demetilasas, acetilasas, desacetilasas, quinasas, fosfatasas, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicinas.

20 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 viral infectante, una cadena T *Agrobacterium tumefaciens*, un plásmido o episoma introducido en la célula o un cromosoma que normalmente no está presente en la célula. Los procedimientos para la introducción de moléculas exógenas en las células son conocidos por los expertos en la técnica e incluyen, pero sin limitación, transferencia mediada por lípidos (es decir, liposomas, que incluye lípidos neutros y catiónicos), electroporación, inyección directa, fusión celular, bombardeo de partículas, coprecipitación de fosfato de calcio, transferencia mediada por DEAE–dextrano y transferencia mediada por vector viral. La molécula exógena es una molécula no de planta, por ejemplo, un anticuerpo de mamífero (*por ejemplo*, humano o humanizado).

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

30 Una molécula de "fusión" es una molécula en la cual dos o más moléculas de subunidades están ligadas, con preferencia, covalentemente. Las moléculas de subunidades pueden ser el mismo tipo químico de molécula, o pueden ser distintos tipos químicos de moléculas. Los ejemplos del primer tipo de molécula de fusión incluyen, pero sin limitación, proteínas de fusión (por ejemplo, una ZFN que comprende una fusión entre un dominio de unión 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). Los ejemplos del segundo tipo de molécula de fusión incluyen, pero sin limitación, una fusión entre un ácido nucleico formador de triplete y un polipéptido, y una fusión entre un ligador del surco menor y un ácido nucleico.

35 La expresión de una proteína de fusión en una célula puede ser el resultado de la administración de la proteína de fusión a la célula o por administración de un polinucleótido que codifica la proteína de fusión a una célula, en que el polinucleótido se transcribe, y la transcripción se traduce para generar la proteína de fusión. En la expresión de una proteína en una célula también pueden intervenir el transempalme, escisión del polipéptido y ligación del polipéptido. Los procedimientos para la administración de polinucleótidos y polipéptidos a las células se presentan en otra parte de la presente descripción.

40 Un "gen", a los fines de la presente descripción, incluye una región 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 estas secuencias reguladoras adyacentes a las secuencias codificadoras y/o transcritas. En consecuencia, un gen incluye, pero no necesariamente se limita a secuencias de promotores, terminadores, secuencias reguladoras de la traducción tales como sitios de unión de ribosomas y sitios de entrada internos del ribosoma, potenciadores, silenciadores, aisladores, elementos limitantes, orígenes de replicación, sitios de unión de la matriz y regiones de control del locus.

45 La "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 de la transcripción directa de un gen (por ejemplo, ARNm, ARNt, ARNr, ARN

antisentido, ribozima, ARNsh, micro ARN, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por traducción de un ARNn. Los productos génicos también incluyen ARN modificados por procesos tales como colocación de capuchón, poliadenilación, metilación y edición, y proteínas modificadas, por ejemplo, por metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristilación y glucosilación.

- 5 La "modulación" de la expresión génica se refiere a un cambio de la actividad de un gen. La modulación de la expresión puede incluir, pero sin limitación, la activación génica y la represión génica.

Las células de "planta" incluyen, pero sin limitación, células de planas monocotiledóneas (monocot) o dicotiledóneas (dicot). Los ejemplos no limitantes de monocotiledóneas incluyen plantas de cereales tales como maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, ananá, cebolla, banana y coco. Los ejemplos no limitantes de dicotiledóneas incluyen tabaco, tomate, girasol, algodón, remolacha azucarera, papa, lechuga, melón, poroto de soja, canola (semilla de colza), y alfalfa. Las células de plantas pueden provenir de cualquier parte de la planta y/o de cualquier etapa de desarrollo de la planta. Por lo tanto, las células de planta pueden ser células de las semillas de la planta. En algunas realizaciones, la planta o célula de planta es, o se deriva de, una planta implicada en la producción de aceites vegetales para usos comestibles y/o industriales (es decir, una "planta productora de aceite").

10 Los ejemplos de plantas productoras de aceite incluyen, pero sin limitación *Brassica sp.* (por ejemplo, *Brassica napus* que incluye cultivares de canola), maíz, soja, crame, mostaza, semilla de ricino, maní, sésamo, algodón, semilla de lino, cártamo, aceite de palma, lino, girasol y coco.

Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificadora dentro o adyacente a un gen, en la que es deseable unirse a una molécula exógena. La unión puede tener por finalidad la escisión de ADN dirigida y/o la recombinación dirigida. Una región de interés puede estar presente por ejemplo, en un cromosoma, un episoma, un genoma de organela (por ejemplo, mitocondrial, cloroplástico), o un genoma viral infeccioso. Una región de interés puede estar dentro de la región codificadora de un gen, dentro de regiones no codificadoras transcritas tales como, por ejemplo, secuencias líderes, secuencias de remolque o intrones, o dentro de regiones no transcritas, sean corriente arriba o corriente abajo respecto de la región codificadora. Una región de interés puede ser tan pequeña como un único par de nucleótidos o de hasta

20 25.000 pares de nucleótidos de longitud, o cualquier valor entero de pares de nucleótidos.

Los términos "unión operativa" y "unido operativamente" (o "unido operablemente") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (elementos como elementos de secuencia), en la que los componentes se disponen de manera tal que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce sobre al menos uno de los demás componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unido operativamente a una secuencia codificadora si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificadora en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora transcripcional generalmente está unida operativamente

30 en cis con una secuencia codificadora, pero no necesita estar directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está unida operativamente a una secuencia codificadora, aun cuando no sean contiguas.

Con respecto a los polipéptidos de fusión, el término "unido operativamente" se puede referir al hecho de que cada uno de los componentes realiza la misma función en unión con el otro componente, como si no estuvieran unidos de este modo. Por ejemplo, con respecto a un polipéptido de fusión en el cual un dominio de unión al ADN ZFP se fusiona a un dominio regulador, el dominio de unión al ADN ZFP y el dominio regular están en unión operativa si, en el polipéptido de fusión, la porción del dominio de unión al ADN ZFP es capaz de unirse a su sitio blanco y/o su sitio de unión, mientras que el dominio regulador es capaz de regular la expresión de ADN en la cercanía del sitio blanco.

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 retiene 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 correspondiente molécula nativa, y/o puede contener una o más sustituciones de aminoácidos o nucleótidos. Los procedimientos para determinar la función de un ácido nucleico (por ejemplo, función codificadora, capacidad para hibridarse a otro ácido nucleico) son bien conocidos en la técnica. De modo similar, los procedimientos para determinar la función proteica son bien conocidos. Por ejemplo, la función de unión a ADN de un polipéptido se puede determinar, por ejemplo, por unión a un filtro, cambio de movilidad electroforética o ensayos de inmunoprecipitación. La escisión de ADN se puede analizar por electroforesis en gel. Véase Ausubel *et al.*, *supra*. La capacidad de una proteína para interactuar con otra proteína se puede determinar, por ejemplo, por coimmunoprecipitación, ensayos de dos híbridos o complementación, tanto genética como

45 50 bioquímica. Véase por ejemplo, Fields *et al.* (1989) *Nature* **340**:245-246; patente U.S. N° 5.585.245 y PCT WO 98/44350.

Como se usa en la presente memoria, "ácidos grasos saturados" incluyen pero sin limitación, ácidos láurico (C12:0), mirístico (C14:0), palmítico (C16:0) y esteárico (C18:0). De modo similar, los "ácidos grasos monoinsaturados" y "ácidos grasos poliinsaturados" incluyen, pero sin limitación ácidos grasos C18 tal como ácido oleico (C18:1), ácido linoleico reducido (C18:2) y ácido linoléico reducido (C18:3).

60

Sitios blanco

Los procedimientos y composiciones descritos incluyen las proteínas de fusión que comprenden un dominio de unión al ADN (*por ejemplo*, ZFP) y un dominio regulador o dominio de escisión (nucleasa) (o un semidominio de escisión), en que el dominio de unión al ADN (*por ejemplo*, dominio de dedos de cinc), mediante la unión a una secuencia en la cromatina celular en un gen involucrado en la síntesis de ácidos grasos, dirige la actividad del dominio regulador transcripcional o dominio de escisión (o semidominio de escisión) en la proximidad de la secuencia y, en consecuencia, modula la transcripción o induce la escisión en la proximidad de la secuencia blanco.

Como se expone en otra parte en esta descripción, un dominio de dedos de cinc se puede manipular genéticamente para unirse prácticamente a cualquier secuencia deseada. Por consiguiente, después de identificar una región de interés que contiene una secuencia en la que se desea la regulación, escisión, o recombinación del gen, uno o más dominios de dedos de cinc se pueden manipular genéticamente para unir una o más secuencias en la región de interés.

La selección de un sitio blanco en una región genómica de interés en la cromatina celular de cualquier gen involucrado en biosíntesis de ácidos grasos (*por ejemplo*, véase Figura 1) para la unión por un dominio de dedos de cinc (*por ejemplo*, un sitio blanco) se puede obtener por ejemplo, de acuerdo con procedimientos descritos en la en la Patente en copropiedad U.S.:Nº No. 6.453.242 (17 de septiembre de 2002), que también describe procedimientos para diseñar ZFP para que se unan 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 usar para la selección de un sitio blanco. Por consiguiente, se puede usar cualquier medio para la selección del sitio blanco en los procedimientos reivindicados

Los sitios blanco se componen generalmente de una pluralidad de subsitios blanco adyacentes. Un subsitio blanco se refiere a la secuencia (usualmente un triplete de nucleótidos, o un cuadruplete de nucleótidos que se puede superponer en un nucleótido con un cuadruplete adyacente) unido por un dedos de cinc individual. Véase, por ejemplo, WO 02/077227. Si la cadena con la que una proteína de dedos de cinc hace mayor contacto se denomina "cadena de reconocimiento primario" o "cadena de contacto primario" de la cadena blanco, algunas proteínas de dedos de cinc se unen a un triplete de tres bases en la cadena blanco y una cuarta base en la cadena no blanco. Un sitio blanco generalmente tiene una longitud de al menos 9 nucleótidos y, por consiguiente, se une mediante un dominio de unión de dedos de cinc que comprende al menos tres dedos de cinc. Sin embargo, también es posible la unión de, por ejemplo, un dominio de unión de 4 dedos a un sitio blanco de 12 nucleótidos, un dominio de unión de 5 dedos a un sitio blanco de 15 nucleótidos o un dominio de unión de 6 dedos a un sitio blanco de 18 nucleótidos. Como será evidente, la unión de los dominios de unión más grandes (*por ejemplo*, 7, 8, 9 dedos y más) a los sitios blanco más largos también es posible.

No es necesario que un sitio blanco sea múltiplo de tres nucleótidos. Por ejemplo, en los casos en que se producen interacciones de cadena cruzada (véase, *por ejemplo*, patente U.S. 6.453.242 WO 02/077227), uno o más de los dedos de cinc individuales de un dominio de unión de múltiples dedos se puede unir a subsitios de cuadruplete superpuestos. Como resultado, una proteína de tres dedos puede unir una secuencia de 10 nucleótidos, en la que el décimo nucleótido es parte de un cuadruplete unido por un dedo terminal, una proteína de cuatro dedos puede unir una secuencia de 13 nucleótidos, en la que el decimotercero nucleótido es parte de un cuadruplete unido por un dedos terminal, etc.

La longitud y naturaleza de secuencias de ligador de aminoácidos entre los dedos de cinc individuales en un dominio de unión de dedos múltiples también afecta a la unión a una secuencia blanco. Por ejemplo, la presencia de un llamado "ligador no canónico", "ligador largo" o "ligador estructurado" entre los dedos de cinc adyacentes en un dominio de unión de dedos múltiples puede permitir que los dedos se unan a los subsitios que no son inmediatamente adyacentes. Los ejemplos no limitantes de tales ligadores se describen, por ejemplo, en las patentes U.S Nros 6.479.626 WO 01/53480. En consecuencia, uno o más subsitios, en un sitio blanco para un dominio de unión de dedos de cinc, puede estar separado entre sí por 1, 2, 3, 4, 5 o más nucleótidos. Para proporcionar un solo ejemplo, un dominio de unión de cuatro dedos se puede unir a un sitio blanco de 13 nucleótidos que comprende, en secuencia, dos subsitios de 3 nucleótidos contiguos, un nucleótido interviniente, y dos subsitios de tripletes contiguos. Véase, también, la solicitud U.S. Nº. 61/130.099 para las composiciones y procedimientos para ligar las nucleasas artificiales que se unen a sitios blanco separados por diferentes números de nucleótidos. La distancia entre las secuencias (por ejemplo, sitios blanco) se refiere al número de nucleótidos o pares de nucleótidos que intervienen entre dos secuencias, medida desde los bordes de las secuencias más cercanas entre sí.

En ciertas realizaciones, se diseñan ZFP con función del factor de transcripción, Para la función de factor de transcripción, la proximidad de unión simple y suficiente al promotor es todo lo que generalmente se necesita. La ubicación exacta respecto del promotor, orientación, y dentro de los límites, la distancia no importa mucho. Esta característica permite una flexibilidad considerable en la elección de los sitios blanco para la construcción de factores de transcripción artificiales. El sitio blanco reconocido por ZFP puede ser cualquier sitio adecuado en el gen blanco que permita la activación o represión de la expresión génica, y una ZFP opcionalmente ligada a un dominio regulador. Los sitios blanco preferidos incluyen las regiones adyacentes, corriente abajo o corriente arriba del sitio de inicio de transcripción. Además, los sitios blanco que se encuentran en las regiones potenciadoras, sitios

represores, sitios de pausa de 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 a p53), sitios en la región que codifica ADNc o en una región codificadora de la marca de la secuencia (EST) expresada.

5 En otras realizaciones, se diseñan ZFP con actividad de nucleasa. La expresión de una ZFN que comprende una proteína de fusión que comprende un dominio de unión de dedos de cinc y un dominio de escisión (o de dos proteínas de fusión, cada una que comprende un dominio de unión de dedos de cinc y un semidominio de escisión), en una célula, efectúa la escisión en la proximidad de la secuencia blanco. En ciertas realizaciones, la escisión depende de la unión de dos moléculas de fusión del dominio de dedos de cinc/semidominio de escisión para separar los sitios blanco. Los dos sitios blanco pueden estar en las cadenas de ADN opuestas, o alternativamente, ambos
10 sitios blanco pueden estar en la misma cadena de ADN.

Dominios de unión al ADN

Cualquier dominio de unión al ADN se puede usar en la práctica de la presente invención. En ciertas realizaciones, el dominio de unión al ADN comprende un dominio de unión de dedos de cinc de uno o más dedos de cinc. Miller *et al.* (1985) *EMBO J.*4:1609-1614; Rhodes (1993) *Scientific American* Feb.:56-65; patente U.S. N° 6.453.242.
15 Generalmente, un dominio de dedos de cinc único tiene aproximadamente 30 aminoácidos de longitud. Los estudios estructurales han demostrado que cada dominio de dedos de cinc (motivo) contiene dos hojas beta (sostenidas en un giro beta que contiene los dos residuos cisteína no variantes) y una hélice alfa (que contiene los dos residuos histidina no variantes), que se mantienen en una particular conformación debido a la coordinación de un átomo de cinc mediante las dos cisteínas y las dos histidinas.

20 Los dedos de cinc incluyen tanto los dedos de cinc canónicos C₂H₂ (es decir, aquellos en los cuales el ión cinc está coordinado por dos residuos cisteína y dos residuos histidina) y dedos de cinc no canónicos, tales como, por ejemplo, los dedos de cinc C₃H y C₂HC (aquellos en los cuales el ión cinc es coordinado por tres residuos de cisteína y un residuo de histidina). Véase también Publicación de patente U.S. N° 2008/0182332) y dedos de cinc C₄ (aquellos en los que el ion de cinc está coordinado por cuatro residuos de cisteína). Véase también el documento
25 WO 02/057293.

Los dominios de unión de dedos de cinc se pueden manipular genéticamente para unirse a una secuencia de elección. Véase, por ejemplo, Beerli *et al.* (2002) *Nature Biotechnol.* 20:135-141; Pabo *et al.* (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan *et al.* (2001) *Nature Biotechnol.* 19:656-660; Segal *et al.* (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo *et al.* (2000) *Curr. Opin. Struct. Biol.* 10:411-416. Un dominio de unión de dedos de cinc manipulado genéticamente (o no natural) puede tener una nueva especificidad de unión, en comparación con una proteína de dedos de cinc natural. Los procedimientos de manipulación genética incluyen, sin limitaciones, un diseño racional y diversos tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden tripletes (o cuatripletas) de secuencias de nucleótidos y secuencias de aminoácidos de dedos de cinc individuales, en los cuales cada secuencia de nucleótidos triplete o cuatriplete se asocia con una o más secuencias de aminoácidos de dedos de cinc que se unen a la secuencia de triplete o cuatriplete particular. Véase, por ejemplo,
30 Patentes U.S. en copropiedad 6.453.242 y 6.534.261. Se describen procedimientos de diseño adicionales, por ejemplo, en las patentes U.S. 6.746.838; 6.785.613; 6.866.997; y 7.030.215.

Los ejemplos de procedimientos de selección, que incluyen sistemas de despliegue en fago y de dos híbridos, se describen en las patentes U.S. 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 WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237.
40

El mejoramiento de la especificidad de unión de los dominios de unión de los dedos de cinc se ha descrito, por ejemplo, en la patente U.S. en copropiedad N° 6.794.136.

Debido a un dedo de cinc individual se une a una secuencia de tres nucleótidos (es decir, triplete) (o una secuencia de cuatro nucleótidos que se puede superponer, por un nucleótido, con el sitio de unión de cuatro nucleótidos de un dedo de cinc adyacente), la longitud de una secuencia para la cual un dominio de unión de un dedo de cinc se manipula genéticamente para unir (por ejemplo, una secuencia blanco) determinará la cantidad de dedos de cinc en un dominio de unión de dedos de cinc manipulado genéticamente. Por ejemplo, para las ZFP en las cuales los motivos de dedo no se unen a los subsitios superpuestos, una secuencia blanco de seis nucleótidos se une a un dominio de unión de dos dedos; una secuencia blanco de nueve nucleótidos se une a un dominio de unión de tres dedos, *etc.* Como se indica en la presente memoria, los sitios de unión para dedos de cinc individuales (es decir, subsitios) en un sitio blanco no necesitan ser contiguos, sino que pueden estar separados por uno o varios nucleótidos, según la longitud y la naturaleza de las secuencias de aminoácidos entre los dedos de cinc (es decir, los ligadores entre dedos) en un dominio de unión de dedos múltiples.
45
50

En un dominio de unión de dedos de cinc de dedos múltiples, los dedos de cinc adyacentes pueden estar separados por secuencias ligadoras de aminoácidos de aproximadamente 5 aminoácidos (los denominados ligadores entre dedos "canónicos") o, alternativamente, por uno o más ligadores no canónicos. Véase, por ejemplo, las patentes en copropiedad U.S. Nros 6.453.242 y 6.534.261. Para los dominios de unión de dedos de cinc manipulados genéticamente que comprenden más de tres dedos, la inserción de ligadores entre dedos más largos ("no
55

canónicos”) entre algunos de los dedos de cinc puede ser preferida ya que puede incrementar la afinidad y/o especificidad de la unión mediante el dominio de unión. Véase, por ejemplo, la patente U.S. N° 6,479,626 y WO 01/53480. En consecuencia, los dominios de unión de dedos de cinc de dedos múltiples también se pueden caracterizar con respecto a la presencia y la ubicación de ligadores entre dedos no canónico. Por ejemplo, un dominio de unión de dedos de cinc de seis dedos que comprende tres dedos (unidos por dos ligadores entre dedos canónicos), un ligador largo y tres dedos adicionales (unidos por dos ligadores entre dedos canónicos) se indica como una configuración 2x3. De modo similar, un dominio de unión que comprende dos dedos (con un ligador canónico entre ellos), un ligador largo y dos dedos adicionales (unidos por un ligador canónico) se indica como una configuración 2x2. Una proteína que comprende tres unidades de dos dedos (en que cada uno de dos dedos se unen mediante un ligador canónico), y en que cada unidad de dos dedos se une a la unidad de dos dedos adyacentes con un ligador largo, se denomina como una configuración 3x2.

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

Como se discutió previamente, un subsitio blanco es una secuencia de tres a cuatro nucleótidos que se une por medio de un dedo de cinc único. Para determinados fines, una unidad de dos dedos se indica como módulo de unión. Un módulo de unión se puede obtener, por ejemplo, mediante la selección de dos dedos adyacentes en el contexto de una proteína de dedos múltiples (generalmente tres dedos) que se unen a una secuencia blanco de seis nucleótidos particular. De modo alternativo, los módulos se pueden construir mediante el ensamblaje de los dedos de cinc individuales. Véase también WO 98/53057 y WO 01/53480.

En una realización, en la presente memoria se describe un dominio de unión de dedos de cinc que comprende una secuencia de aminoácidos como se muestra en la Tabla 1. En otra realización, la descripción proporciona un polinucleótido que codifica un dominio de unión de dedos de cinc, en el que el dominio de unión de dedos de cinc comprende una secuencia de aminoácidos como se muestra en la Tabla 1.

De modo alternativo, el dominio de unión al ADN se puede derivar de una nucleasa. Por ejemplo, se conocen las secuencias de reconocimiento de las endonucleasas de asentamiento y meganucleasas tales como I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-Scell, I-PpoI, I-ScellIII, I-Crel, I-TeV, I-TeVII y I-TeVIII. Véase también la Patente U.S. N° 5.420.032; patente U.S. 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 and the New England Biolabs catalogue. Además, la especificidad de unión de ADN de las endonucleasas de asentamiento y meganucleasas se pueden manipular genéticamente para unirse a sitios blanco 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; Publicación de patente U.S. N° 20070117128.

Dominios reguladores

Los dominios de unión a ADN (por ejemplo, ZFP) descritos en la presente memoria se pueden asociar opcionalmente con dominios reguladores para la modulación de la expresión génica. La ZFP se puede asociar de forma covalente o no covalente con uno o más dominios reguladores, alternativamente dos o más dominios reguladores, con los dos o más dominios que son dos copias del mismo dominio o dos dominios diferentes. Los dominios reguladores se pueden unir covalentemente a la ZFP, por ejemplo, a través de un ligador de aminoácidos, como parte de una proteína de fusión. Las ZFP también se pueden asociar con un dominio regulador a través de un dominio de dimerización no covalente, por ejemplo, un cierre 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 se puede asociar con la ZFP en cualquier posición adecuada, que incluye el extremo C o N de la ZFP.

Los dominios reguladores comunes para la 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 del ADN y sus factores y modificadores asociados; proteínas asociadas a la cromatina y sus modificadores (por ejemplo, quinasas, acetilasas y desacetilasas); y enzimas que modifican el ADN (por

ejemplo, metiltransferasas, topoisomerasas, helicasas, ligasas, quinasas, fosfatasas, polimerasas, endonucleasas) y sus factores y modificadores asociados.

Los polipéptidos del factor de transcripción a partir de los cuales se puede obtener un dominio regulador incluyen los 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 involucrado en la transcripción; los factores de transcripción en general se revisan en Barnes and Adcock, *Clin. Exp. Allergy* 25 Suppl. 2:46-9 (1995) and Roeder, *Methods Enzymol.* 273:165-71(1996)). Los factores de transcripción del receptor de la hormona nuclear se describen, por ejemplo, en Rosen et al., *J. Med. Chem.* 38:4855-74 (1995). La familia de factores de transcripción C/EBP se revisa en Wedel et al., *Immunobiology* 193:171-85 (1995). Los coactivadores y correpresores que median la regulación de la transcripción por receptores de hormonas nucleares se revisan en, por ejemplo, Meier, *Eur. J. Endocrinol.* 134(2):158-9 (1996); Kaiser et al., *Trends Biochem. Sci.* 21:342-5 (1996); and Utley et al., *Nature* 394:498-502 (1998)). Los factores de transcripción GATA, que están involucrados en la regulación de la hematopoyesis, se describen en, por ejemplo, Simon, *Nat. Genet.* 11:9-11 (1995); Weiss et al., *Exp. Hematol.* 23:99-107. La proteína de unión a la caja TATA (TBP) y sus polipéptidos TAP asociados (que incluyen TAF30, TAF55, TAF80, TAF10, TAF150 y TAF250) se describen en Goodrich and Tjian, *Curr. Opin. Cell Biol.* 6:403-9 (1994) and Hurley, *Curr. Opin. Struct. Biol.* 6:69-75 (1996). La familia STAT de factores de transcripción se revisa en, por ejemplo, Barahmand-Pour et al., *Curr. Top. Microbiol. Immunol.* 211:121-8 (1996). Los factores de transcripción involucrados en la enfermedad se revisan en Aso et al., *J. Clin. Invest.* 97:1561-9 (1996).

En un aspecto, el dominio de represión de KRAB de la proteína KRX-1 humana se usa como un represor transcripcional (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 usa con KRAB (Friedman et al., *Genes Dev.* 10:2067-2078 (1996)). Alternativamente, KAP-1 se puede usar solo con una ZFP. Otros factores de transcripción y dominios del factor de transcripción preferidos que actúan como represores transcripcionales incluyen MAD (ver, *por ejemplo*, 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); and Cultraro et al., *Mol. Cell. Biol.* 17:2353-2359 (1997)); FKHR (cabeza de tenedor en el gen de rhabdosarcoma; Ginsberg et al., *Cancer Res.* 15:3542-3546 (1998); Epstein et al., *Mol. Cell. Biol.* 18:4118-4130 (1998)); EGR-1 (producto génico 1 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)); el dominio represor del factor represor 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 anfífilico ERF3 (factor de respuesta al etileno 3), EAR (Ohta, M., et al. (2001), *Célula de planta* 13:1959-1968).

En un aspecto, el dominio de activación de VP16 HSV se usa como un activador de transcripción (véase, por ejemplo, Hagmann et al., *J. Virol.* 71:5952-5962 (1997)). Otros factores de transcripción preferidos que pueden 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 and Barik, *J. Virol.* 72:5610-5618 (1998) and Doyle and Hunt, *Neuroreport* 8:2937-2942 (1997));; EGR-1 (producto de gen de respuesta al crecimiento temprano 1; Yan et al., *Proc. Natl. Acad. Sci USA* 95:8298-8303 (1998); and Liu et al., *Cancer Gene Ther.* 5:3-28 (1998)); y la proteína reguladora de la vía biosintética de antocianina del maíz, C1 (S.A. Goff, y col., (1991), *Genetics and Development* 5:298-309).

Las quinasas, fosfatasas y otras proteínas que modifican los polipéptidos involucrados en la regulación de genes también son útiles como dominios reguladores para las ZFP. Tales modificadores a menudo están involucrados en la activación o desactivación de la transcripción mediada, por ejemplo, por hormonas. Las quinasas involucradas 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 Boulikas, *Crit. Rev. Eukaryot. Gene Expr.* 5:1-77 (1995), mientras que se revisan las fosfatasas en, por ejemplo, Schonthal and Semin, *Cancer Biol.* 6:239-48 (1995). Las tirosina quinasas nucleares se describen en Wang, *Trends Biochem. Sci.* 19:373-6 (1994).

Como se describe, también se pueden obtener dominios útiles 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 en, por ejemplo, Cooper, *Oncogenes*, 2nd ed., The Jones and Bartlett Series in Biology, Boston, Mass., Jones and Bartlett Publishers, 1995. Los factores de transcripción se revisan en Wasylk et al., *Eur. J. Biochem.* 211:7-18 (1993) and Crepieux et al., *Crit. Rev. Oncog.* 5:615-38 (1994). Los oncogenes Myc se revisan en, por ejemplo, Ryan et al., *Biochem. J.* 314:713-21 (1996). Los factores de transcripción jun y fos se describen en, por ejemplo, The Fos and Jun Families of Transcription Factors, Angel and Herrlich, eds. (1994). El oncogén max se revisa en Hurlin et al., *Cold Spring Harb. Symp. Quant. Biol.* 59:109-16. La familia del gen 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 ZFP pueden incluir dominios reguladores obtenidos de las enzimas de reparación de ADN y sus factores y modificadores asociados. Los sistemas de reparación de ADN se revisan en, por ejemplo, Vos, *Curr. Opin. Cell Biol.* 4:385-95 (1992); Sancar, *Ann. Rev. Genet.* 29:69-105 (1995); Lehmann, *Genet. Eng.* 17:1-19 (1995); and Wood, *Ann. Rev. Biochem.* 65:135-67 (1996). Los sistemas de reordenamiento de ADN y sus factores y modificadores asociados también se pueden usar como dominios reguladores (véase, *por ejemplo*, Gangloff et al., *Experientia* 50:261-9 (1994); Sadowski, *FASEB J.* 7:760-7 (1993)).

De forma similar, los dominios reguladores se pueden derivar de enzimas que modifican el ADN (por ejemplo, ADN metiltransferasas, topoisomerasas, helicasas, ligasas, quinasas, fosfatasas, polimerasas) y sus factores y modificadores asociados. Las helicasas se revisan en Matson et al., *Bioessays*, 16:13-22 (1994), y las metiltransferasas se describen en Cheng, *Curr. Opin. Struct. Biol.* 5:4-10 (1995). Las proteínas asociadas a cromatina y sus modificadores (por ejemplo, quinasas, acetilasas y desacetilasas), tales como histona desacetilasa (Wolffe, *Science* 272:371-2 (1996)) también son útiles como dominios para la adición a la ZFP de elección. En un aspecto preferido, el dominio regulador es una ADN metil transferasa que actúa como un represor transcripcional (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 and Caiafa, *J. Biol. Chem.* 273:16517-16520 (1998)). En otra realización preferida, las endonucleasas tales como *Fok1* se usan como represores transcripcionales, que actúan a través de la escisión génica (véase, *por ejemplo*, WO95/09233 y PCT/US94/01201).

Los factores que controlan la estructura, el movimiento y la localización de la cromatina y el ADN y sus factores y modificadores asociados; los factores derivados de microbios (por ejemplo, procariotas, eucariotas y virus) y los factores que los asocian con o modifican también se pueden usar para obtener proteínas quiméricas. En un aspecto, las recombinasas e integrasas se usan como dominios reguladores. En un aspecto, la histona acetiltransferasa se usa como un activador transcripcional (véase, por ejemplo, Jin and 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, se usa la histona desacetilasa como un represor transcripcional (véase, *por ejemplo*, Jin and Scotto, *Mol. Cell. Biol.* 18:4377-4384 (1998); Syntichaki and 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)).

Se pueden incluir dominios ligadores entre dominios polipeptídicos, por ejemplo, entre dos ZFP o entre una ZFP y un dominio regulador. Tales ligadores son típicamente secuencias polipeptídicas, tales como secuencias de poli-gly de entre aproximadamente 5 y 200 aminoácidos. Los ligadores pueden ser subsecuencias de aminoácidos flexibles o rígidas que se sintetizan como parte de una proteína de fusión recombinante. Véase, por ejemplo, la Patente U.S. N° 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).

Alternativamente, los ligadores flexibles se pueden diseñar racionalmente usando un programa de computadora capaz de modelar tanto los sitios de unión al ADN como los péptidos mismos (Desjarlais and Berg, *Proc. Nat. Acad. Sci. USA* 90:2256-2260 (1993), Desjarlais and Berg, *Proc. Nat. Acad. Sci. USA* 91:11099-11103 (1994) o mediante procedimientos de despliegue en fagos.

Un ligador químico se usa para conectar secuencias de dominio producidas de manera sintética o recombinante. Tales ligadores flexibles son conocidos por los expertos en la técnica. Por ejemplo, los ligadores de poli(etilenglicol) están disponibles en Shearwater Polymers, Inc. Huntsville, Alabama. Estos ligadores tienen opcionalmente uniones amida, uniones sulfhidrilo o uniones heterofuncionales. Además de la unión covalente de las ZFP con los dominios reguladores, se pueden usar procedimientos no covalentes para producir moléculas con ZFP asociadas con dominios reguladores.

Dominios de escisión

Como se indicó anteriormente, el dominio de unión al ADN también se puede asociar con un dominio de escisión para formar una nucleasa de dedos de cinc (ZFN). La porción del dominio de escisión de las proteínas de fusión descritas en la presente se puede obtener de cualquier endonucleasa o exonucleasa. Los ejemplos de endonucleasas de las cuales se puede derivar un dominio de escisión incluyen, pero sin limitación, endonucleasas de restricción y endonucleasas de asentamiento. Véase, por ejemplo, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; y Belfort *et al.* (1997) *Nucleic Acid Res.* 25:3379-3388. Se conocen otras enzimas adicionales que escinden ADN (*por ejemplo*, S1 Nucleasa; nucleasa de porotos mung; ADNasa pancreática I; nucleasa de micrococos; endonucleasa HO de levaduras; véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Una o más de estas enzimas (o fragmentos funcionales de estos) se pueden usar como una fuente de los dominios de escisión y semidominios de escisión.

De forma similar, un semidominio de escisión se puede derivarse de cualquier nucleasa o porción de esta, como se expone anteriormente, que requiere dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Alternativamente, se puede usar una proteína única que comprende dos semidominios de escisión. Los dos semidominios de escisión se pueden derivar de la misma endonucleasa (o fragmentos funcionales de la misma), o

5 cada semidominio de escisión se puede derivar de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios blanco para las dos proteínas de fusión están dispuestos preferentemente uno con respecto al otro, de modo que la unión de las dos proteínas de fusión a sus respectivos sitios blanco sitúa a los semidominios de escisión en una orientación espacial entre sí que permite a los semidominios de escisión para formar un dominio de escisión funcional, *por ejemplo*, por dimerización. Por lo tanto, en ciertos aspectos, los bordes cercanos de los sitios blanco están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, cualquier número entero de nucleótidos o pares de nucleótidos puede intervenir entre dos sitios blanco (por ejemplo, de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios blanco.

10 Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y tienen capacidad para unirse al ADN en forma específica de secuencia (en un sitio de reconocimiento), y escindir ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (por ejemplo, tipo IIS) escinden el ADN en sitios retirados del sitio de reconocimiento y tienen dominios de unión y escisión separables. Por ejemplo, la enzima de tipo Type IIS *Fok I* cataliza la escisión de ADN de cadena doble, a 9 nucleótidos de su sitio de reconocimiento en una cadena y 13 nucleótidos de su sitio de reconocimiento en la otra. Véase, por ejemplo, patentes US 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. En consecuencia, en un aspecto, las proteínas de fusión comprenden el dominio de escisión (o semidominio de escisión) de al menos una enzima de restricción Tipo IIS y uno o más dominios de unión de dedos de cinc, que pueden estar o no manipulados genéticamente.

20 Un ejemplo de enzima de restricción de tipo IIS, cuyo dominio de escisión es separable del dominio de unión, es *Fok I*. Esta enzima particular es activa como dímero. Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA 95:10,570-10,575. En consecuencia, a los fines de la presente descripción, la porción de la enzima *Fok I* usada en las proteínas de fusión descritas se considera un semidominio de escisión. En consecuencia, para la escisión de cadena doble dirigida y/o el reemplazo dirigido de secuencias celulares usando fusiones de dedos de cinc-*Fok I*, se pueden usar dos proteínas de fusión, cada una que comprende un semidominio de escisión de *FokI*, para reconstituir un dominio de escisión catalíticamente activo. Alternativamente, también se puede usar una molécula de polipéptido única que contiene un dominio de unión de dedos de cinc y dos semidominios de escisión *Fok I*. Los parámetros para la escisión dirigida y la alteración de secuencia dirigida mediante las fusiones dedo de cinc-*Fok I* se proveen en otra parte de esta descripción.

30 Un dominio de escisión o semidominio de escisión puede ser cualquier porción de una proteína que retiene la actividad de escisión, o que retiene la capacidad de multimerizarse (*por ejemplo*, dimerizar) para formar un dominio de escisión funcional.

35 Los ejemplos de sistemas de restricción tipo III se describen en la publicación Internacional WO 07/014275. Las enzimas de restricción adicionales también contienen dominios de unión y escisión separables, y éstas se contemplan en la presente descripción. Véase, por ejemplo, Roberts et al. (2003) Nucleic Acids Res. 31:418-420.

40 En ciertos aspectos, el dominio de escisión comprende uno o más semidominios manipulados genéticamente (también denominados como mutantes del dominio de dimerización) que minimizan o evitan la homodimerización, como se describe en la Publicación de patente U.S. Nros 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 blancos para influir sobre la dimerización de los semidominios de escisión de *FokI*.

Los ejemplos de semidominios de escisión manipulados genéticamente de *Fok I* que forman los heterodímeros obligados incluyen un par en que un primer semidominio de escisión incluye las mutaciones en los residuos de aminoácidos en las posiciones 490 y 538 del *Fok I* y un segundo semidominio de escisión incluye las mutaciones en los residuos de aminoácidos 486 y 499.

45 En consecuencia, en ciertos aspectos, la mutación en 490 reemplaza Glu (E) con Lys (K); la mutación en 538 reemplaza Iso (I) con Lys (K); la mutación en 486 reemplaza Gln (Q) con Glu (E); y la mutación en la posición 499 reemplaza Iso (I) con Lys (K). En forma específica, los semidominios de escisión manipulados genéticamente descritos en la presente memoria se prepararon mediante la mutación de las posiciones 490 (E→K) y 538 (I→K) en un semidominio de escisión para producir un semidominio de escisión manipulado genéticamente denominado "E490K:I538K" y mediante la mutación de posiciones 486 (Q→E) y 499 (I→L) en otro semidominio de escisión para producir un semidominio de escisión manipulado genéticamente denominado "Q486E:I499L". Los semidominios de escisión manipulados genéticamente descritos en la presente son mutantes del heterodímero obligados en que se minimiza o anula la escisión aberrante. Véase, *por ejemplo*, Ejemplo 1 de la Publicación de patente U.S. N° 2008/0131962. Véase, también Szczepk et al. (2007) Nat Biotechnol 25:786-793. En ciertas realizaciones, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 486, 499 y 496 (numerados con respecto a *FokI* tipo salvaje), por ejemplo mutaciones que reemplazan el residuo de Gln (Q) tipo salvaje en la posición 486 con un residuo de Glu (E), el residuo Iso (I) tipo salvaje en la posición 499 con un residuo Leu (L) y el residuo de Asn (N) tipo salvaje en la posición 496 con un residuo de Asp (D) o Glu (E) (también denominado como dominios "ELD" y "ELE", respectivamente). En otros aspectos, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 490, 538 y 537 (numerados en relación con

FokI tipo salvaje), por ejemplo, mutaciones que reemplazan el residuo de Glu (E) tipo salvaje en la posición 490 con un residuo de Lys (K), el residuo de Iso (I) tipo salvaje en la posición 538 con un residuo de Lys (K), y el residuo de His (H) tipo salvaje en la posición 537 con un residuo de Lys (K) o un residuo de Arg (R) (también denominado como dominios "KKK" y "KKR", respectivamente). En otros aspectos, el semidominio de escisión manipulado genéticamente comprende mutaciones en las posiciones 490 y 537 (numerado en relación con FokI tipo salvaje), por ejemplo, las mutaciones que reemplazan el residuo de Glu (E) tipo salvaje en la posición 490 con un residuo de Lys (K) y el residuo de His (H) tipo salvaje en la posición 537 con un residuo de Lys (K) o un residuo de Arg (R) (también denominado como dominios "KIK" y "KIR", respectivamente). (Véase solicitud provisional internacional 61/337.769 presentada el 8 de febrero de 2010). En otros aspectos, el semidominio de escisión manipulado genéticamente comprende las mutaciones "Sharkey" y/o "Sharkey" (véase Guo et al, (2010) J. Mol. Biol. doi:10.1016/j.jmb.2010.04.060).

Los semidominios de escisión manipulados genéticamente descritos en la presente memoria se pueden preparar usando cualquier procedimiento adecuado, por ejemplo, por mutagénesis dirigida al sitio de semidominios de escisión tipo salvaje (Fok I) como se describe en las publicaciones de patente U.S. Nros. 20050064474 y 20080131962.

Otra enzima de restricción tipo IIS preferida es BfiI (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 de unión al ADN y unirse operativamente a un dominio de unión al ADN de dedos de cinc para crear una ZFN.

Proteínas de fusión

Los procedimientos para el diseño y la construcción de las proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos por los expertos en la técnica. Por ejemplo, los procedimientos para el diseño y la construcción de las proteínas de fusión que comprenden los dominios de unión al ADN (*por ejemplo*, dominios del dedo de cinc) y dominios reguladores o de escisión (o semidominios de escisión), y polinucleótidos que codifican tales proteínas de fusión, se describe en las patentes U.S. en copropiedad 6.453.242 y 6.534.261 y publicaciones de la solicitud de patente U.S. 2007/0134796 y 2005/0064474.

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 (ver más adelante para la descripción adicional con respecto a vectores y procedimientos para introducir los polinucleótidos en las células).

Como se indicó anteriormente, en ciertas realizaciones, la proteína de fusión comprende una proteína de dedos de cinc que se une a un sitio blanco en un gen involucrado en biosíntesis de ácidos grasos y al menos un dominio regulador transcripcional, por ejemplo un dominio de activación o represión.

En otros aspectos de los procedimientos descritos en la presente memoria, una nucleasa de dedos de cinc comprende una proteína de fusión que comprende un dominio de unión de dedos de cinc o un semidominio de escisión de la enzima de restricción *FokI*, y dos de estas proteínas de fusión se expresan en una célula. La expresión de dos proteínas de fusión en una célula pueden provenir de la administración de dos proteínas a la célula; administración de una proteína y un ácido nucleico que codifica una de las proteínas a la célula; administración de dos ácidos nucleicos, cada uno que codifica una de las proteínas, a la célula; o por la administración de un ácido nucleico único, que codifica ambas proteínas, a la célula. En aspectos adicionales, una proteína de fusión comprende una cadena polipeptídica simple que comprende dos semidominios de escisión y un dominio de unión a dedos de cinc. En este caso, una proteína de fusión única se expresa en la célula y, sin desear estar ligado por teoría alguna, se considera que escinde el ADN como resultado de la formación de un dímero intramolecular de los semidominios de escisión.

En ciertos aspectos, los componentes de las nucleasas de dedos de cinc (*por ejemplo*, fusiones ZFP-*FokI*) se disponen de manera tal que el dominio de dedos de cinc está más cerca del amino terminal de la proteína de fusión, y el semidominio de escisión está más cerca del carboxilo terminal. Esto refleja la orientación relativa del dominio de escisión en dominios de escisión dimerizante naturales tales como los derivados de la enzima *FokI*, en los que el dominio de unión de ADN está más cerca de estos aspectos, la dimerización de los semidominios de escisión para formar una nucleasa funcional se realiza mediante la unión de las proteínas de fusión a los sitios en cadenas de ADN opuestas, con los extremos 5' de los sitios de unión proximales entre sí.

En aspectos adicionales, los componentes de las proteínas de fusión (*por ejemplo*, las fusiones ZFP-*FokI*) se disponen de manera tal que el semidominio de escisión está más cerca del amino terminal de la proteína de fusión, y el dominio de dedos de cinc está más cerca del carboxilo terminal. En estos aspectos, la dimerización de los semidominios de escisión para formar una nucleasa funcional se realiza por la unión de las proteínas de fusión a sitios en cadenas de ADN opuestas, con los extremos 3' de los sitios de unión proximales entre sí.

En otros aspectos adicionales, una primera proteína de fusión contiene el semidominio de escisión más cerca del amino terminal de la proteína de fusión, y el dominio de dedo de cinc más cerca del carboxilo terminal, y una segunda proteína de fusión se disponen de manera tal que el dominio de dedos de cinc está más cerca del amino terminal de la proteína de fusión, y el semidominio de escisión está más cerca del carboxilo terminal. En estos

aspectos, ambas proteínas de fusión se unen a la misma cadena de ADN, con el sitio de unión de la primera proteína de fusión que contiene el dominio de dedo de cinc más cercano del carboxilo terminal ubicado en el lado 5' del sitio de unión de la segunda proteína de fusión que contiene el dominio de dedo de cinc más cerca del amino terminal.

5 En ciertos aspectos, proteínas de fusión descritas, la secuencia de aminoácidos entre el dominio de dedos de cinc y el dominio de escisión (o semidominio de escisión) se indica como "ligador ZC". El ligador ZC se debe distinguir de los ligadores entre dedos descritos anteriormente. Véase, *por ejemplo*, publicaciones de patente U.S. 20050064474A1 y 20030232410, y Publicación de patente internacional WO05/084190 para detalles sobre la obtención de ligadores ZC que optimizan la escisión.

10 En una realización, la descripción proporciona una ZFN que comprende una proteína de dedos de cinc que tiene una o más de secuencias de aminoácidos de la hélice de reconocimiento que se muestran en una fila única de la Tabla 1. En otro aspecto, se proporciona en la presente memoria un vector de expresión ZFP que comprende una secuencia de nucleótidos que codifica una ZFP que tiene las hélices de reconocimiento mostradas en la Tabla 1 o Tabla 10.

15 Regulación de la expresión génica

Una variedad de ensayos se puede utilizar 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*, mediante la medición, por ejemplo, de los niveles de proteína o ARNm, niveles de producto, actividad enzimática, activación o represión transcripcional de un gen indicador, usando, 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 matriz de oligonucleótidos), ensayos colorimétricos, ensayos de amplificación, ensayos de actividad enzimática, ensayos fenotípicos, y similares.

Las ZFP se analizan típicamente primero para determinar la actividad *in vitro* utilizando ensayos ELISA y a continuación, usando células renales. La ZFP a menudo se analiza primero utilizando un sistema de expresión transitorio con un gen indicador y luego se analiza la regulación del gen endógeno blanco en las células y en plantas enteras, tanto *in vivo* como *ex vivo*. La ZFP se puede expresar de modo recombinante en una célula, expresar de forma recombinante en las células trasplantadas en una planta, expresar de forma recombinante en una planta transgénica, así como administrarse como una proteína a la planta o célula utilizando vehículos de administración como se describe a continuación. Las células pueden estar inmovilizadas, estar en solución, inyectarse en una planta, o ser de origen natural en una planta transgénica o no transgénica.

La modulación de la expresión génica se analiza usando uno de los ensayos *in vitro* o *in vivo* descritos en la presente memoria. Las muestras o ensayos se tratan con una ZFP y se comparan con muestras de control sin el compuesto de prueba, para examinar el grado de modulación. Para la regulación de la expresión génica endógena, la ZFP típicamente tiene una K_d de 200 nM o menos, más preferentemente 100 nM o menos, más preferentemente 50 nM, con máxima preferencia 25 nM o menos.

Los efectos de las ZFP se pueden medir mediante el examen de cualquiera de los parámetros descritos anteriormente. Cualquier expresión génica, fenotípica o cambio 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 crecimiento de plantas, cambios transcripcionales en 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 los segundos mensajeros intracelulares, como cGMP.

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

En otro aspecto, la regulación de ZFP de la expresión génica endógena se determina *in vitro* mediante la medición del nivel de expresión del ARNm del gen blanco. El nivel de expresión génica se mide usando amplificación, por ejemplo, usando PCR, LCR o ensayos de hibridación, por ejemplo, hibridación Northern, protección con RNasa, transferencia de puntos. El nivel de proteína o ARNm se detecta usando agentes de detección marcados directa o indirectamente, por ejemplo, ácidos nucleicos marcados fluorescentemente o radiactivamente, anticuerpos marcados radiactivamente o enzimáticamente y similares, como se describe en la presente memoria.

Alternativamente, se puede idear un sistema de gen indicador usando el promotor del gen blanco unido operativamente a un gen indicador tal como luciferasa, proteína fluorescente verde, CAT o β -gal. El constructo indicador típicamente se cotransfecta en una célula cultivada. Después del tratamiento con la ZFP de elección, la cantidad de transcripción, traducción o actividad del gen indicador se mide de acuerdo con técnicas estándar conocidas por los expertos en la materia.

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, se pueden utilizar plantas que expresan transitoriamente la ZFP de elección, o a la que se ha administrado la ZFP en un vehículo de administración. La regulación de la expresión de genes endógenos se analiza usando uno de los ensayos descritos en la presente memoria.

Procedimientos para la escisión dirigida

Los procedimientos y las composiciones descritas se pueden usar para escindir ADN en una región de interés en la cromatina celular (*por ejemplo*, en un sitio deseado o predeterminado en un genoma, *por ejemplo*, dentro de o adyacente a un gen involucrado en biosíntesis de ácidos grasos). Para tal escisión de ADN dirigida, un dominio de unión de dedos de cinc se manipula genéticamente para unirse a un sitio blanco en o cerca del sitio de escisión predeterminado, y una proteína de fusión que comprende el dominio de unión de dedos de cinc manipulado genéticamente y un dominio de escisión se expresa en una célula. Después de la unión de la porción de de dedos de cinc de la proteína de fusión al sitio blanco, el ADN se escinde cerca del sitio blanco en el dominio de escisión. El sitio exacto de escisión puede depender de la longitud del ligador de ZC.

De forma alternativa, dos proteínas de fusión, cada una que comprende un dominio de unión de dedos de cinc y un semidominio de escisión, se expresan en una célula y se unen a los sitios blanco que se yuxtaponen de manera tal que se reconstituye un dominio de escisión funcional y se escinde el ADN en la proximidad de los sitios blanco. En un aspecto, la escisión se produce entre los sitios blanco de los dos dominios de unión de dedos de cinc. Uno o ambos dominios de unión de dedos de cinc se pueden manipular genéticamente.

Para la escisión dirigida usando un polipéptido de fusión del dominio de unión de dedos de cinc-dominio de escisión, el sitio de unión puede abarcar el sitio de escisión, o el borde cercano del sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25, 50 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) del sitio de escisión. La ubicación exacta del sitio de unión, con respecto al sitio de escisión, dependerá del dominio de escisión particular, y la longitud del ligador de ZC. Para los procedimientos en que se usan dos polipéptidos de fusión, cada uno que comprende un dominio de unión de dedos de cinc y un semidominio de escisión, los sitios de unión generalmente abarcan el sitio de escisión. El consecuencia el borde cercano del primer sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en un lado del sitio de escisión, y el borde cercano del segundo sitio de unión puede tener 1, 2, 3, 4, 5, 6, 10, 25 o más nucleótidos (o cualquier valor entero entre 1 y 50 nucleótidos) en el otro lado del sitio de escisión. Los procedimientos para mapear los sitios de escisión *in vitro* e *in vivo* son conocidos por los expertos en la técnica.

En consecuencia, los procedimientos descritos en la presente memoria pueden emplear un dominio de unión de dedos de cinc manipulado genéticamente fusionado a un dominio de escisión. En estos casos, el dominio de unión se manipula genéticamente para unirse a una secuencia blanco, en o cerca de donde se desea la escisión. La proteína de fusión, o un polinucleótido que codifica la misma, se introduce en una célula de planta. Una vez que se introduce en, o expresa en la célula, la proteína de fusión se une a la secuencia blanco y escinde en o cerca de la secuencia blanco. El sitio exacto de la escisión depende de la naturaleza del dominio de escisión y/o la presencia y/o naturaleza de las secuencias ligadoras entre los dominios de unión y escisión. En los casos en que se usan dos proteínas de fusión, cada una que comprende un semidominio 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 de la distancia entre los sitios de unión de las dos proteínas de fusión (*ver*, *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 ligador ZC en cada proteína 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 al longitud del ligador ZC en cada proteína de fusión. Véase, también, la Publicación de patente U.S. 20050064474A1 y las Publicaciones de patente internacional WO05/084190WO05/014791 y WO03/080809.

En ciertos aspectos, el dominio de escisión comprende dos semidominios de escisión, lo que son parte de un polipéptido único que comprende un dominio de unión, un primer semidominio de escisión y un segundo semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o diferentes secuencias de aminoácidos, siempre que actúen para escindir el ADN.

Los semidominios de escisión también se pueden proporcionar en moléculas separadas. *Por ejemplo*, dos polipéptidos de fusión se pueden introducir en una célula, en que cada polipéptido comprende un dominio de unión y un semidominio de escisión. Los semidominios de escisión pueden tener la misma secuencia de aminoácidos o diferentes secuencias de aminoácidos, siempre que actúen para escindir el ADN. Además, los dominios de unión se unen a las secuencias blanco que están normalmente dispuestas de manera tal, que después de la unión de los polipéptidos de fusión, los dos semidominios de escisión se presentan en una orientación espacial entre sí que permite la reconstitución de un dominio de escisión (*por ejemplo*, *por la* dimerización de los semidominios), de este modo se ubican los semidominios entre sí formar un dominio de escisión funcional, lo que produce la escisión de la cromatina celular en una región de interés. En general, la escisión mediante los dominios de escisión reconstituidos se produce en un sitio ubicado entre las dos secuencias blanco. Una o ambas proteínas se pueden manipular genéticamente para unirse a su sitio blanco.

Las dos proteínas de fusión se pueden unir en la región de interés en la misma u opuesta polaridad, y sus sitios de unión (*es decir*, sitios blanco) se puede separar por cualquier número de nucleótidos, *por ejemplo*, de 0 a 200 nucleótidos o cualquier valor entero entre ellos. En ciertos aspectos, los sitios de unión para dos proteínas de fusión, cada uno que comprende un dominio de unión de dedos de cinc y un semidominio de escisión, se pueden ubicar separados entre 5 y 18 nucleótidos, *por ejemplo*, separados por 5-8 nucleótidos, o separados por 15-18 nucleótidos, o separados por 6 nucleótidos, o separados por 16 nucleótidos, medidos desde el borde de cada sitio de unión más cercano al otro sitio de unión, y la escisión se produce entre los sitios de unión.

El sitio en el cual se escinde el ADN generalmente se encuentra entre los sitios de unión de las dos proteínas de fusión. La ruptura de la doble cadena de ADN a menudo es el resultado de dos rupturas de cadenas simples, o "mellas" compensadas por 1, 2, 3, 4, 5, 6 o más nucleótidos, (*por ejemplo*, la escisión de ADN de cadena doble por *Fok I* nativa es el resultado de rupturas de cadenas simples separadas por 4 nucleótidos). En consecuencia, la escisión no necesariamente ocurre en sitios exactamente opuestos de cada cadena de ADN. Además, la estructura de las proteínas de fusión y la distancia entre los sitios blanco puede influir sobre si la escisión se produce junto a un par de nucleótidos único, o si la escisión se produce en varios sitios. Sin embargo, para muchas aplicaciones, que incluyen la recombinación dirigida y la mutagénesis dirigida, (véase *infra*) la escisión dentro de un rango de nucleótidos generalmente es suficiente, y no se requiere la escisión entre pares de bases particulares.

Como se indicó con anterioridad, la proteína de fusión se puede introducir como polipéptidos y/o polinucleótidos. *Por ejemplo*, dos polinucleótidos, cada uno que comprende secuencias que codifican uno de los polipéptidos mencionados anteriormente, se pueden introducir en la célula, y cuando los polipéptidos se expresan y cada uno se une a su secuencia blanco, se produce la escisión en la secuencia blanco o cerca de ella. Alternativamente, se introduce un polinucleótido único que comprende secuencias que codifican ambos polipéptidos de fusión en una célula. Los polinucleótidos pueden ser ADN, ARN o cualquier forma modificada o análogos o ADN y/o ARN.

Para mejorar la especificidad de escisión, las composiciones adicionales también se pueden emplear en los procedimientos descritos en la presente memoria. *Por ejemplo*, los semidominios de escisión únicos pueden exhibir actividad de escisión de cadena doble limitada. En procedimientos en los que dos proteínas de fusión, cada una que contiene un dominio de dedos de cinc de tres dedos y un semidominio de escisión, se introducen en la célula, cada proteína específica un sitio blanco de aproximadamente 9 nucleótidos. Aunque la secuencia blanco agregada de 18 nucleótidos es probable que sea única en un genoma de mamífero y plantas, cualquier sitio blanco de 9 nucleótidos dado aparece, en promedio, aproximadamente 23.000 veces en el genoma humano. *Por lo tanto*, se puede producir la escisión no específica, debido a la unión específica de sitio de un semidominio único. En consecuencia, los procedimientos descritos en la presente memoria contemplan el uso de una mutante dominante negativa de un semidominio tal como *Fok I* (o un ácido nucleico que codifica el mismo) que se expresa en una célula junto con las dos proteínas de fusión. La mutante dominante negativa es capaz de dimerizarse pero es incapaz de escindir, y también bloquea la actividad de escisión de un semidominio al cual se dimeriza. Al proporcionar la mutante dominante-negativa en exceso molar para las proteínas de fusión, solo las regiones en las que se unen ambas proteínas de fusión tendrán una concentración local suficientemente alta de semidominios de escisión funcionales para que se produzca la dimerización. En sitios en los que solo se une una de las dos proteínas de fusión, su semidominio de escisión forma un dímero con el semidominio mutante negativo dominante, y no se produce escisión indeseable ni específica.

40 Vector de expresión

Un ácido nucleico que codifica una o más proteínas de fusión (*por ejemplo*, ZFP) como se describe en la presente memoria se puede clonar en un vector para la transformación en células procariotas o eucariotas para la replicación y/o expresión. Los vectores pueden ser vectores procarióticos (*por ejemplo*, plásmidos, o vectores de lanzadera, vectores de insectos o vectores eucarióticos). También se puede clonar un ácido nucleico que codifica una proteína de fusión en un vector de expresión para la administración a una célula de planta.

A fin de expresar las proteínas de fusión, las secuencias que codifican las proteínas de fusión generalmente se subclonan en un vector de expresión que contiene un promotor para dirigir la transcripción. Los promotores bacterianos o eucarióticos adecuados son bien conocidos en la técnica y han sido descritos, *por ejemplo*, en Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2nd ed. 1989; 3rd ed., 2001); Kriegler, Gene Transfer and Expression: A Laboratory Manual Patrón Press, New York (1990); and Current Protocols in Molecular Biology (Ausubel *et al.*, *supra*). Los sistemas de expresión bacteriana para expresar ZFP están disponibles, *por ejemplo*, en *E. coli*, *Bacillus sp.*, y *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983)). Los kits para dichos sistemas de expresión están disponibles en el comercio. Los sistemas de expresión eucarióticos para células de mamíferos, levaduras e insectos son bien conocidos por los expertos en la técnica y también están disponibles en el comercio.

El promotor usado para dirigir la expresión de un ácido nucleico que codifica la proteína de fusión depende de la aplicación particular. *Por ejemplo*, generalmente se usa un promotor constitutivo fuerte adecuado para la célula huésped para la expresión y purificación de las ZFPs.

En contraste, cuando se administra una ZFP in vivo para la regulación de un gen de planta (ver, la sección siguiente de "Administración de ácidos nucleicos a las células de planta"), se usa un promotor constitutivo o inducible, según

el uso particular de la ZFP. Los ejemplos no limitantes de promotores vegetales incluyen secuencias promotoras derivadas de ubiquitina-3 de *A. thaliana* (ubi-3) (Callis, et al., 1990, J. Biol. Chem. 265-12486-12493); manopinasintetasa de *A. tumefaciens* (Δ mas) (Petolino et al., patente U.S. N° 6.730.824); y/o virus del mosaico de las nervaduras de yuca (CsVMV) (Verdaguer et al., 1996, Plant Molecular Biology 31:1129-1139). Véase, también, Ejemplos.

Además del promotor, el vector de expresión generalmente contiene una unidad de transcripción o casete de expresión que contiene todos los elementos adicionales requeridos para la expresión del ácido nucleico en las células huésped, sean eucarióticas o procarióticas. Un típico casete de expresión en consecuencia contiene un promotor unido operativamente, por ejemplo, a una secuencia de ácido nucleico que codifica la ZFP y las señales requeridas, por ejemplo, para la eficiente poliadenilación del transcripto, terminación de la transcripción, sitios de unión del ribosoma o terminación de la traducción. Los elementos adicionales del casete pueden incluir, por ejemplo, potenciadores, señales de empalme heterólogas, y/o señal de localización nuclear (NLS).

El vector de expresión particular usado para transportar la información genética en una célula se selecciona con respecto al uso pretendido de las ZFP por ejemplo, la expresión en plantas, animales, bacterias, hongos, protozoos, etc. (véase vectores de expresión descritos más adelante). Los vectores de expresión bacterianos y animales estándares son conocidos en la técnica y se describen en detalle, por ejemplo, en la Publicación de la patente U.S. 20050064474A1 y las publicaciones de patente internacionales WO05/084190WO05/014791 y WO03/080809.

Los procedimientos de transfección estándares se pueden usar para producir líneas celulares bacterianas, de mamíferos, levaduras o insectos que expresan grandes cantidades de proteína, que luego se pueden purificar mediante técnicas estándares (véase, por ejemplo, Colley et al., J. Biol. Chem. 264:17619-17622 (1989); *Guide to Protein Purification, in Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). La transformación de células eucarióticas y procarióticas se realiza de acuerdo con técnicas estándares (véase, por ejemplo, Morrison, J. Bact. 132:349-351 (1977); Clark-Curtiss and Curtiss, *Methods in Enzymology* 101:347-362 (Wu et al., eds., 1983).

Se puede usar cualquiera de los procedimientos conocidos para introducir secuencias de nucleótidos extrañas en dichas células huésped. Estos incluyen el uso de transfección de fosfato de calcio, polibreno, fusión de protoplastos, electroporación, procedimientos ultrasónicos (por ejemplo, sonoporación), liposomas, microinyección, ADN desnudo, vectores plasmídicos, vectores virales, tanto episódicos como integradores y cualquier otro procedimiento bien conocido para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético extraño en una célula huésped (véase, por ejemplo, Sambrook et al., *supra*). Solo es necesario que el procedimiento de ingeniería genética particular usado sea capaz de introducir con éxito al menos un gen en la célula huésped capaz de expresar la proteína elegida.

Administración de ácido nucleico a las células de plantas

Como se indicó anteriormente, se pueden introducir constructos de ADN en (por ejemplo, en el genoma de) una planta huésped deseada mediante una variedad de técnicas convencionales. Para las revisiones de dichas técnicas véase, por ejemplo, Weissbach and Weissbach *Methods for Plant Molecular Biology* (1988, Academic Press, N.Y.) Section VIII, pp. 421-463; and Grierson and Corey, *Plant Molecular Biology* (1988, 2d Ed.), Blackie, London, Ch. 7-9.

Por ejemplo, el constructo de ADN genómico se puede introducir directamente en el ADN genómico de la célula de planta mediante técnicas tales como electroporación y microinyección de protoplastos de células de planta, o los constructos de ADN se pueden introducir directamente al tejido vegetal mediante procedimientos biolísticos, tales como bombardeo de partículas de ADN (véase, por ejemplo, Klein et al (1987) *Nature* 327:70-73). Alternativamente, los constructos de ADN se pueden combinar con adecuadas regiones flanqueantes de T-ADN e introducir en un vector huésped de *Agrobacterium tumefaciens* convencional. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, que incluye el desarmado y uso de vectores binarios, están bien descritas en la bibliografía científica. Véase, por ejemplo Horsch et al., (1984) *Science* 233:496-498, and Fraley et al., (1983) *Proc. Nat'l. Acad. Sci. USA* 80:4803.

Además, se puede lograr la transferencia génica mediante el uso de bacterias distintas de *Agrobacterium* o virus tales como *Rhizobium* sp. NGR234, *Sinorhizobium meliloti*, *Mesorhizobium loti*, virus de papa X, virus mosaico de coliflor y virus mosaico de nervaduras de yuca y/o virus de mosaico de tabaco, Véase, por ejemplo, Chung et al. (2006) *Trends Plant Sci.* 11(1):1-4.

Las funciones de virulencia del huésped *Agrobacterium tumefaciens* dirigirán la inserción del constructo y el marcador adyacente hacia el interior del ADN de la célula de planta cuando la célula es infectada por la bacteria mediante un vector de ADN-T binario (Bevan (1984) *Nuc. Acid Res.* 12:8711-8721) o el procedimiento de cocultivo (Horsch et al (1985) *Science* 227:1229-1231). Generalmente, el sistema de transformación de *Agrobacterium* se usa para manipular genéticamente plantas dicotiledóneas (Bevan et al (1982) *Ann. Rev. Genet* 16:357-384; Rogers et al (1986) *Methods Enzymol.* 118:627-641). El sistema de transformación de *Agrobacterium* también se puede usar para transformar, así como transferir, ADN a plantas monocotiledóneas y células de plantas. Véase, Patente U.S. N° 5.591.616; Hernalsteen et al., (1984) *EMBO J* 3:3039-3041; Hooykass-Van Slogteren et al., (1984) *Nature* 311:763-

764; Grimsley et al. (1987) *Nature* 325:1677-179; Boulton et al., (1989) *Plant Mol. Biol.* 12:31-40.; and Gould et al., (1991) *Plant Physiol.* 95:426-434.

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

Los procedimientos y composiciones descritos se pueden usar para insertar secuencias exógenas en una ubicación predeterminada en un genoma de célula de planta. Esto es útil en la medida en que la expresión de un transgén introducido en un genoma de planta depende esencialmente de su sitio de integración. En consecuencia, se pueden insertar genes que codifican, por ejemplo, nutrientes, antibióticos o moléculas terapéuticas, mediante recombinación dirigida, en regiones de un genoma vegetal favorables para su expresión.

Las células de planta transformadas que se producen por cualquiera de las técnicas de transformación anteriores se pueden cultivar para regenerar una planta entera, que posea el genotipo transformado y en consecuencia el fenotipo deseado. Dichas técnicas de regeneración dependen de la manipulación de ciertas fitohormonas en un medio de crecimiento del cultivo tisular, generalmente que dependen de un marcador biocida y/o herbicida que se ha introducido junto con las secuencias de nucleótido deseadas. La regeneración de la planta a partir de protoplastos cultivados es describe en Evans, et al., "Protoplasts Isolation and Culture" in *Handbook of Plant Cell Culture*, pp. 124-176, Macmillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985 También se puede lograr la regeneración a partir de callos vegetales, explantes, órganos, polen, embriones o partes de estos. Dichas técnicas de regeneración se describen en general en Klee et al. (1987) *Ann. Rev. of Plant Phys.* 38:467-486.

Los ácidos nucleicos introducidos en una célula de planta se pueden usar para conferir rasgos deseados en esencialmente cualquier planta. Una amplia variedad de plantas y sistemas de células de planta se pueden manipular genéticamente para obtener las características fisiológicas y agronómicas deseadas descritas en la presente memoria mediante el uso de los constructos de ácidos nucleicos de la presente descripción y los diversos procedimientos de transformación mencionados anteriormente. Las plantas y células de planta blanco para manipulación genética incluyen, sin limitaciones, las plantas monocotiledóneas y dicotiledóneas, tales como cultivos que incluyen cultivos de granos (por ejemplo, trigo, maíz, arroz, mijo, cebada), cultivos de frutas (por ejemplo, tomate, manzana, pera, frutilla, naranja), cultivos de forraje (por ejemplo, alfalfa), cultivos de hortalizas de raíz (por ejemplo, zanahoria, patata, remolacha azucarera, batata), cultivos de hortalizas de hoja (por ejemplo, lechuga, espinaca); plantas de flor (por ejemplo, petunia, rosa, crisantemo), coníferas y pinos (por ejemplo, pino de mar, abeto); plantas usadas en fitomedicamentos (por ejemplo, plantas que acumulan metales pesados); cultivos oleaginosos (por ejemplo, girasol, semilla de colza, soja) y plantas usadas con fines experimentales (por ejemplo, *Arabidopsis*). En consecuencia, los procedimientos y las composiciones descritas se utilizan en una amplia variedad de plantas, que incluyen pero sin limitación, 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 que el casete de expresión se incorpora de forma estable en plantas transgénicas y se confirma que es operable, se puede introducir en otras plantas mediante cruzamiento sexual. Se puede usar cualquiera de numerosas técnicas de cultivo estándar, de acuerdo con la especie para cruzar.

Una célula de planta, callo, tejido o planta transformada se puede identificar y aislar mediante la selección o análisis del material de la planta manipulado genéticamente por rasgos codificados por los genes marcadores presentes en ADN de transformación. Por ejemplo, la selección se puede realizar por cultivo del material vegetal manipulado genéticamente en medios que contienen una cantidad inhibitoria de antibióticos o herbicidas a los cuales el constructo de gen transformante confiere resistencia. Además, las plantas y células de plantas transformadas también se pueden genes de la β -glucuronidasa), luciferasa o C1), que pueden estar presentes en los constructos de ácido nucleico recombinante. Estas metodologías de selección y detección son bien conocidas por los expertos en la materia.

También se pueden usar procedimientos físicos y bioquímicos para identificar transformantes de plantas o células de planta que contienen constructos génicos insertados. Estos procedimientos incluyen pero sin limitaciones: 1) análisis Southern o amplificación por PCR para detectar y determinar la estructura del inserto de ADN recombinante; 2) transferencia Northern, protección de S1 ARNs, extensión de cebador o amplificación por PCR de transcriptasa inversa para detectar y examinar transcritos de ARN de los constructos génicos; 3) ensayos enzimáticos para detectar actividad de enzimas o ribozima, cuando dichos productos génicos son codificados por el constructo génico;

4) electroforesis en gel de proteínas, técnicas de transferencia Western, inmunoprecipitación, o inmunoensayos ligados a enzima, cuando los productos de la construcción génica son proteínas. Otras técnicas, tales como hibridación *in situ*, tinción de enzimas e inmunotinción, también se pueden usar para detectar la presencia o expresión del constructo recombinante en órganos y tejidos vegetales específicos. Los procedimientos para llevar a cabo todos estos ensayos son bien conocidos en la técnica.

Los efectos de la manipulación génica mediante los procedimientos descritos en la presente memoria se pueden observar, por ejemplo, mediante transferencias Northern del ARN (por ejemplo, ARNm) aislado de los tejidos de interés. Generalmente, si el ARNm está presente o la cantidad de ARNm ha aumentado se puede suponer que el correspondiente gen endógeno se está expresando en una tasa mayor que antes. Se pueden usar otros procedimientos para medir la actividad génica y/o CYP74B. Se pueden usar diferentes tipos de ensayos enzimáticos, según el sustrato usado y el procedimiento para detectar el incremento o la disminución de un producto de reacción o subproducto. Además, los niveles de proteína y/o CYP74B expresados se pueden medir por inmunoquímica, es decir, ELISA, RIA, EIA y otros ensayos basados en anticuerpos bien conocidos por los expertos en la técnica, tales como ensayos de detección por electroforesis (con tinción o transferencia western). El transgén se puede expresar en 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 vegetales, sustancialmente a lo largo de todo el ciclo de vida. Sin embargo, también se aplica cualquier modo de expresión combinatoria.

La presente descripción también abarca semillas de las plantas transgénicas descritas anteriormente, en las que a semilla tiene el constructo del transgén o gen, y que incluye las semillas con los perfiles de aceite modificados. La presente descripción también abarca la progenie, clones, líneas celulares o células de las plantas transgénicas descritas anteriormente en las que dicha progenie, clon, línea celular o célula posee el constructo transgénico o génico.

La administración de cantidades efectivas se produce por cualquiera de las vías normalmente usadas para introducir ZFP en contacto último con la célula vegetal tratada. Las ZFP se administran de cualquier manera adecuada, preferentemente con portadores aceptables. Los procedimientos adecuados para administrar dichos moduladores están disponibles y son bien conocidos por los expertos en la técnica, y si bien se puede usar más de una vía para administrar una composición particular, a menudo una vía particular puede proporcionar una reacción más inmediata y más efectiva que otra vía.

También se pueden usar portadores y se determinan en parte por la composición particular que se administra, así como por el procedimiento particular usado para administrar la composición. En consecuencia, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas disponibles.

Aplicaciones

Como se indicó anteriormente, la modulación dirigida de los genes involucrados en la síntesis de ácidos grasos se puede usar para generar de manera rápida y eficiente aceites de planta del perfil de ácidos grasos deseado.

La alteración de las composiciones de ácidos grasos de los aceites de planta de las plantas productoras de aceite, tales como, por ejemplo, canola, tiene profundas implicaciones para la producción de alimentos y, por consiguiente, para la salud de la dieta. Por ejemplo, las variedades *Brassica* de semillas oleaginosas de calidad canola con niveles reducidos de ácidos grasos saturados en el aceite de semilla se pueden usar para producir productos alimenticios que promuevan la salud cardiovascular. Los procedimientos y composiciones descritos en la presente memoria se pueden usar para generar plantas productoras de aceite con bajo contenido de ácido palmítico y/o esteárico, que reducen los niveles de ácidos grasos saturados en el aceite de planta. Del mismo modo, la generación de plantas productoras de petróleo con niveles aumentados de contenido de ácido oleico también reducirá la cantidad de ácidos grasos saturados en el aceite de planta.

Además, los procedimientos y las composiciones que se describen en la presente memoria se pueden usar para aumentar el contenido de ácido palmítico, por ejemplo en aceites producidos a partir de tales plantas. Los aceites ricos en contenido de ácido palmítico son particularmente útiles en la formulación de las margarinas.

Además, la alteración dirigida de los perfiles de ácidos grasos como se describe en la presente memoria se puede usar para generar plantas (y aceites de planta) que tienen bajo contenido de ácido linolénico. Los aceites con bajo contenido de ácido linolénico muestran aumento de estabilidad; los alimentos elaborados usando estos aceites no se estropean en términos de sabor u olor tan rápido como los alimentos elaborados con materiales vegetales con altas concentraciones de ácido linolénico.

Ejemplos

Los siguientes son ejemplos de aspectos específicos descritos en la presente memoria.

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

Ejemplo 1: Identificación de secuencia blanco en *Brassica napus* L

1.1 Identificación de secuencia blanco

En este ejemplo, se identificaron las secuencias de ADN endógeno que codifican la enzima β -cetoacil-ACP sintetasa II (KAS II) nativa de *Brassica napus* L (canola). Estos genes se seleccionaron como ejemplos de blancos para demostrar la regulación transcripcional a través de factores de transcripción de proteínas con dedos de cinc manipulados genéticamente (ZFP TF), lo que produce la modificación deseada de la biosíntesis de ácidos grasos y los perfiles de aceite de semilla alterados concomitantes. La enzima β -cetoacil-ACP sintetasa II cataliza la conversión de 16:0-ACP a 18:0-ACP y la posterior formación de ácido oleico (Ohlrogge and Browse, 1995, The Plant Cell, 7:957-970). Una mutante *fab1* de *Arabidopsis thaliana* del gen β -cetoacil-ACP sintetasa II produjo una reducción de 65% en la actividad enzimática y aumentos concomitantes del contenido de ácido esteárico en un 7% y un 3% en hojas y raíces, respectivamente (Wu et al., 1994, Plant Physiology, 106:143-150). Un transgén de β -cetoacil-ACP sintetasa II de soja transformado de expresión estable disminuyó el contenido de ácido palmítico de semilla en 0.8% y la introducción del mismo gen en tabaco disminuyó el contenido de ácido palmítico en 2% (Publicación de patente japonesa N° 501446/1995).

Se han informado secuencias de ADNc de β -cetoacil-ACP sintetasa II a partir de múltiples especies de planta que incluyen *Arabidopsis thaliana* (GenBank:AF318307) y *Brassica napus* (GenBank:AF244520). El alineamiento de las secuencias de ADNc de *A. thaliana* y *B. napus* (GenBank:AF318307 y AF244520, respectivamente) mostró que la secuencia AF244520 estuvo incompleta (Figura 7). A esta secuencia de ADN de *B. napus* truncada le faltaban varios cientos pares de bases en el extremo 5'. Además, debido a que *B. napus* es una especie anfiploide resultante de la combinación de los conjuntos cromosómicos 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 puede existir más de un gen de β -cetoacil-ACP sintetasa II en esta especie. Se identificaron y se obtuvieron las secuencias 5' UTR adicionales presentes en las secuencias de ADNc. Estas secuencias génicas de 5' β -cetoacil-ACP sintetasa II sirvieron como blancos para la regulación por aumento transcripcional a través de ZFP-TF en los ejemplos presentes.

1.2 Aislamiento de ARN total

El ARN total se aisló de semilla inmadura del genotipo de *Brassica napus* (canola) Nex710 (Certificado de cultivo 99-7049208-501) 15 días después de la floración (DAF) usando el RNEASY@PLANT MINI KIT de Qiagen (Qiagen, Valencia, CA). El ARNm total se trató con ADNasa libre de ARNasa según las recomendaciones del fabricante para eliminar cualquier ADN contaminante que podría amplificarse durante la RT-PCR cuantitativa.

1.3 5' RACE y análisis de secuencia

La amplificación rápida de extremos de ADNc (RACE) específicos para el extremo 5' del ADNc de β -cetoacil-ACP sintetasa II de *B. napus* (GenBank AF244520) se realizó usando el kit FIRSTCHOICE® RLM-RACE de Ambion (Ambion, Austin, TX) por recomendaciones del fabricante. Para obtener la secuencia de ADNc 5', un adaptador de ARN sintético se apareó a la región de ARNm sin terminación 5'. Un cebador suministrado con el kit (5'-CGCGATCCGAACACTGCGTTTGCTGGCTTTGATGAAA - 3') (SEQ ID NO:1) y un segundo cebador (5'-CTCGAGCTGCTACTGCTAGTTCGGTGGAGGAGCC - 3') (SEQ ID NO:2) que se diseñó para unirse dentro de la secuencia de ADNc de *B. napus* parcial (GenBank AF244520) se usaron para amplificar un fragmento para determinar la secuencia corriente arriba desconocida. La amplificación 5' RACE identificó aproximadamente de 500 pares de bases de nuevas secuencias de *B. napus*. Se identificaron cóntigos de cuatro secuencias de ADNc 5' claramente diferentes de los genes de β -cetoacil-ACP sintetasa II (SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:46 y SEQ ID NO:30) cuando se alinearon las secuencias amplificadas (Figura 8). Estas secuencias 5' mostraron altos niveles de homología con las secuencias de *B. rapa* (GenBank:AC189461) y *B. oleracea* (GenBank:BH723504) de la misma región.

Las secuencias de unión de ZFP TF se identificaron en las secuencias de 5' cDNA de genes de β -cetoacil-ACP sintetasa II. Las secuencias de la región corriente arriba de los genes de β -cetoacil-ACP sintetasa II (Tabla 2) sirvieron como blancos para la unión de ZFP TF (Tabla 1). Constructos de plásmido:pDAB4695, pDAB4696, pDAB4697, y pDAB4698 que contienen los diseños de ZFP TF manipulados genéticamente se usaron para transformar de forma estable *B. napus* como se describe en las siguientes secciones. Se planteó la hipótesis de que las ZFP TF manipuladas genéticamente después de la expresión en células de planta se pueden unir a sus blancos de β -cetoacil-ACP sintetasa II endógenos, lo que produce la expresión de ARNm modificada del blanco.

Ejemplo 2: Diseño de los dominios de ADN de ZFP específicos para el gen de β -cetoacil-ACP sintetasa II

Las proteínas de dedos de cinc se diseñaron contra varios sitios blanco en la región promotora del gen de β -cetoacil-ACP sintetasa II de *B. napus* y en la región 5' traducida y no traducida. Véase, Figura 2. Las hélices de reconocimiento para los diseños representativos de ZFP se muestran a continuación en la Tabla 1. Los sitios blanco de los diseños de dedos de cinc se muestran a continuación en la Tabla 2.

Tabla 1: Diseños de dedos de cinc de β -cetoacil-ACP sintetasa II

ZFP	F1	F2	F3	F4	F5	F6
14025	RSDNLSV (SEQ ID NO:5)	QKINLQV (SEQ ID NO:6)	RSDLTSE (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: Sitio blancos de dedos de cinc β -cetoacil-ACP sintetasa

ZFP	Sitio blanco (5' a 3')	Sitio blanco/unión de ZFP presente en las SEQ ID Nos.
14025	cGTGGAGACGtCAAAGa (SEQ ID NO:24)	3, 4, 46 y 30
14033	aAGGAAGGCGAGAAAAGGg (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 β -cetoacil-ACP sintetasa II se incorporaron en el vector de expresión de dedos de cinc que codifica una proteína que tiene al menos un dedo con una estructura CCHC. Véase, Publicación de patente U.S. N° 2008/0182332. En particular, el último dedos de cada proteína tenía una estructura CCHC (arquitectura).

Las secuencias que codifican los dedos de cinc se fusionaron luego a secuencias que codifican un dominio de activación de VP16 y una señal de localización nuclear *opaque-2*. La expresión de las proteínas de fusión fue impulsada por un promotor constitutivo relativamente fuerte tal como un promotor derivado del promotor Ubiquitina 10 de *Arabidopsis thaliana* (AtUbi10). Los ejemplos de vectores se muestran en la Tabla 3 a continuación.

Ejemplo 3: Regulación por aumento mediado por ZFP TF del gen β -cetoacil-ACP sintetasa II nativo en *B. napus*

Con el fin de evaluar la funcionalidad de las proteínas de dedos de cinc diseñadas en células de plantas, se utilizaron procedimientos para la expresión de tales proteínas en células de planta vivas. El ADN que codifica las proteínas del dedos de cinc se puede administrar a las células de planta en condiciones en las que el ADN no se incorpora al genoma de la célula vegetal. Por lo tanto, la molécula de ADN se mantiene transitoriamente en las células de planta y actúa como un molde para la expresión génica. Alternativamente, el ADN que codifica las proteínas del dedos de cinc se puede administrar a las células de planta en condiciones que permitan que el ADN se incorpore en el genoma de la célula vegetal, lo que da como resultado la transgénesis de los genes que codifican la proteína del dedos de cinc de modo que la molécula de ADN se mantiene de forma estable en las células de planta y actúa como molde para la expresión génica. Un experto en la técnica puede utilizar la expresión transitoria o transgénica de proteínas con dedos de cinc que codifican ADN con el fin de evaluar la funcionalidad de estas proteínas en células de plantas vivas.

25 3.1 Diseño del constructo

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

Tabla 3: Descripción del constructo para ZFP TF dirigidos a *B. napus* KAS II

S.N.	ZFP	Constructo No.	Casete génico
1	14025	pDAB4695	Atubi10/ZFP1-Vp16/AtuORF23/CsVMV/pat/AtuORF1
2	14033	pDAB4696	Atubi10/ZFP2-Vp16/AtuORF23/CsVMV/pat/AtuORF1
3	14035	pDAB4697	Atubi10/ZFP3-Vp16/AtuORF23/CsVMV/pat/AtuORF1
4	14047	pDAB4698	Atubi10/ZFP4-Vp16/AtuORF23/CsVMV/pat/AtuORF1

AtUbi10 = Promotor de Ubiquitina 10 *Arabidopsis thaliana*, CsVMV = promotor del virus de mosaico de las nervaduras de yuca, ZFP = gen de proteína de dedos de cinc, *pat* = gen de fosfinotricina acil transferasa ATuORF1 = 3' UTR 1 *Agrobacterium tumefaciens* y AtuORF23 = 3' UTR 23 *Agrobacterium tumefaciens*.

5 pDAB4695 es un plásmido binario que contiene los casetes de expresión génica de 14025 v3/VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 (Región de unión de la matriz (Thompson et al., 1997, WO9727207)) ::Promotor v2 AtUbi10 (Promotor de Ubiquitina-10 de *Arabidopsis thaliana* (Callis, et al., 1990, J. Biol. Chem. 265-12486-12493)) :: Fusión 14025 v3 dedos de cinc/VP16 ::Atu ORF23 3'UTR v1 (Marco de lectura abierto 23, región 3' no traducida de *Agrobacterium tumefaciens* (Gelvin et al., 1987, EP222493))::Promotor v2 CsVMV (Promotor del virus de mosaico de las nervaduras de yuca (Verdaguer et al., 1996, Plant Molecular Biology 31:1129-1139))::*pat* v5 (Fosfinotricina acetil transferasa (Wohlleben et al., 1988, Gene 70:25-37))::AtuORF 1 3'UTR v4 (Marco de lectura abierto 1, región 3' no traducida *Agrobacterium tumefaciens* (Huang et al., J. Bacteriol. 172:1814-1822)). El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión 14025 v3 Dedos de cinc/VP16 en pDAB3916 por medio de los sitios de restricción NcoI- SacI. El constructo resultante que se marcó como pDAB8221 contenía el casete de expresión del gen Promotor v2 AtUbi10 ::Fusión 14025 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1. pDAB8221 se clonó en el binario pDAB7309 por medio de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4695 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

20 pDAB4696 es un plásmido binario que contiene los casetes de expresión génica de 14033 v3/VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor v2 AtUbi10 ::Fusión 14033 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1 ::Promotor v2 CsVMV ::*pat* v5 ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la Fusión 14033 v3 Dedos de cinc/VP16 en pDAB3916 por medio de los sitios de restricción NcoI - SacI. El constructo resultante que se marcó como pDAB8222 contenía el casete de expresión del gen Promotor v2 AtUbi10 ::Fusión 14033 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1. pDAB8222 se clonó en el binario pDAB7309 por medio de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4696 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

30 pDAB4697 es un plásmido binario que contiene los casetes de expresión génica de 14035 v3/VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor v2 AtUbi10 ::Fusión 14035 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1 ::Promotor v2 CsVMV ::*pat* v5 ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la Fusión 14035 v3 Dedos de cinc/VP16 en pDAB3916 por medio de los sitios de restricción NcoI- SacI. El constructo resultante que se marcó como pDAB8223 contenía el casete de expresión del gen Promotor v2 AtUbi10 ::Fusión 14035 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1. pDAB8223 se clonó en el binario pDAB7309 por medio de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4697 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

35 pDAB4698 es un plásmido binario que contiene los casetes de expresión génica de 14047 v3/VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor v2 AtUbi10 ::Fusión 14037 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1 ::Promotor v2 CsVMV ::*pat* v5 ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la Fusión 14037 v3 Dedos de cinc/VP16 en pDAB3916 por medio de los sitios de restricción NcoI- SacI. El constructo resultante que se marcó como pDAB8224 contenía el casete de expresión del gen Promotor v2 AtUbi10 ::Fusión 14037 v3 Dedos de cinc/VP16 ::Atu ORF23 3'UTR v1. pDAB8224 se clonó en el binario pDAB7309 por medio de una reacción L-R Gateway® (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAB4698 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

3.2 Transformación de *Agrobacterium*

Se prepararon células de *Agrobacterium* para la electroporación usando el protocolo descrito en Weigel D., Glazebrook J. *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2002:pg123-124. Se realizaron modificaciones menores en el protocolo que permitieron el crecimiento óptimo de *Agrobacterium* (es decir, el medio LB se sustituyó con YEP). Independientemente, se añadieron 1,5-3 µg de ADN plasmídico para cada constructo a 50 µl de células de *Agrobacterium tumefaciens* C58 ::Z707 y se mezclaron suavemente. La mezcla se transfirió a cubetas GENE PULSER® de 0,2 cm frías (BioRad Hercules, CA) y se colocó sobre hielo. Las cubetas se colocaron luego en una gradilla GENE PULSER® fría (BioRad, Hercules, CA) y se sometieron a electroporación en las siguientes condiciones: salida de capacitancia de 25 µFarad, extensor de capacitancia 960 µFarad, resistencia de 200 ohmios y voltaje de 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 and Co., Franklin Lakes, NJ) o equivalente. Las células transformadas se incubaron luego a 28 °C durante 1 hora. Después de la incubación, 50 µl, 100 µl y 200 µl de células se sembraron en placas de medio YEP separadas (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 contienen antibióticos según corresponda. Las placas se cultivaron invertidas a 28 °C durante aproximadamente 36-48 horas. Se recogieron colonias individuales y se 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 contienen antibióticos según correspondía, a 28 °C durante aproximadamente 36 horas.

La cepa de *Agrobacterium tumefaciens* se confirmó mediante una prueba de cetolactosa. Las colonias putativamente transformadas se sembraron en estrías en agar con lactosa y se incubaron a 28 °C durante 48 horas. Las placas luego se inundaron con la solución de Benedict. Las placas se controlaron; las cepas sembradas en estrías que convirtieron 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*, Volume 44. Humana Press, 1995:9-13).

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

3.3 Transformación de *B. napus* con ZFP TF

Preparación de los segmentos de hipocótilo: las semillas del genotipo de *B. napus*, Nex 710, se esterilizaron en superficie con blanqueador comercial 10% durante 10 minutos y se enjuagaron 3 veces con agua destilada estéril. Las semillas se secaron con una toalla de papel estéril y luego se colocaron en una bandeja de Phyta que contenía "medio de germinación" que consiste en la mitad de la 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 en régimen de crecimiento ajustado a 23 °C, y un fotoperíodo de 16 h luz/8 h de oscuridad.

El día 5, se verificó la esterilidad de las plántulas y se colocó la bandeja Phyta dentro de una campana de flujo laminar EDGE GARD® (The Baker Company, Sanford, ME) para mantener la esterilidad. Usando pinzas estériles y tijeras de disección, las plantas se retiraron de la bandeja Phyta y las regiones aéreas (meristemas y cotiledones) y las raíces se desprendieron y descartaron. Los hipocótilos se colocaron en una placa Petri de 100 x 25 mm que contiene agua destilada estéril que se requiere 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 usando un bisturí #10. Los segmentos de hipocótilo se colocaron en un "medio de inducción de callo" que consiste en medio MS que contiene 30 g/l de sacarosa, 1 mg/l de quinina y 1 mg/l de 2,4-D solidificado con 7 g/l de agar TC que incluía un papel de filtro estéril. Las placas se colocaron en una cuba transparente STERILITE® y se mantuvieron bajo el mismo régimen de crecimiento durante 3 días, como un pretratamiento.

El día 8, los segmentos luego se transfirieron a una placa de Petri estéril de 100 x 25 que contiene 20 ml de "medio de cultivo líquido" que consiste en medio basal Linsmaier y Skoog (1965) que contiene 30 g/l de glucosa, vitaminas Gamborg de ½ concentración (1968), 215,2 mg/l de quinina y 221,04 mg/12,4-D para un pretratamiento de 1 hora. El "medio de cultivo líquido" se retiró de los segmentos de hipocótilo y 40 ml de suspensión de *Agrobacterium* (que contiene pDAB4695, pDAB4696, pDAB4697 o pDAB4698 en Z707) a 50 Klett se agitaron en vórtex brevemente y se vertieron en la placa de Petri de 100 x 25 mm que contiene 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 volvieron a colocar sobre el "medio de inducción de callo" más papeles de filtro, se devolvieron a la cuba STERILITE®, se cubrieron con una tapa oscura y se devolvieron a la sala de cultivo bajo el mismo régimen de crecimiento que antes, durante un período de cultivo de 3 días. Después de 3 días, los hipocótilos se colocaron directamente en un "medio de inducción de callos" que contiene 1 mg/l de HERBIACE®, se volvieron a colocar en la cuba con una tapa transparente y se devolvieron a la sala de cultivo, manteniendo el mismo régimen de crecimiento que antes. Después de 1 semana, los hipocótilos se transfirieron directamente al "medio de

inducción de callos" con selección a 3 mg/l de HERBIACE® durante 2 semanas para el desarrollo de callos adicionales y se transfirieron al "medio de inducción de callos" con selección a 5 mg/l de HERBIACE durante 2-8 semanas para el desarrollo de callos adicionales. Una vez que estuvo disponible una cantidad suficiente de callos, se envió para análisis moleculares.

- 5 Regeneración de plantas de callos de canola envejecida: se colocó tejido de callo de canola en 'medio de regeneración de brotes' que consiste en medio MS que contiene 30 g/l de sacarosa, 3 mg/l de benzaminopurina, 0,5 g/l de MES [ácido 2- (N-morfolino)etansulfónico], 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®, las placas se envolvieron con cinta de 3M y se colocaron en régimen de crecimiento ajustado a 23 °C, y un fotoperíodo de 16 horas luz/8 horas oscuridad. Los tejidos se movieron luego al 'medio de elongación de brotes' que consista en medio MS que contiene 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®. Las plántulas se transfirieron a un 'medio de enraizamiento' que consiste en medio ½ MS que contiene 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®.
- 10
- 15 Una vez que las raíces se establecieron en la plántula in vitro, se trasplantaron a macetas de 5 ¼" que contienen Metro Mix 360. Las plantas se cubrieron con una copa Solo transparente y se colocaron en un conviron para aclimatación. Después de 48-72 horas, se retiraron las copas para permitir la circulación de aire. Después de un total de siete días, las macetas se cambiaron del conviron a una bahía de invernadero para su posterior desarrollo. Las plantas se cultivaron en un fotoperíodo de 16:8 horas, con temperatura diurna y nocturna entre 22 y 24 °C. Cuando el tallo primario de floración comenzó a alargarse y formar capullos florales, la planta entera se cubrió con una bolsa de autofecundación para evitar el cruzamiento. Las semillas derivadas de las autopolinizaciones se cosecharon aproximadamente cuatro meses después del trasplante.
- 20

- Preparación de *Agrobacterium*: El *Agrobacterium* de un patrón de glicerol se sembró en estrías, cuatro días antes del tratamiento, sobre "medio de cultivo bacteriano semisólido" que consiste 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 espectinomicina y 250 mg/L de estreptomina y solidificada con 15 g/L de Bacto Agar y se cultivaron durante dos días en un incubadora (Incubadora Fisher Scientific Isotemp) a 28 °C. Después de 2 días, se colocó un pequeña asa de siembra de *Agrobacterium* en un matraz con deflector descartable estéril de 500 ml que contiene 150 ml de 'medio de cultivo bacteriano líquido' (el mismo que el anterior menos el agente solidificante), 250 mg/l de estreptomina y 100 mg/l de espectinomicina, cultivada durante 16 días. horas durante la noche a 28 °C en la oscuridad en un agitador cerrado (agitador de incubador refrigerado New Brunswick Scientific Innova 4330) a 200 rpm. Después de 16 horas, el cultivo de *Agrobacterium* se retiró del agitador y se alicuotó en tubos de centrifuga de 50 ml (uno que contiene 35 ml para la preparación y dos que contienen 50 ml para la revalidación). Los tubos de centrifuga se colocan en una centrifuga (centrifuga Beckman Modelo J2-21) y se centrifugan a 6.000 rpm durante 15 minutos y posteriormente se resuspender en el 'medio de cultivo líquido' hasta una densidad final de Klett 50 con un filtro rojo.
- 25
- 30
- 35

Ejemplo 4: Análisis de las muestras de callos transformados

- Las muestras de callo de *B. napus* transformadas putativamente se analizaron en busca de alteraciones en las expresiones de ARNm para el gen endógeno de β-cetoacil-ACP sintetasa II, el gen endógeno de tubulina y el transgén de ZFP TF. Los niveles de ARNm de tubulina se usaron como un control interno para normalizar la expresión de ARNm de ZFP TF y β-cetoacil-ACP sintetasa II. Los datos de expresión de ARNm de ZFP TF se usaron para confirmar la presencia de al menos un transgén de ZFP TF funcional en callos transformados.
- 40

4.1 Preparación de la muestra de callo

- Aproximadamente 40-50 muestras de callo de ocho semanas de *B. napus* se obtuvieron después de la transformación de tejido de hipocótilo Nex710. Se transformaron individualmente con los constructos de ZFP TF, pDAB4695, pDAB4696, pDAB4697 y pDAB4698 como se describió anteriormente en el Ejemplo 3.3. Las muestras de control se obtuvieron por transformación de una constructo binario de control que contiene un casete de expresión del gen *pat* (Promotor v2 AtUbi10 ::*pat* v3 ::Atu ORF1 3'UTR v3). Todas las muestras se cultivaron en medio de cultivo celular complementado con HERBIACE hasta su recolección.
- 45

- El ARN total se preparó a partir del tejido de callo fresco usando el kit QIAGEN RNEASY 96 (Qiagen, Valencia, CA). El ARN se trató con ADNasa libre de ARNasa de acuerdo con las instrucciones del kit para eliminar cualquier contaminante de ADN genómico. La síntesis de la primera cadena se estableció de acuerdo con las instrucciones del fabricante de la enzima de transcriptasa inversa Superscript® III (Invitrogen, Carlsbad, CA) y se activó usando hexámeros aleatorios. Las cadenas de ADNc sintetizadas se diluyeron en agua a relaciones de 1:10 y 1:50 (esto proporciona un molde suficiente para amplificar por PCR múltiples blancos). Cada alicuota se guardó a -20 °C indefinidamente.
- 50
- 55

4.2 Análisis de la expresión de ARNm de β-cetoacil-ACP sintetasa II

Se prepararon mezclas de reacción qRT-PCR para la amplificación del ADNc de la β-cetoacil-ACP sintetasa II de la siguiente manera: 7,5 µl de tampón Master 2X LC480 Probes (Roche Diagnostic, Indianapolis, IN), 0,3 µl de cebador

directo específico del gen (SEQ ID NO :28:5'-TTGACTCGAGCTGCTACTGC-3'; posiciones de nucleótidos 544 - 563 para la SEQ ID NO:3 en el alineamiento de la Figura 8) a partir del patrón de 10 µM, 0,3 µl de cebador inverso específico del gen (SEQ ID NO:29:5'- TTTCCATATCCATCGCAACA - 3 '); posiciones de nucleótidos 588 - 607 para SEQ ID NO:3 en el alineamiento de la Figura 8) a partir de 10 µM de patrón, 0,15 µl de sonda UPL # 25 de LIGHTCYCLER® 480 Probes Master, (Roche Diagnostic, Indianapolis, IN), 1,5 µl de polivinilpirrolidona-40 10 % (p/v) (PVP-40) y 3,9 µl de agua. La sonda UPL (Roche Diagnostics, Indianápolis, USA) es un ácido nucleico bloqueado y, por lo tanto, tiene una Tm superior a la calculada de otro modo. Todos los componentes se volvieron a colocar en el congelador antes de manipular los estándares y los desconocidos. Se demarcó y marcó una microplaca de 384 pocillos, se añadieron 13,5 µl de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3.000 rpm en una centrífuga de microplacas Qiagen. Se añadieron 1,5 µl de cadenas de ADNc sintetizado diluido y descongelado. Además, se añadieron 1,5 µl de estándares de número de copias de ADN plasmídico a pocillos separados en una serie de diluciones de concentraciones más bajas a más altas, estos estándares se compararon con el ADNc de β-cetoacil-ACP sintetasa II (sintetizado a partir de ARNm total) para cuantificar el número de copias. Se prepararon series estándar de número de copias de ADN de β-cetoacil-ACP sintetasa por clonación del amplicón blanco en un plásmido pCR2.1 (Invitrogen, Carlsbad, CA) y se realizó una serie de diluciones para cuantificar el número de copias. Se fijó firmemente un sello laminado a la placa y se centrifugó como se describió previamente. Un programa de PCR se realizó de la siguiente manera: i. Activar 95 °C durante 5 minutos; ii. Desnaturalizar 95 °C durante 10 segundos @ 4.8 °C/segundos; iii. Aparear/extender 60 °C durante 25 segundos a 2.5 °C/ segundos; iv. Adquirir 72 °C durante 1 segundos @ 4.8 °C/segundos; Repetir la etapa ii - iv, 40-50 más veces; Enfriar a 38 °C durante 5 segundos. El ADN se amplificó en instrumentación de PCR en tiempo real LC480 (Roche, Indianapolis, IN) o equivalente. El tamaño del amplicón fue de 64 pares de bases. Las secuencias de cebador directo e inverso empleadas en este ensayo de PCR coincidieron perfectamente con las secuencias correspondientes presentes en dos de los blancos del gen de β-cetoacil-ACP sintetasa II, SEQ ID NOs:3 y 30 (Figura 8). Por lo tanto, este ensayo representó la expresión cuantitativa de dos blancos del gen de β-cetoacil-ACP sintetasa II.

4.3 Análisis de la expresión de ARNm de tubulina

El gen de tubulina, un gen nativo de *Brassica napus* (GenBank:AF258790 y GenBank:DU106489), se usó como estándar de referencia para normalizar con exactitud la señal de expresión de ARNm a través de los genes en ensayos de qRT-PCR. El ADNc sintetizado para el ensayo qRT-PCR de β-cetoacil-ACP sintetasa II (descrito en el ejemplo 4.1) también se usó en el ensayo qRT-PCR de de tubulina como se describe a continuación.

Se montó un qRT-PCR con 0,3 µL de un cebador directo específico del gen (SEQ ID NO:31:5'-ACAGCGATTGCCTACAAGG -3') (patrón de 10 µM) y 0,3 µL de un cebador inverso específico del gen (SEQ ID NO:32:5'-AGATGGTTAAGATCACCAAAGG -3') (patrón de 10 µM), 1,5 µL de PVP-40 10% (p/v), 3,9 µL de agua y 7,5 µL de mezcla maestra 2X LIGHTCYCLER® 480 SYBR Green I (Roche, Indianapolis, IN) para detectar y cuantificar ADN. Se demarcó y marcó una microplaca de 384 pocillos, se añadieron 13,5 µl de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3.000 rpm en una centrífuga de microplacas Qiagen. Se retiró la lámina de sellado y se añadieron 1,5 µl de cadenas de ADNc sintetizado diluido y descongelado. Adicionalmente, se añadieron 1,5 µl de estándares de número de copias de ADN plasmídico a pocillos separados en una serie de diluciones de concentraciones más bajas a más altas, estos estándares se compararon con el ADNc de tubulina (sintetizado a partir de ARNm total) para cuantificar el número de copias. Se prepararon series estándar de número de copias de ADN de tubulina por clonación del amplicón blanco en un plásmido pCR2.1 (Invitrogen, Carlsbad, CA) y se realizó una serie de diluciones para cuantificar el número de copias. Se fijó firmemente un sello laminado a la placa y se centrifugó como se describió previamente. Un programa de PCR se realizó de la siguiente manera: i. Activar 95 °C durante 5 minutos; ii. Desnaturalizar 95 °C durante 10 segundos @ 4.8 °C/segundos; iii. Aparear/extender 55 °C durante 20 segundos a 2.5 °C/ segundos; iv. Adquirir 72 °C durante 20 segundos @ 4.8 °C/segundos; las etapas ii - iv se repitieron 39 veces más; vi. Enfriar a 38 °C durante 5 segundos. El ADN se amplificó en instrumentación de PCR en tiempo real LC480 o equivalente. Un amplicón de de 307 pares de bases se amplificó en esta reacción. Este cebador inverso abarca un intrón de 78 pb basado 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 pueden ser de mayor peso molecular y se pueden diferenciar fácilmente de los de ADNc mediante la corrida de los amplicones en un gel de agarosa.

4.4 Análisis de la expresión de ARNm en ZFP TF

La expresión del ARNm de ZFP TF se cuantificó a partir de muestras de ADNc que se sintetizaron originalmente para ARNm de β-cetoacil-ACP sintetasa II (descrito anteriormente en el Ejemplo 4.1). Se diseñó un ensayo de PCR TaqMan a partir del casete de ZFP mediante el anclaje de cebadores a la secuencia NLS *opaque-2* en la secuencia 5' y VP16 en el 3'. El ensayo de PCR se montó usando: 0,5 µl de cebador directo "OP2_NLS_F1" "OP2_NLS_F1" (SEQ ID NO:33:5'- AAGGAAGAGGAAGGAGTCTAACAG -3') a partir del patrón de 10 µM, 0,5 µL de cebador inverso "VP16_R1" (SEQ ID 34:5'- CTTCTGCTCTCCACCGTA -3') a partir del patrón de 10 µM, 0,25 µL de sonda "VP16_MGB_185" (SEQ ID NO:45:5'-TTGATGGTGAAGATGT -3') a partir del patrón de 5 µM, 1,5 µL de PVP-40 10% (p/v), 1,5 µL de tampón de PCR 10x Hot Start, 1,0 µL de MgCl₂ 25 mM, 1,2 µL de dNTP (2,5 mM cada uno), 0,15 µL de polimerasa Hot Start Taq (Qiagen, Valencia, CA), y 6,9 µL de agua, el cóctel se amplificó usando LIGHTCYCLER® 480 (Roche Diagnostics, USA). Se demarcó y marcó una microplaca de 384 pocillos, se añadieron

13,5 µl de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la microplaca. La placa se centrifugó durante 1 minuto a 3.000 rpm en una centrifuga de microplacas Qiagen. Se fijó firmemente un sello laminado a la placa y se centrifugó como se describió previamente. Un programa de PCR se realizó de la siguiente manera: i. Activar 95 °C durante 5 minutos; ii. Desnaturalizar 95 °C durante 20 segundos @ 4,8°C/segundos; iii. Aparear/extender 60°C durante 20 segundos @ 2,5°C/segundos; iv. Adquirir 72°C durante 55 segundos @ 4,8°C/segundos; las ii - iv se repitieron 44 veces; vi. Enfriar a 38°C durante 5 segundos. El ADN se amplificó en instrumentación de PCR en tiempo real LC480 o equivalente. El tamaño del amplicón fue de 775 pares de bases. Se determinó el número de copia de ADNc /ARNm de ZFP TF usando una serie estándar en la que el amplicón blanco se clonó en un plásmido y se prepararon series de dilución para ensayos de PCR como se describe para los ejemplos de β-cetoacil-ACP sintetasa II y tubulina anteriores.

4.5 Análisis de la expresión de las muestras de callo

La expresión de ARNm de ZFP TF se midió en muestras de callos transgénicos de 8 semanas transformados putativamente con TF ZFP y que crecen en medio de selección HERBIACE® (véase el Ejemplo 3). La expresión de ARNm cuantitativa de ZFP TF, medida con el ensayo de qRT-PCR, se detectó en muestras transformadas con constructos que contienen solo TF ZFP pero no en muestras que contenían el constructo de control. Estos datos indicaron que el ensayo era específico para la expresión de ZFP TF y que los callos de control contenían ek transgén de ZFP TF.

Las relaciones de β-cetoacil-ACP sintetasa II y la expresión de ARNm de tubulina se calcularon, sobre la base de los resultados de qRT-PCR, para discernir las diferencias de expresión entre muestras de callos. La regulación por aumento del ARNm más alta de β-cetoacil-ACP sintetasa II se observó en muestras de callos de canola transformadas con el constructo pDAB4695 (Tabla 4). Esta regulación por aumento de 27% fue estadísticamente significativa (p = 0.05) respecto de las muestras de control.

Tabla 4: Expresión de ARNm KASII en muestras de callo de *B. napus* transformadas con diferentes constructos de ZFP TF

Constructo	Número de callos analizados	Media de expresión de KasII/Tubulina	STDEV de 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 las muestras de planta T₀ transformadas

5.1 Análisis de ARNm de β-cetoacil-ACP sintetasa II, Tubulina y ZFP TF

Expresión

Se cultivaron plantas de canola T₀ que contienen el constructo del transgén pDAB4695 hasta la etapa de planta de 6 hojas y se analizaron las expresiones de ARNm del gen endógeno de β acil-ACP sintetasa II, gen endógeno de tubulina y transgén ZFP TF. Las plantas transgénicas se compararon con muestras de control que se habían transformado con un constructo binario que contiene un casete de expresión del gen *pat* (Promotor v2 AtUbi10 ::*pat* v3 ::Atu ORF1 3'UTR v3). Seis perforaciones de hojas de un perforador de papel de tamaño estándar de cada planta de muestreo en hielo y se extrajo el ARNm total con el kit de extracción de ARN RNeasy de 96 pocillos Qiagen de acuerdo con las instrucciones del fabricante. (Qiagen, Carlsbad, CA). Los análisis de expresión de ARNm de β-cetoacil-ACP sintetasa II, tubulina y ZFP se completaron usando los protocolos descritos anteriormente en el Ejemplo 4.

La expresión del ARNm de β-cetoacil-ACP sintetasa II varió entre las plantas T₀ transgénicas independientes. Se observó una regulación por aumento de ARNm superior a 3 veces entre los 16 eventos analizados (Figura 3). Todas las plantas fueron positivas para la expresión de ZFP TF. Además, se usaron 27 plantas de control para calcular una

línea base de expresión de ARNm de β -cetoacil-ACP sintetasa II. La expresión media del ARNm de β -cetoacil-ACP sintetasa II del control inicial se comparó con la expresión del ARNm de β -cetoacil-ACP sintetasa II de eventos de plantas ZFP TF individuales para calcular el nivel de aumento de la expresión (Figura 3). Estas plantas se cultivaron hasta la madurez para obtener la semilla T₁. Antes de la floración, todas las plantas se cubrieron individualmente con bolsas de autofecundación para facilitar la autopolinización de las flores dentro de una planta. Las semillas T₁ de estas plantas se recolectaron aproximadamente 4 meses después de su transferencia al invernadero.

Los eventos 3, 6 y 12.2 (de aquí en adelante denominado como el evento 12), que representan diferentes rangos de expresión de ARNm de β -cetoacil-ACP sintetasa II, se seleccionaron para un estudio adicional.

Ejemplo 6: Análisis de muestras de plantas T₁ transformadas

6.1 Análisis de ácidos grasos de la semilla T₁

El análisis de ácidos grasos de semilla individual se realizó en 24 semillas T₁ individuales, por evento, para estudiar los efectos de ZFP TF sobre la alteración en los contenidos de ácidos grasos (Tabla 5). Se empleó un procedimiento de análisis de éster metílico de ácidos grasos (FAME) basado en el procedimiento AOCS Ce2 - 66 (97) y todos los números en la Tabla 5 se muestran como un porcentaje de ácidos grasos totales presentes en las semillas de canola.

Tabla 5: Perfil de ácidos grasos de semillas T₁ individuales medido análisis FAME

Eventos	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C18 total
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
Desv. Est (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
Desv. Est (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

Eventos	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C18 total
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
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
Desv. Est (n = 24)	0,6	0,1	0,6	1,8	1,5	0,3	0,54
Nex710 Media (n=216)	3,8	0,4	2,3	78,1	9,5	2,5	92,5
Desv. Est (n=216)	0,2	0,0	0,6	1,8	1,6	0,7	0,45

5 Para la preparación de la muestra, las semillas individuales se colocaron en tubos cluster etiquetados en una placa de extracción de 96 pocillos que contiene una bola de acero de 1/8 "(Small Parts, Miramar, FL) Se taparon los tubos y las semillas se molieron en seco en un GenoGrinder (SPEX CertiPrep Group, Metuchen, NJ) durante 3,0 minutos a 1300 golpes/minuto. Se retiraron las tapas y se añadieron 0,6 ml de heptano a cada pocillo. Los pocillos se taparon de nuevo y se volvieron a colocar en el GenoGrinder para una molienda adicional durante 2,0 minutos a 1200 golpes/min. Las muestras luego se retiraron y centrifugaron a 3700 rpm durante 10,0 minutos a 6 °C. Usando un Beckman Coulter MC Robot, el sobrenadante se transfirió a una placa de 96 pocillos con insertos de vidrio (MicroLiter, Suwanee, GA). Se añadieron 40 µl de metóxido de sodio 1% a la muestra. El metóxido de sodio se diluyó a partir de una solución patrón 30% con metanol (Fluka/Sigma Aldrich, St. Louis, MO). Las placas se taparon con una placa revestida de teflón y se dejaron incubar a temperatura ambiente durante 4 horas antes del análisis de GC.

15 Las muestras se analizaron para determinar el contenido de ácidos grasos en un Agilent 6890 GC-FID (Agilent Technologies, Santa Clara, CA) equipado con una columna J & W Scientific DB-23, 15 metros x 0,25 mm de ID y 0,25 µm de espesor de película (J & W Scientific, Folsom, CA). La temperatura inicial del horno fue de 200 °C y esta temperatura se mantuvo durante la corrida. La entrada se ajustó a una relación de división de 1:100 y una temperatura de 280 °C. Se mantuvo una tasa de flujo en rampa de 0,8 ml/min de helio durante los dos minutos iniciales. Luego se aumentó el flujo a una tasa de 1,0 ml/min a 2,5 ml/min y se mantuvo durante 1,5 minutos. El detector se ajustó a 300 °C con una composición de gas de portador constante y flujo de columna de 30 ml/min, flujo de hidrógeno de combustible de 30 ml/min y flujo de oxidante de 400 ml/min. Se utilizó un volumen de inyección de 2 µl para todas las muestras.

25 Los picos de éster metílico de ácido graso individual se identificaron por comparación con los tiempos de retención de los estándares de referencia del éster metílico (GLC n° 428, Nu-Chek-Prep, Inc., Elysian, MN) usando Waters Corp. Empower Software. Se calcularon las áreas porcentuales individuales para todos los analitos en el estándar de referencia en función de las áreas del pico de cromatografía integradas totales. También se tomó un blanco de heptano para identificar cualquier contaminación en el GC.

30 Una comparación de la semilla T₁ con los tres niveles más bajos de C16:0 (que muy probablemente son plantas ZFP TF positivas) y los tres niveles más altos de C16:0 (que son más probablemente las plantas ZFP TF nulas) indicaron que los cambios en el contenido de C16:0 se puede deber 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 fue a pesar de la variabilidad en los contenidos de 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., y col., Theor. Appl. Genet 2006, 113:497-507) en el genotipo Nex710 que se transformó. Por lo tanto, se llevó a cabo la siguiente etapa para identificar las plantas ZFP TF positivas y nulas hermanas en una población segregante T₁ de cada evento para detectar el cambio real en el perfil de ácidos grasos en trasfondos genéticos similares.

6.2 Análisis de presencia de ZFP TF en plantas T₁

Se analizaron 100-150 plántulas T₁ de los tres eventos transgénicos 3, 6 y 12 para identificar las plantas ZFP TF positivas y nulas hermanas (Tabla 6). Las plantas se analizaron para detectar la presencia de al menos un casete de longitud completa del transgén ZFP TF. El ADN genómico total de las hojas de estas 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 a una concentración final del 1%. El ADN_g purificado se cuantificó mediante un protocolo de cuantificación de ADN PICOGREEN® (Molecular Probes, Inc., Eugene, OR). Las muestras de ADN se analizaron para determinar la presencia de cómo mínimo una copia de longitud completa del transgén de ZFP TF en un ensayo de PCR que contiene los siguientes reactivos: 5 µl de tampón Taq 10X EX (Takara Bio Inc., Otsu, Siga, Japón), 5 µl de polivinilpirrolidona-40 10%, 1,0 µl de cebador directo ubi10 (SEC ID N°:35:5'-GGTCAACGGATCAGGATATTCTTG -3') a partir del patrón de 10 µM, 1,0 µl de cebador inverso AtuORF23 (SEC ID N°:36:5'- CCATGTTGGCAAAGGCAACC -3'), patrón 10 µM, 4 uL de dNTP (2.5 mM cada uno), 0.2 µl de TaKaRa EX Taq™ Hot Start, 31.8 µl de agua y 2 µl de ADN a una concentración de 10 ng/µl. Las condiciones de ciclado de PCR fueron las siguientes: 94 °C durante 2 min, 10 ciclos de PCR touch-down (decreciente) con 98 °C durante 10 segundos, 65 °C durante 20 segundos con una temperatura decreciente de 0,5 °C en cada ciclo a 60 °C seguido a 72 °C durante 3:00 min. Esto fue seguido por 35 ciclos de 98 °C durante 10 segundos, 60 °C durante 20 segundos y 72 °C durante 3 minutos y una extensión final de 72 °C durante 10 minutos. Se corrieron 5 µl de cada reacción en un gel de agarosa 1% para detectar la presencia del tamaño esperado de la banda ZFP TF de 2662 pb.

Se desarrolló un ensayo de PCR en tiempo real cuantitativa para el gen *pat*, que está unido molecularmente a TF ZFP en el constructo pDAB4695, y se aplicó para identificar plantas *pat* positivas y nulas para confirmar su segregación genética con el casete ZFP TF al que está unido molecularmente.

Tabla 6: Identificación de plantas ZFP TF positivas y nulas hermanas en la población T₁ segregante de los tres eventos

Eventos	Semilla T ₁ total plantada	Plantas T ₁ ZFP TF/nulas identificadas	Plantas T ₁ ZFP TF/nulas avanzadas a semilla
3	150	115/7	12/7
6	100	66/27	5/5
12	150	126/6	8/6

La mezcla de reacción para el ensayo de TaqMan RT-PCR *pat* estaba compuesta de los siguientes: 3 µL de mezcla maestra Roche LIGHTCYCLER® Cycloer 480 II Probes 2X, 0,6 µL de PVP-40 10%, 0,2 µL de cada uno del cebador directo de *pat* (SEQ ID NO:37:5'-ACAAGAGTGGATTGATGATCTAGAGAGGT -3') a partir del patrón de 10 µM, 0,2 µL de cebador inverso PAT (SEQ ID NO:38:5'- CTTTGATGCCTATGTGACACGTAACAGT -3') a partir del patrón de 10 µM, 0,2 µL de sonda (SEQ ID NO:39:5'-6FAM-CCAGCGTAAGCAATACCAGCCACAACACC-inactivador) a partir del patrón de 5 µM, y se multiplexó con los siguientes reactivos para el estándar interno de referencia de HMG (GenBank:AF127919), de la siguiente manera: 0,2 µL de cada cebador directo de HMG (SEQ ID NO:40:5'-CCTCTCTACCACCGTCTCACATG -3') a partir del patrón de 10 µM, 0,2 µL de cebador inverso HMG (SEQ ID NO:41:5'- GATCTGGCCGGACTGTTTCA -3') a partir del patrón de 10 µM, 0,2 µL de sonda (SEQ ID NO:42:5'-6FAM-CGCTCCTCAGCTACCACCTCAACCA-inactivador-3'), a partir del patrón de 50 µM, 1 µL de ADN y 0,2 µL de agua. Las condiciones de ciclado de PCR fueron: 1 ciclo de 95 °C durante 5 min, 40 ciclos de 95°C durante 10 segundos, 60°C durante 40 segundos y 72°C durante 1 segundo. Todas las muestras se amplificaron por triplicado en placas de 384 pocillos justo con 1 a 4 copias de estándares de ADN genómico transgénico en una máquina Roche LIGHTCYCLER® 480.

Sobre la base de las relaciones de segregación de las plantas ZFP TF positivas y nulas dentro de cada evento, los eventos 3, 6 y 12 contenían aproximadamente 2, 1 y 2 inserciones respectivamente del transgén de ZFP TF en el genoma de *B. napus* (Tabla 6, Columna 3) Estos datos coinciden con los datos de segregación de genes *pat* en los tres eventos.

6.3 Regulación por aumento de ARNm de β-cetoacil-ACP sintetasa II en las plantas T₁

Todas las plantas T₁ ZFP TF positivas y nulas hermanas mostradas en la Tabla 6, Columna 3 se sometieron luego al análisis de expresión de ARNm para el gen de β-cetoacil-ACP sintetasa endógeno, transgén de ZFP TF y gen de tubulina endógeno. Este último sirvió como un gen de referencia para normalizar las expresiones del gen de β-cetoacil-ACP sintetasa II y ZFP TF. Seis perforaciones de hoja de cada una etapa de la planta de 6 hojas se

muestraron en hielo y se extrajo el ARN total utilizando el kit QIAGEN RNeasy®. La síntesis de cadena de ADNc y las diluciones se completaron como se describe en el Ejemplo 4.1.

El análisis de qRT-PCR del ADNc sintetizado se completó para los tres genes. El análisis de expresión del ARNm de β -cetoacil-ACP sintetasa II y tubulina se realizó como se describe en los Ejemplos 4.2 y 4.3, respectivamente. El ensayo ZFP TF se modificó. La reacción de PCR se montó de la siguiente manera: Las condiciones de reacción fueron las siguientes: 7,5 μ l de tampón maestro 2X LC480 Probes (Roche Diagnostic, Indianápolis, IN), 0,3 μ l de cebador directo específico del gen #1 (SEQ ID NO:43:5'- TCGATCTTGATATGTTGGGAGA -3') (patrón 10 μ M), 0,3 μ l de cebador inverso específico del gen #2 (SEQ ID NO:44:5'- AGGTGCAGAATCATGTGGTG - 3') (patrón 10 μ M), 0,15 μ L de sonda UPL #85, 1,5 μ l de PVP-40 10% (p/v) y 3,9 μ l de agua. Se demarcó una microplaca de 384 pocillos y se añadieron 13,5 μ l de la mezcla descrita anteriormente por pocillo. Se fijó una lámina de sellado a la placa, y la placa se centrifugó durante 1 minuto a 3000 rpm. La película se retiró y se añadieron 1,5 μ L de las primeras cadenas diluidas y descongeladas por pocillo. Además, se añadieron estándares de ADNc a pocillos de control separados. El sello laminado se selló sobre la placa y se repitió la etapa de centrifugación. Se corrió un programa de PCR con las siguientes condiciones; 1. Activar 95 °C durante 5 minutos, ii. Desnaturalizar 95 °C durante 10 segundos (@ 4,8 °C/segundos), iii. Aparear/extender 60°C durante 25 segundos (@ 2,5°C/segundos), iv. Adquirir 72°C durante 1 segundos (@ 4,8°C/segundos), v. Repetir las etapas ii - iv 39-49 veces más, vi. Enfriar a 38°C durante 5 segundos.

La expresión de ARNm de ZFP TF fue 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 expresaron ARNm de ZFP TF lo que indica que el ensayo ZFP TF era específico para la presencia de transgén de ZFP TF. Los niveles de ARNm del gen de la β -cetoacil-ACP sintetasa II, el blanco para ZFP TF, se regularon por aumento en las plantas ZFP TF positivas de todos los eventos (Figura 5). Se observó hasta un aumento de 3-4 veces en la expresión de ARNm de β -cetoacil-ACP sintetasa II mediante comparaciones por pares de las plantas ZFP TF positivas y las correspondientes nulas hermanas dentro de cada evento. La regulación por aumento de los niveles de ARNm de β -cetoacil-ACP sintetasa II fue relativa al aumento en los niveles de expresión de ARNm de ZFP TF dentro de cada evento (Figura 4), es decir, más bajo en el evento 6 y más alto en el evento 3.

6.4 Análisis de ácidos grasos de las hojas de las plantas T₁

Se recogió un segundo conjunto de muestras de hojas T₁ para análisis de ácidos grasos concurrentes a la recolección de muestras de análisis de ARNm descritas en el Ejemplo 6.3. Sin embargo, solo un subconjunto de las muestras se analizó para ácidos grasos solo para aquellas plantas que avanzaron a la madurez como se muestra en la Tabla 6, columna 4. El análisis de éster metílico de ácido graso (FAME) del material foliar se completó de la siguiente manera. Se recogió material de hoja de 10-100 mg en hielo y luego se liofilizó 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 contiene heptadecanoína como un estándar 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 usando una columna capilar BPX 70 de SGE (15 m x 0,25 mm x 0,25 μ m). Cada FAME se identificó por su tiempo de retención y se cuantificó mediante la inyección de una mezcla de referencia de aceite de colza de Matreya LLC como estándar de calibración. La completitud de la reacción se verificó mediante el control de la presencia de FAMES endógenos en una cuarta extracción/derivación.

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

Tabla 7: Perfil de ácidos grasos de hoja T₁

Nombre de muestra	Plantas analizadas	C16:0		C18 total	
		Media	Desv. Est	Media	Desv. Est
12 nulo	6	10,98	0,31	69,93	0,91
12ZFP	8	10,29	0,39	70,94	1,02
3 nulo	7	10,79	0,56	69,84	1,18
3ZFP	12	10,34	0,29	71,25	0,74

Nombre de muestra	Plantas analizadas	C16:0		C18 total	
		Media	Desv. Est	Media	Desv. Est
6 nulo	5	11,94	0,47	70,36	1,91
6ZFP	5	10,44	0,62	72,16	2,88

Ejemplo 7: Análisis de semilla T₂ transgénica ZFP TF

5 Un subconjunto de plantas ZFP TF positivas que demostraron las expresiones más altas de ARNm de β -cetoacil-ACP sintetasa II y plantas nulas hermanas dentro de cada evento que avanzaron a la madurez para obtener las semillas T₂ (Tabla 6, columna 4). La excepción fue el evento 3, en el que también avanzaron algunas plantas positivas para ZFP TF con rangos de expresión más bajos. Estos podrían representar la otra inserción segregante del transgén pDAB4695. Cada planta T₁ se cubrió con una bolsa de malla para facilitar la autopolinización dentro de una planta determinada. Se recolectaron las semillas y se analizaron los niveles de ácidos grasos.

10 El análisis de ácidos grasos T₂ se realizó en una mezcla de 24 semillas por planta como se describe para el análisis de semillas T₁ en el Ejemplo 6.1. Los resultados mostraron una disminución en el contenido de C16:0 en plantas ZFP TF positivas de los tres eventos por comparación por pares con sus correspondientes plantas nulas hermanas (Tablas 8A y 8B). Esta disminución en los contenidos de C16:0 fue del 12.6%, 11.2% y 22.3% para los eventos 12, 3 y 6, respectivamente. También se observó una disminución constante en el contenido de C16:1. Correspondiente a la disminución en los contenidos de C16:0 y C16:1, se observó un aumento concomitante en el contenido de C18 total (C18:0 + C18:1 + C18:2 + C18:3) en plantas ZFP TF positivas en una comparación por pares con las plantas nulas hermanas correspondientes. Las tres comparaciones por pares de las plantas ZFP TF positivas y las correspondientes nulas hermanas fueron estadísticamente significativas ($p < 0.001$) para aumentos en el contenido de C16 total (C16:0 + C16:1) y contenido de C16 total/C18 total en el análisis con el software estadístico JMP (Figura 6, A y B). Aunque los tres eventos fueron variables en su contenido de C16:0 basado en observaciones de plantas nulas hermanas (Tablas 8A y 8B), ZFP TF fue eficaz en la disminución del contenido de C16:0 a niveles aproximadamente similares; entre ~ 2.93 a 3.05. Mientras que todas las plantas ZFP TF positivas de los eventos 3 y 12 establecen las semillas (Tablas 8 y 9), solo una planta que contiene ZFP TF del evento 6 estableció la semilla. Esta planta particular del evento 6 fue la de contenido más alto de C16:0 sobre la base de los datos de ácidos grasos de las hojas T₁. Las Tablas 8A y 8B muestran la comparación del perfil de ácidos grasos T₂ de tres eventos transgénicos de *B. napus* de pDAB4695 ZFP TF positivo y las correspondientes plantas nulas hermanas. Estos datos se obtuvieron con el análisis FAME y se describe en la Sección 6.1. Se analizaron los ácidos grasos C12:0 a C24:1 y se muestran los principales en estas tablas. Todos los números se representaron como un porcentaje de ácidos grasos totales presentes en las semillas de *B. napus*.

Tabla 8A: Análisis de ácidos grasos C16:0, C16:1, C18:0, C:18:1, C:18:2 y C:18:3

Muestras	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
12 Nulo Media	3,49	0,26	1,29	76,59	11,45	3,73
Desv. Est (n=6)	0,06	0,01	0,05	0,25	0,21	0,04
12ZFP Media	3,05	0,20	1,33	76,53	11,63	3,87
DEV EST (n=8)	0,11	0,02	0,04	0,39	0,33	0,06
3Nulo Media	3,30	0,25	1,31	74,34	13,18	3,94
Desv. Est (n=7)	0,17	0,02	0,10	1,04	0,86	0,25
3ZFP Media	2,93	0,18	1,30	74,10	13,62	4,14

ES 2 646 138 T3

Muestras	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
DEV EST						
(n=12)	0,12	0,01	0,08	1,80	1,45	0,25
6Nulo Media	3,77	0,38	2,02	73,46	12,42	3,66
Desv. Est (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

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
12Nulo Media	0,53	1,26	0,05	0,33	0,02	0,14	0,14
Desv. Est (n=6)	0,03	0,07	0,01	0,04	0,01	0,02	0,02
12ZFP Media	0,55	1,37	0,06	0,36	0,02	0,15	0,16
DEV EST (n=8)	0,02	0,05	0,01	0,02	0,01	0,01	0,02
3Nulo Media	0,59	1,49	0,06	0,43	0,02	0,15	0,21
Desv. Est (n=7)	0,07	0,11	0,01	0,07	0,01	0,03	0,04
3ZFP Media	0,57	1,57	0,07	0,42	0,03	0,15	0,19
DEV EST (n=12)	0,04	0,21	0,02	0,07	0,01	0,04	0,04
6Nulo Media	0,80	1,49	0,07	0,54	0,03	0,30	0,24
Desv. Est (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

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

Tabla 9

Muestras	C18 total*	Total LC**	Sat totales***
12 Nulo Media	93,06	95,52	5,83

Muestras	C18 total*	Total LC**	Sat totales***
DEV EST (n=6)	0,13	0,05	0,09
12ZFP Media	93,36	96,01	5,48
DEV EST (n=8)	0,11	0,13	0,09
3 Nulo Media	92,78	95,74	5,82
DEV EST (n=7)	0,39	0,17	0,32
3ZFP Media	93,16	96,14	5,41
DEV EST (n=12)	0,41	0,13	0,19
6 Nulo Media	91,55	95,01	7,47
DEV EST (n=4)	0,36	0,29	0,40
6ZFP (n=1)	92,85	96,18	5,17

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

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

***Sat totales representa todos los ácidos grasos saturados mostrados en las Tablas 8A y 8B.

Combinando todos los ácidos grasos de cadena larga (C18 total, C20 total, C22 total y C24 total), se observó un aumento de 0.4% a 1.2% en las plantas ZFP TF positivas de diferentes eventos en comparación con las nulas hermanas. Entre los ácidos grasos de cadena más larga, se observó un aumento de C20:1 en un 25% en plantas ZFP TF positivas para el evento 6. También se observó una disminución en el contenido de ácidos grasos saturados totales en plantas que contienen ZFP TF en todos los eventos. Varió de una reducción del 6% en el evento de 12 a 31% de reducción en el evento 6.

En resumen, se ejemplificó la regulación por aumento transcripcional mediada por ZFP TF del gen de β -cetoacil-ACP sintetasa II en plantas T₀ y T₁ y la disminución concomitante del contenido de ácidos grasos totales de C16 y aumento de C18 en la semilla T₂. Estos datos demuestran la aplicación exitosa de nuevas tecnologías ZFP TF para dirigir y modificar genes específicos en la vía de biosíntesis de ácidos grasos, lo que produce cambios predichos en los perfiles de aceite de las semillas, que son heredables a través de generaciones de progenie.

Ejemplos 8-15: Regulación por aumento/Regulación por disminución

Los plástidos son el sitio principal para la biosíntesis *de novo* de ácidos grasos en plantas superiores (Ohlrogge et al., 1979, Proc Natl Acad Sci U S A 76:1194-8; Thelen and 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 acil-ACP tioesterasas (FAT) mediante la hidrolización de las proteínas portadoras de acil-acilo (acil-ACP) para liberar ácidos grasos libres. Hay dos clases de enzimas FAT en plantas, FATA y FATB. La clase FATA prefiere 18:1-ACP *in vitro* mientras que la clase FATB prefiere los sustratos acil-ACP saturados, tales como 16:0-ACP y 18:0-ACP, pero también muestra otra especificidad de sustrato heterogéneo Doermann et al, 1995, Arch. Biochem. Biophys. 316:612-618; Voelker et al., 1997, Plant Physiology 114:669-677; Salas and Ohlrogge, 2002, Archives of Biochemistry and Biophysics, 403:25-34). De modo similar a FATA, también se considera que FATB 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).

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

El principal cultivo de semillas oleaginosas, *Brassica napus*, es una especie anfiploide estrechamente relacionada con la especie modelo *Arabidopsis*, pero tiene un tamaño del genoma aproximadamente ocho veces mayor que el de *Arabidopsis*. Como resultado, se informaron seis genes *FatB* en esta especie (documento WO 2009/007091) en comparación con uno en *Arabidopsis*. Esta complejidad hace que sea difícil comprender la función de los genes *FatB* individuales para su manipulación. Por el contrario, la presencia de un mayor número de genes hace que sea más fácil manipular la expresión de múltiples genes hacia el objetivo de ácidos grasos deseado en lugar de inactivar uno por completo. Se sabe que los factores de transcripción de dedos de cinc se unen esencialmente a cualquier secuencia de ADN y modulan la función génica de los genes dirigidos (Van Eenennaam et al., *Metab. Eng.* 2004 6:101-108; Sanchez 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 puede ayudar a manipular la expresión de uno a múltiples genes *FatB* hacia el aceite de semilla de canola "más saludable" deseado que tenga menor contenido de ácidos grasos saturados.

Los ejemplos 8-15 siguientes demuestran la regulación por aumento transcripcional de los genes *FatB* en *Brassica napus* L. lo que produce cambios en el perfil de ácidos grasos en las semillas.

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

8.1 Identificación de secuencia blanco

Se empleó un protocolo modificado de Genome Walking para descubrir y caracterizar las secuencias promotoras de los genes *FatB* de la variedad Nex710 de *Brassica napus*. Estas secuencias de nucleótidos se aislaron e identificaron para el diseño y la producción de las ZFP TF, que se pueden usar para modificar la expresión génica del gen *FatB*. Se construyeron nueve bibliotecas Genome Walking usando el kit de Clontech Genome Walking (Palo Alto, CA). Estas bibliotecas se usaron para obtener ~ 1 kb de secuencia corriente arriba del sitio de inicio de la transcripción para los genes *FatB*. Se diseñó y sintetizó un par de cebadores directos específicos de genes *FatB* anidados sobre la base de las secuencias EST *FatB* disponibles de bases de datos públicas. Se diseñó y sintetizó un par de cebadores inversos anidados para hibridar con las secuencias adaptadoras. Se amplificaron múltiples fragmentos de PCR de las nueve bibliotecas. Los fragmentos de PCR resultantes se clonaron en vectores TOPO (Invitrogen, Carlsbad, CA, EE. UU.) Y se secuenciaron para identificar las secuencias del promotor *FatB*. Se identificaron dos genes *FatB* diferentes, *FatB4* y *FatB*, basados en este enfoque. Se observó que aunque las secuencias codificadoras de estos genes *FatB* estaban altamente conservadas, las secuencias divergían en las regiones 5'UTR y promotoras.

La expresión de *FatB* y la cuantificación de los genes *FatB4* y *FatB* a través del análisis de RT-PCR cuantitativo en las semillas de *B. napus* en desarrollo sugirió que los genes estaban altamente expresados en la semilla de *B. napus*. El ARN total se aisló, se cuantificó y los sitios de iniciación de la transcripción se mapearon como se describió anteriormente en el Ejemplo 1. Las secuencias promotoras de estos dos genes se utilizaron para diseñar ZFP TF para la modificación de la expresión transcripcional.

Ejemplo 9: Diseño de dominios de unión al ADN de ZFP específicos para los genes *FatB*

Se diseñaron proteínas de dedos de cinc contra varios sitios blanco del gen *FatB*. Los sitios blanco 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 blanco para ZFP TF de *FatB*

ZFP	Sitio blanco <i>FatB4</i> (5' a 3')	Sitio blanco <i>FatB</i> (5' a 3')	Especificidad de <i>FatB</i>
13685*	aaCGAAAGgAGATCGAGAGAGg agagag (SEQ ID NO:47)	Sin sitio de unión	<i>FatB4</i>
13714	Sin sitio de unión	cgAAAGGGAGATCGAGAGAG gcaccgca (SEQ ID NO:48)	<i>FatB</i>

ZFP	Sitio blanco <i>FatB4</i> (5' a 3')	Sitio blanco <i>FatB</i> (5' a 3')	Especificidad de <i>FatB</i>
13722	aaGGAGAAcTTTAGGGTTTGgg agact (SEQ ID NO:49)	aaGGAGAA/TTTAGGGTTTGG ggagact (SEQ ID NO:50)	FatB4 y <i>FatB</i>
13743**	ctCCGAAGAGATTGGCGTAAcac ttcgt (SEQ ID NO:51)	ctCCGAAGAGATTGGCGTAAc cttcatt (SEQ ID NO:52)	FatB4 y <i>FatB</i>

Tabla 10B: Regiones de la hélice de reconocimiento 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	RSDHLSL (SEQ ID NO:82)	HSNTRKN (SEQ ID NO:83)	RSDHLSQ (SEQ ID NO:84)	NSASRKN (SEQ ID NO:85)	QSGNLAR (SEQ ID NO:16)	QSGHLSR (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 *FatB*, pero la secuencia flanqueante 3' es diferente entre los dos genes.

*** los números representan los dedos de cinc (1 es dedos 1, 2 es dedos 2, etc.)

Ejemplo 10: Regulación por aumento mediada por ZFP TF de los genes de *FatB* nativos en *B. napus*

- 5 Para ejemplificar la-regulación por aumento mediada por ZFP TF de *FatB* dentro de las células de *B. napus*, se construyeron constructos que contienen cuatro diseños de genes ZFP TF (Tablas 10 y 11), estos genes se administraron de forma estable en células de *B. napus* a través de la transformación mediada por *Agrobacterium*.

10.1 Diseño del constructo

- 10 Los plásmidos binarios ZFP TF diseñados y construidos se enumeran en la Tabla 11 a continuación. La expresión de los genes ZFP TF fue dirigida por un promotor constitutivo relativamente fuerte, tal como un promotor derivado del promotor del virus del mosaico de las nervaduras de la yuca (CsVMV). La transformación de la planta mediada por *Agrobacterium* se realizó como se describe en el Ejemplo 3.2 para incorporar de forma estable TF ZFP en el genoma de *B. napus*.

Tabla 11: Descripción de ZFP TF dirigidas a los genes *FatB* de *B. napus*

S.N.	ZFP	Constructo No.	Casete del gen
1	13685	pDAB4689	RB7 MAR/CsVMV/Op-2 NLS-ZFP 13685 VP16/AtuORF23 3' UTR/AtUbi10/Pat/AtuORF1 3' UTR

S.N.	ZFP	Construc ^{to} No.	Casete del gen
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

pDAB4689 es un plásmido binario que contiene la secuencia de localización nuclear *opaque-2* (Op-2) o los casetes de expresión del gen NLS /13685/ VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 (Región de unión de la matriz (Thompson et al., 1997, WO9727207)) ::Promotor v2 CsVMV (Promotor del virus de mosaico de las nervaduras de la yuca (Verdaguer et al., 1996, Plant Molecular Biology 31:1129-1139)) ::NLS *opaque-2* (Van Eenennaam et al., 2004, Metabolic Engineering 6:101-108; Holmes-Davis et al., 2005, Plant Molecular Biology 57:411-423)/13685 Dedos de cinc/VP16 (Jamieson et al., Biochem. Biophys. Res. Commun. 2006, 348:873-879) Fusión ::AtuORF23 3'UTR v1 (Marco de lectura abierto 23, región 3' no traducida de *Agrobacterium tumefaciens* (Gelvin et al., 1987, EP222493))::Promotor v2 AtUbi10 (Promotor de Ubiquitina-10 de *Arabidopsis thaliana* (Callis, et al., 1990, J. Biol. Chem. 265-12486-12493)) ::*pat* v5 (Fosfinotricina acetil Transferasa (Wohlleben et al., 1988, Gene 70:25-37)) ::AtuORF 1 3'UTR v4 (Marco de lectura abierto 1, región 3' no traducida de *Agrobacterium tumefaciens* (Huang et al., J. Bacteriol. 172:1814-1822)). El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión NLS *opaque-2* /13685 Dedos de cinc/VP16 en pDAB3912 por medio de los sitios de restricción NcoI - SacI. El constructo resultante que se marcó como pDAB8215 contenía el casete de expresión del gen fusión de Promotor v2 CsVMV ::NLS *opaque-2* /13685 Dedos de cincNP16 Fusión ::AtuORF23 3'UTR v1. pDAB8215 se clonó en el binario pDAB4668 por medio 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.

pDAB4690 es un plásmido binario que contiene los casetes de expresión génica de NLS *opaque-2*/13714NP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor v2 CsVMV ::Fusión NLS *opaque-2*/13714 Dedos de cinc/VP16 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión del NLS *opaque-2*/13714 Dedos de cinc/VP16 en pDAB3912 por medio de los sitios de restricción NcoI - SacI. El constructo resultante que se marcó como pDAB8216 contenía el casete de expresión del gen de Promotor v2 CsVMV :: Fusión de NLS *opaque-2* /13714 Dedos de cinc/VP16 ::AtuORF23 3'UTR v1. pDAB8216 se clonó en el binario pAB4668 por medio 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.

pDAB4691 es un plásmido binario que contiene los casetes de expresión génica de NLS *opaque-2*/13722NP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor v2 CsVMV ::Fusión NLS *opaque-2*/13722 dedos de cincLS/VP16 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 (Fosfinotricina Acetil transferasa ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión de NLS *opaque-2*/13722 dedos de cincLS/VP16 en pDAB3912 por medio de los sitios de restricción NcoI - SacI. El constructo resultante que se marcó como pDAB8217 contenía el casete del gen Promotor v2 CsVMV ::Fusión de NLS *opaque-2* /13722 dedos de cincLS/VP16 ::AtuORF23 3'UTR v1. pDAB8217 se clonó en el binario pDAB4668 por medio 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.

pDAB4692 es un plásmido binario que contiene los casetes de expresión génica de NLS *opaque-2*/13743/VP 16 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 (Región de unión de la matriz ::Promotor v2 CsVMV ::Fusión NLS *opaque-2*/13743 Dedos de cinc/VP16 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 (Promotor de Ubiquitina-10 *Arabidopsis thaliana* ::*pat* v5 ::AtuORF 1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión de NLS *opaque-2*/13743 Dedos de cinc/VP16 en pDAB3912 por medio de los sitios de restricción NcoI - SacI. El constructo resultante que se marcó como pDAB8218 contenía el casete de expresión génica del Promotor v2 CsVMV ::NLS *opaque-2* /13743 Dedos de cinc/VP16 ::AtuORF23 3'UTR v1. pDAB8218 se clonó en el binario pDAB4668 por medio 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 callo transformado

Las líneas de callos de *B. napus* transgénicas se produjeron mediante la transformación de cuatro constructos de ZFP TF, pDAB4689 - pDAB4692, y un constructo de control, pDAB8210, a través de la transformación de plantas mediada por *Agrobacterium* como se describe en el Ejemplo 3.2. El constructo de control, pDAB8210, estaba compuesta por un casete de expresión del gen *pat* (Promotor AtUbi10/*pat*/AtuORF 1 3' UTR) y un casete ZFN no blanco (promotor CsVMV/ZFN/AtuORF23 3' UTR). Se extrajo el ARN total y se sintetizó ADNc de todas las líneas como se describe en el Ejemplo 4.1. La única modificación del protocolo descrito anteriormente fue el uso de cebadores Oligo dT para cebar las reacciones de ADNc en lugar de hexámeros aleatorios y la modificación correspondiente de la reacción de ADNc, según las especificaciones del fabricante.

11.1 Análisis de expresión de ARNm de *FatB4* y *FatB*11.1.1 Ensayo PCR (qRT-PCR) de tiempo real cuantitativa *FatB4* para *B. napus*

Se preparó una mezcla de PCR de la siguiente manera para la amplificación de ADNc de *FatB4*: 1,5 µL de tampón PCR 10X Hot Start (Qiagen, Valencia, USA), 1,2 µL de dNTP 10 mM, 1 µL de MgCl₂ 25 mM, 0,15 µL de Qiagen Hot Start Taq (5 U/µL), 0,5 µL de cebador zz143 FW del patrón 10 µM (SEQ ID NO:53:CTTTGAACGCTTATCTTCCTC) (patrón 10 µM), 0,5 µL de cebador REV *FATB_R4* del patrón 10 µM (SEQ ID NO:54:TTCCACAACATCTCCCAAG), 0,25 µL de Sonda TaqMan MGB (Life Technologies, Carlsbad, California) *FatB4_MGB_Probe_4* del patrón de 5 µM (SEQ ID NO:55:FAM-CTCAGGCTCCACCC), 1,5 µL de PVP-40 10% (p/v), y H₂O a 13,5 µL por reacción. Los cuadrantes apropiados de una microplaca de 384 pocillos se demarcaron y se cargaron con 13,5 µL de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la placa. La placa se centrifugó durante 1 minuto a 3,000 rpm en una centrífuga de microplacas Qiagen. A continuación, se añadieron 1,5 µL de las primeras cadenas de ADNc sintetizado diluido y descongelado a pocillos apropiados, seguido por la adición de 1,5 µL de estándares de ADNc de plásmido desde la concentración más baja a más alta 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 corrió en el sistema PCR de tiempo real LIGHTCICLOR® 480 (Roche Diagnostics, Indianápolis, IN, USA), usando el siguiente programa: 1) activar 95 °C durante 15 minutos; 2) desnaturalizar 95 °C durante 20-30 segundos (@ 4,8°C/segundos); 3) aparear 60 °C durante 20-30 segundos (@ 2,5°C/segundos); 4) adquirir 72°C durante 45-60 segundos (@ 4,8°C/segundos); repetir etapa 2) -- 4), 39-49 veces mas; 5) enfriar a 38°C durante 5 segundos para detener la reacción.

El amplicón de *FatB4* tenía 678 bp de tamaño. Esta secuencia abarca un intrón de 79 pb basado en la secuencia genómica. Además, el cebador inverso se diseñó para que abarcara un intrón, de este modo se favorece la amplificación específica de solo ADNc de *FatB4*, así se elimina la amplificación del ADN genómico.

11.1.2 Ensayo de qRT-PCR de *FatB* para *B. napus*

Se preparó una mezcla de PCR de la siguiente manera para la amplificación de ADNc de *FatB*: 1,5 µL de tampón PCR 10X Hot Start (Qiagen, Valencia, USA), 1,2 µL de dNTP 10 mM, 1 µL de MgCl₂ 25 mM, 0,15 µL de Qiagen Hot Start Taq (5U/µl), 0,5 µL de cebador FW zz145 de patrón de 10 µM (SEQ ID NO:56:CTTTGAAAGCTCATCTTCCTC), 0,5 µL de cebador REV *FATB_R4* del patrón de 10 µM (SEQ ID NO:57:TTCCACAACATCTCCCAAG), 0,25 µL de Sonda TaqMan MGB (Life technologies, Carlsbad, California) *FatB_MGB_Probe_1* del patrón de 5 µM (SEQ ID NO:58:FAM-AACCTTCATCCTCCCA), 1,5 µL de PVP-40 10% (p/v)), y H₂O a 13,5 µL por reacción. El ensayo restante y las especificaciones del ciclo PCR fueron como se describió para el ensayo de qRT-PCR *FatB4* en el Ejemplo 11.1.1.

El amplicón producido tenía 678 bp de tamaño. Esta secuencia de ADNc abarca un intrón de 76 pb basado en la secuencia genómica. Además, el cebador inverso se diseñó para que abarcara un intrón, de este modo se favorece la amplificación específica de *FatB*, así se elimina la amplificación del ADN genómico.

11.2 Análisis de expresión de ARNm de tubulina

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

11.3 Análisis de expresión de ARNm de ZFP TF

Se completó el análisis de ARNm de *ZFP TF* de la muestra de T₀ como en el Ejemplo 4.4. El tamaño de amplicón en este ensayo fue menor de 1 Kb de acuerdo con el tamaño de ZFP en cada uno de los diseños. La cuantificación de la expresión de *ZFP TF* de la planta T₀ se realizó con un ensayo basado en el dominio de activación de VP16 de la siguiente manera. Se preparó una mezcla de qRT-PCR VP16 como sigue para la amplificación de cDNA de *ZFP TF*: 7,5 µl de tampón 2X LIGHTCICLOR® 480 Probes Master (Roche Diagnostics, Indianápolis, IN, USA), 0,3 µL de cebador VP16_UPL_F1 del patrón de 10 µM (SEQ ID NO:59:TCGATCTTGATATGTTGGGAGA), 0,3 µL de VP16_UPL_R1 del patrón de 10 µM (SEQ ID NO:60:AGGTGCAGAATCATGTGGTG), 0,15 µL de UPL Probe#85 (Roche Diagnostics, Indianápolis, IN, USA), 1,5 µL de PVP-40 10% (p/v)), y H₂O a 13,65 µL. Los cuadrantes apropiados de una microplaca de 384 pocillos se demarcaron y se cargaron con 13,5 µL de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la placa. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrífuga de microplacas Qiagen. Se añadieron primeras cadenas de ADNc sintetizado diluido y descongelado.

A 1,5 µL, seguido por la adición de 1,5 µL estándares de ADNc plasmídico de las concentraciones más bajas a más altas en los pocillos control. Se fijó firmemente una lámina de sellado a la placa y se centrifugó. Se corrió un programa de PCR en un sistema de PCR tiempo real LIGHTCICLOR® 480 (Roche Diagnostics, Indianapolis, IN, USA) usando las siguientes condiciones: 1) activar 95°C durante 5 minutos; 2) desnaturalizar 95°C durante 10 segundos (@ 4,8°C/segundos); 3) aparear/extender 60°C durante 25 segundos (@ 2,5°C/segundos); 4) adquirir 72°C durante 1 segundos (@ 4,8°C/segundos); 5) repetir las etapas 2-4, 39-49 veces más. Finalmente, la reacción se enfrió a 38 °C durante 5 segundos para detener la reacción.

El tamaño del amplicón *VP16* fue 68 bp, este tamaño de fragmento corresponde al fragmento *VP16* que se esperaba fuera producido por la amplificación de PCR.

11.4 Análisis de expresión de las muestra de callo

Las líneas de callos transgénicos que crecen en la selección de HERBIACE® se analizaron para determinar los niveles de expresión de *FatB4*, *FatB*, *tubulina* y *ZFP TF* mediante qRT-PCR. El gen de tubulina sirvió como un gen de referencia para normalizar la expresión de los niveles de ARNm de *FatB4*, *FatB* y *ZFP TF*. Se calcularon las relaciones *FatB4/tubulina* y *FatB/tubulina* para normalizar las expresiones de ARNm. Los resultados mostraron una regulación por aumento estadísticamente significativa de ARNm de *FatB4* y ARNm de *FatB* para los diseños de constructo pDAB4691 ($p = 0,005$; Fig. 9). Las líneas de callos pDAB4692 mostraron una tendencia a la regulación por aumento en el ARNm de *FatB4* y *FatB*, pero la tendencia no fue estadísticamente significativa. Las líneas de callos pDAB4689 también se analizaron en un experimento separado, que mostró una tendencia de regulación por aumento, pero no fue estadísticamente significativa (datos no mostrados). Las líneas celulares de control no amplificaron un amplicón específico de *ZFP TF*, lo que confirma que el ensayo de *ZFP TF* era específico para las líneas que expresaban solo *ZFP TF*.

Ejemplo 12: Análisis de expresión de las muestras de plantas T_0 transgénicas

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

El ARN total se extrajo y se sintetizó el ADNc de todas las líneas como se describe en el Ejemplo 11. Se preparó una mezcla de PCR de la siguiente manera para la amplificación de ADNc de *actina* (Bo Yang et al., 2007, Plant Science 173:156-171; gen de actina Genbank acceso n°. AF111812): 7,5 µL de tampón 2X LIGHTCICLOR® 480 SYBR Green I Master, 0,3 µL de cebador BN_Actin_F del patrón de 10 µM (SEQ ID NO:61:ACGAGCTACCTGACGGACAAG), 0,3 µL de BN_Actin_R del patrón de 10 µM (SEQ ID NO:62:GAGCGACGGCTGGAAGAGTA), 1,5 µL de PVP-40 10% (p/v), y q/s con H₂O a 13,5 µL. Los cuadrantes apropiados de una microplaca de 384 pocillos se demarcaron y se cargaron con 13,5 µL de mezcla maestra por pocillo. Una lámina de sellado se unió suavemente a la placa. La placa se centrifugó durante 1 minuto a 3000 rpm en una centrífuga de microplacas Qiagen. Se añadieron primeras cadenas de ADNc sintetizado diluido y descongelado. a 1,5 µL, seguido por la adición de 1,5 µL estándares de ADNc de las concentraciones bajas a altas en los pocillos control. Se fijó firmemente una lámina de sellado a la placa y se centrifugó. Se corrió un programa de PCR en un sistema de PCR tiempo real LIGHTCICLOR® 480 (Roche Diagnostics, Indianapolis, IN, USA) usando las siguientes condiciones: 1) activar 95°C durante 10 minutos; 2) desnaturalizar 95°C durante 10 segundos (@ 4,8°C/segundos); 3) aparear/extender 60°C durante 20 segundos (@ 2,5°C/segundos); 4) adquirir 72°C durante 20 segundos (@ 4,8°C/segundos); 5) repetir las etapas 2-4, 39-49 más veces. Finalmente, la reacción se enfrió a 38°C durante 5 segundos para detener la reacción. El amplicón producido fue de 80 bp de tamaño.

12.2 Análisis de expresión de ARNm de *FatB4* y *FatB*

Las plantas de *B. napus* que se habían transformado con pDAB4689-pDAB4691 se analizaron para determinar la expresión de ARNm aumentada de *FatB4* y *FatB*. Además, se usaron dos tipos de controles. Un control no transgénico que consiste en plantas Nex710 sirvió como control negativo. También se analizó un segundo control transgénico de plantas de *B. napus* Nex710 que se habían transformado con pDAB8210. El diseño del constructo pDAB8210 consistió en dos casetes de expresión génica. El casete de expresión del gen *pat* (Promotor AtUbi10 ::gen *ZFN* ::AtuORF23 3' UTR). No se esperaba que el ZFN expresado en pDAB8210 se uniera y escindiera el genoma de las plantas de *B. napus* Nex710 y no debería alterar el fenotipo de estas plantas.

Los callos transgénicos putativos que crecen en la selección de HERBIACE® se regeneraron en plantas (Ejemplo 3.3). Se recogieron muestras de hojas de las plantas de 6 hojas *B. napus* que crecen en el invernadero para el análisis de expresión de ARNm. Seis perforaciones de hoja de cada planta se aislaron y se colocaron en hielo antes de la extracción de ARN. El ARN total se extrajo utilizando el kit QIAGEN RNEASY. La síntesis de ADNc y las diluciones posteriores se completaron como se describe en el Ejemplo 11. El protocolo de análisis de expresión para *FatB4* y *FatB* mediante qRT-PCR se describió anteriormente. El análisis de expresión de *ZFP TF* fue más/menos con el ensayo. El ensayo qRT-PCR del gen de referencia de *actina* se describe anteriormente.

La expresión de la hoja T_0 para los genes *FatB4* y *FatB* varió entre los constructos (Fig. 10). La regulación por aumento más alta de ARNm de *FatB4* y *FatB* se observó en eventos de plantas transgénicas pDAB4691 que produjo un aumento general de 2,0 - 2,5 veces en expresiones de ARNm en comparación con controles no transgénicos Nex710 y transgénicos pDAB8210. La expresión de ambos genes fue estadísticamente significativa (por debajo de p

= 0.05) en comparación con los controles. Por lo tanto, se continuó la caracterización adicional de los constructos con pDAB4691.

Ejemplo 13: Análisis de las plantas T₁ transgénicas

13.1 Análisis de ácido graso de semilla T₁

- 5 Se observó una diferencia significativa en el contenido de C18:0 entre la semilla de contenido de aceite "bajo" y "alto" del evento pDAB4691-003-049.1 ("evento 49") (Tabla 12). El contenido de C18:0 de la semilla de la categoría "Bajo" fue similar al de las líneas Nex710 (control no transgénico) y pDAB8210 (control transgénico). El contenido de C18:0 en la semilla "Alto" se correlacionó con la expresión de *ZFP TF* que dio como resultado la regulación por aumento de los genes de *FatB*. Algunos aumentos en C20:0, C22:0 y C24:0 también se observaron sobre la base del flujo de C18:0 en ácidos grasos de cadena más larga. El contenido de C16:0 en la semilla individual no pareció cambiar dentro de ninguna de las categorías "Bajo" o "Alto". Este resultado es indicativo del hecho de que pDAB4691 *ZFP TF* se une y regula por aumento los genes *FatB* específicos para reacciones enzimáticas C18:0-ACP a C18:0 en lugar de las reacciones C16:0-ACP a C16:0 (véase Ejemplo 1, Fig. 1).

Tabla 12: Perfil de ácidos grasos de las semillas T₁ individuales con análisis FAME

Muestras*	Sat totales	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
Evento 49 Media (n=24)	6,90	3,70	0,40	2,03	0,67	0,31	0,15
Evento 49 Desv. Est (n=24)	0,51	0,14	0,05	0,37	0,10	0,04	0,03
Nex710 Media (n=24)	6,47	3,70	0,45	1,79	0,57	0,26	0,11
Nex710 Desv. Est (n=24)	0,35	0,24	0,06	0,17	0,06	0,04	0,03
8210 Media (n=150)	6,27	3,51	0,34	1,63	0,61	0,33	0,14
8210 Desv. Est (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 individuales T ₁ del evento 49 y Nex710, y 144 semillas de pDAB8210 compuesto por 6 eventos. Las categorías "Bajo" y "Alto" representan el contenido de FA específico de la semilla individual cuando se clasifican según el contenido C18:0.							

15

13.2 Análisis de presencia de *ZFP TF* en las plantas T₁

Se plantaron 100 semillas T₁ del evento 49 en el invernadero. Noventa y siete plantas germinaron en plántulas. Se estimó el número de copias de *ZFP TF* usando un ensayo qRT-PCR de *pat* como se describe en el Ejemplo 6.2 para material vegetal aislado de la etapa de 2-4 hojas. El evento 49 contenía múltiples inserciones (~ 3 inserciones) y 0-7

gen *pat*, y por lo tanto ZFP TF, las copias se segregaron en la población T₁. Solo se identificó una planta nula de la población T₁.

13.3 Regulación por aumento de ARNM de *FatB4* y *FatB* en las plantas T₁

5 El análisis de expresión del gen de *FatB* nativo se realizó en las 97 plantas de la población T₁ segregante. Las plantas Nex710 no transgénicas que se plantaron al mismo tiempo se incluyeron en el estudio como controles. Se recogieron perforaciones de seis hojas por planta y se colocaron en hielo hasta que se pudo completar la extracción de ARN. El ARN total se extrajo utilizando el kit QIAGEN RNEASY. La síntesis de ADNc y las diluciones posteriores se completaron como se describe en el Ejemplo 11. El análisis de expresión de *FatB4*, *FatB*, *tubulina* y *VP16* (ZFP TF) se realizó como se describió anteriormente.

10 El análisis estadístico de las relaciones de *FatB4/FatB* y *tubulina* mostró tendencias lineales significativas con la expresión de *tubulina* (Fig. 11) lo que indica que los aumentos en *FatB4/FatB* no tenían una relación 1:1 con aumentos en el control endógeno, *tubulina*. Como tal, las relaciones no se usaron en este análisis, y la *tubulina* se incluyó como una covariable en los modelos de expresión de *FatB4/FatB*. Para eliminar cualquier colinealidad entre *tubulina* y *VP16* (ZFP TF), las matrices de incidencia para las ecuaciones del modelo se ortogonalizaron.

15 El siguiente modelo se ajustó a todos los conjuntos de datos en el estudio actual: donde el estado_i es el estado transgénico de y_{ijk} (transgénico o nulo/control);

$$y_{ijk} = estado_i + VP16\ transg\acute{e}nico^* + e_{ijk}$$

20 VP16 transgénico* es una regresión lineal en VP16 anidado dentro de líneas transgénicas; y e_{ijk} es el residuo aleatorio. Debido que hubo tres mediciones repetidas para cada muestra, se utilizó una estructura residual correlacionada:

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

Donde I_n es la matriz de identidad con rango igual al número de muestras y φ es la correlación entre los residuos de las mediciones repetidas.

25 Los resultados del análisis para el evento 49 detectaron una regulación por aumento significativa en *FatB4* con regresiones anidadas de la expresión de *FatB4* en *VP16* que muestran pendientes positivas altamente significativas (Fig. 12, Tabla 13). Para *FatB*, las regresiones anidadas mostraron pendientes positivas para la regulación por aumento, y la pendiente fue estadísticamente significativa. Para los eventos de regulación por aumento se debe tener en cuenta que las plantas transformadas y de control no provenían de la misma línea, por lo que podría haber
30 cierta confusión entre los efectos de la línea y la transformación.

Tabla 13: Resultados para los eventos de regulación por aumento

Evento	Gen	Efecto ^A	Solución	Valor p
4691-3-049,001	<i>FatB4</i>	Transformado-control	10331,11	<0,00001
		VP16 transformado*	0,764E-05	<0,00001
	<i>FatB</i>	Transformado-control	5084,09	<0,00001
		VP16 transformado*	0,308E-05	<0,00001

^ATransformado-control = la diferencia en la expresión de la media de mínimos cuadrados entre las líneas transformadas y las líneas que no portan transgenes. VP16 Transformado* = regresión de la expresión de *FatB4/FatB* en la expresión de VP16 anidada en líneas transgénicas.

Las plantas T₁ del Evento 49 se clasificaron luego sobre la base de las relaciones de expresión de *FatB4/tubulina* y luego de *FatB/tubulina* para encontrar las plantas de mayor expresión para el avance a la generación de T₂ en el invernadero. Ocho de las plantas de *FatB4* de mayor expresión se seleccionaron para el avance hasta la madurez. Seis de estas plantas también fueron las de mayor expresión para *FatB*. El número de copias de ZFP TF en estas plantas de mayor expresión varió de 2 a 4 copias, mientras que el número de copias que se segregó en la progenie de T₁ entera fue de 0 a 7 copias. Para las plantas de control, una planta nula de las plantas del evento 49 y diez plantas control Nex710 también avanzaron a la madurez para la recolección de semillas T₂.

13.4 Análisis de ácidos grasos de las semillas T₂

El análisis de ácidos grasos (FA) se realizó en un granel de 24 semillas de todas las plantas como se describe en el Ejemplo 6.1. El perfil de FA de una planta nula se combinó con diez plantas de control Nex710 para los cálculos de perfil de FA control. En 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 fue estadísticamente significativo a p = 0.05 (Fig. 13). Los FA de cadena larga, tal como C20:0, C22:0 y C24:0, también mostraron algunos aumentos que llevan al aumento en el contenido de FA saturados totales en un 5%. Debido a que la enzima *FatB* cataliza la conversión de FA libre C18:0 ACP en C18:0, se acumulará más C18:0 como resultado de la regulación por aumento transcripcional de *FatB4* y *FatB* (Fig. 1). El pDAB4691 (Ejemplo 6, Tablas 1 y 2), es específico tanto para los genes *FatB4* como para *FatB*, lo que dio como resultado un cambio significativo en el perfil FA (Tablas 14A y 14B).

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

S.N.	Planta ID	Sat totales	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 de la columna 2 representan números de plantas analizadas.

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

S.N.	Planta ID	Sat totales	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 de la columna 2 representan números de plantas analizadas.

No se observaron cambios significativos en el contenido total de C16 con la regulación por aumento génica de *FatB4* y *FatB* nativos. Es probable que la catálisis de C16:0-ACP a C16:0 ocurra con otros genes *FatB* pero no con *FatB4* y *FatB*, los blancos ejemplificados para la regulación por aumento usando el diseño pDAB4691 ZFP TF (Fig. 1).

El contenido de C16:0 de la semilla individual no pareció cambiar en las categorías "Bajo" y "Alto" de semilla T₁ (Tabla 12), lo que indica que el diseño pDAB4691 ZFP TF se une más específicamente al gen *FatB* para la reacción enzimática C18:0-ACP a C18:0 en lugar de la reacción C16:0-ACP a C16:0 (Fig. 1).

Se debe observar que la regulación por aumento de los genes de β -cetoacil-ACP sintetasa II nativa (Ejemplos 1-6) inducidos por ZFP TF produjo diferentes cambios en el perfil de FA en comparación con las regulaciones por aumento del gen de *FatB4* y *FatB* nativo. Por ejemplo, la regulación por aumento del gen de β -cetoacil-ACP sintetasa II provocó una reducción estadísticamente significativa en los contenidos de C16:0 y C16:1 y un aumento concomitante en el contenido total de C18 en comparación con sus segregantes nulos (Ejemplo 6, Fig. 6, Tablas 5 y 8). Comparativamente, la regulación por aumento de los genes de *FatB* mediada por ZFP-TF provocó aumentos estadísticamente significativos en el contenido de C18:0 pero ningún cambio aparente en el contenido de C16:0 (Fig. 14). De nuevo, estos cambios en los perfiles FA a través de la regulación por aumento mediada por ZFP TF de los genes de β -cetoacil-ACP sintetasa II y *FatB* coinciden según la vía de biosíntesis de FA (Fig. 1).

Ejemplo 14: Diseños de constructos ZFP TF para la transformación de plantas

Se diseñaron y construyeron cuatro constructos para la regulación por disminución transcripcional de los genes *FatB* en *B. napus* L (Tabla 15). Los mejores diseños de ZFP de regulación por aumento de *FatB*, 13722 y 13714, se emplearon para la demostración de la regulación por disminución regulación por disminución de *FatB* (Ejemplos 9 - 13). Estos diseños de ZFP se fusionaron a una señal de localización nuclear *opaque-2* y un dominio de represión que consiste en KRAB1 (Hanna-Rose and Hansen, 1996, Trends in Genetics, 12:229-234) o NtERF3 (Ohta et al., The Plant Cell, 2001, 13:1959-1968), para construir ZFP TF funcionales. Todas las ZFP TF se expresaron bajo promotores específicos de semilla; se usaron el promotor de proteína de transferencia lipídica *Arabidopsis thaliana* (promotor AtLTP170) (Genbank ID:NC_003076) o el promotor de β -Faseolina de *Phaseolus vulgaris* (promotor PvPhas) (Patente U.S:Nº 5.591.605).

Tabla 15: Detalles del constructo de ZFP TF para la regulación por disminución dirigida de los genes de *FatB*

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

* Op-2 = Señala de localización nuclear *Opaque-2*

pDAS5203 es un plásmido binario que contenía los casetes de expresión génicos NLS *opaque-2* /13722/KRAB1 y *pat*. Este constructo incluye los siguientes elementos génicos; Promotor 170 AtLTP ::Fusión NLS *opaque-2* /13722 dedos de cincLS/KRAB1 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 ::AtuORF1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la fusión NLS *opaque-2*/13722 dedos de cincLS/KRAB1 en un vector donante Gateway. El constructo resultante contenía el casete de expresión génica Promotor 170 AtLTP ::Fusión NLS *opaque-2*/13722 dedos de cincLS/KRAB1 ::AtuORF23 3'UTR v1. Este constructo se clonó en un plásmido binario por medio de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAS5203 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5204 es un plásmido binario que contenía los casetes de expresión génicos NLS *opaque-2*/13714/KRAB1 y *pat*. Este constructo incluye los siguientes elementos génicos; Promotor 170 AtLTP ::Fusión NLS *opaque-2*/13714 Dedos de cinc/KRAB1 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 ::AtuORF1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la Fusión NLS *opaque-2*/13714 Dedos de cinc/KRAB1 en un vector donante Gateway. El constructo resultante contenía el casete de expresión génica Promotor 170 AtLTP ::Fusión NLS *opaque-2*/13714 Dedos de cinc/KRAB1 ::AtuORF23 3'UTR v1. Este constructo se clonó en un plásmido binario por medio de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAS5204 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5212 es un plásmido binario que contiene los casetes de expresión génica NLS *opaque-2*/13722/KRAB1 y *pat*. Este constructo incluye los siguientes elementos génicos; Promotor PvPhas ::Fusión NLS *opaque-2*/13722 dedos de cincLS/KRAB1 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 ::AtuORF1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene la Fusión NLS *opaque-2*/13722 dedos de cincLS/KRAB1 en un vector donante Gateway. El constructo resultante contenía el casete de expresión génica de Promotor PvPhas ::Fusión NLS *opaque-2*/13722 dedos de cincLS/KRAB1 ::AtuORF23 3'UTR v1. Este constructo se clonó en un plásmido binario por medio de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAS5212 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

pDAS5227 es un plásmido binario que contiene los casetes de expresión génica de NLS *opaque-2*/13722/NtERF3 y *pat*. Este constructo incluye los siguientes elementos génicos; RB7 MAR v3 ::Promotor 170 AtLTP ::Fusión NLS *opaque-2*/13722 dedos de cincLS/NtERF3 ::AtuORF23 3'UTR v1 ::Promotor v2 AtUbi10 ::*pat* v5 ::AtuORF1 3'UTR v4. El plásmido binario se construyó mediante la clonación de un fragmento de ADN que contiene el casete de expresión génica NLS *opaque-2*/13722 dedos de cincLS/NtERF3 en un vector donante Gateway. El constructo

resultante contenía el casete de expresión génica Promotor 170 AtLTP ::Fusión NLS *opaque-2*/13722 dedos de cinc LS/NtERF3::AtuORF23 3'UTR v1. Este constructo se clonó en un plásmido binario por medio de una reacción L-R Gateway (Invitrogen, Carlsbad, CA). Esta reacción produjo pDAS5227 y se confirmó mediante digestiones con enzimas de restricción y reacciones de secuenciación.

- 5 Los constructos 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 que crecen en el invernadero se autopolinizaron y las semillas T₁ se recolectaron cuatro meses después del trasplante.

Ejemplo 15: Análisis de semilla T₁ transgénica para 4 constructos de ZFP TF

- 10 Las semillas T₁ se obtuvieron a partir de eventos transgénicos (líneas) 8, 20, 37 y 15 de los constructos pDAS5203, pDAS5204, pDAS5212 y pDAS5227, respectivamente. Se usaron cinco semillas de plantas Nex710 como controles. El análisis de FA se realizó en 24 semillas individuales de cada uno de los eventos transgénicos de T₁, como se describió anteriormente en el Ejemplo 6.1, para identificar primero el constructo ZFP TF que altera más efectivamente el perfil de FA.

15.1 Análisis de ácidos grasos (FA) de las semillas T₁

- 15 Los constructos pDAS5203, pDAS5212 y pDAS5227 produjeron eventos transgénicos con los perfiles de FA saturados totales más bajos en comparación con el Nex710 control (Tabla 16). Las diferencias de pares entre cada uno de los dos eventos del constructo, pDAS5203 y pDAS5212, y el Nex710 control fueron significativas ($p < 0.005$). Sin embargo, la diferencia entre pares de pDAS5227 exhibió una fuerte tendencia ($p = 0.06$). Uno de los eventos de pDAS5203 contenía el contenido de FA saturada total más bajo de 5.3% (no mostrado) en comparación con un promedio de 6.38% en el Nex710 control. Esta es una reducción del 17% en los FA saturados totales.
- 20

Tabla 16A: Perfil de FA de semillas T₁ de múltiples eventos transgénicos del constructo

Constructo No.	Tipo de análisis	Sat totales	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	Desv. Est (n=37)	0,21	0,01	0,20	0,02	0,14	1,36	1,03	0,41
pDAS5212	valor 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	Desv. Est (n=20)	0,29	0,01	0,15	0,02	0,16	0,78	0,66	0,30
pDAS5204	valor 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	Desv. Est (n=8)	0,36	0,01	0,25	0,05	0,13	1,76	1,31	0,47
pDAS5203	valor p	0,02	0,06	0,45	0,96	0,0001	0,96	0,98	0,49

ES 2 646 138 T3

Constructo No.	Tipo de análisis	Sat totales	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
pDAS5227	Promedio (n=15)	6,09	0,03	3,63	0,32	1,32	75,43	12,32	3,73
pDAS5227	Desv. Est (n=15)	0,35	0,00	0,29	0,09	0,12	1,69	1,36	0,37
pDAS5227	valor 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	Desv. Est (n=5)	0,61	0,01	0,15	0,03	0,49	3,00	2,76	0,43

* Los números entre paréntesis muestran el número de eventos incluidos en el análisis.
 ** valor de p igual o menor de 0,05 se consideró estadísticamente significativo y se calculó con el software JMP.

Tabla 16B: Perfil de FA de semillas T₁ de múltiples eventos transgénicos del constructo

Constructo No.	Tipo de análisis	Sat totales	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	Desv. Est (n=37)	0,21	0,04	0,11	0,01	0,03	0,05	0,04
pDAS5212	valor 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	Desv. Est (n=20)	0,29	0,06	0,17	0,01	0,05	0,05	0,03
pDAS5204	valor 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	Desv. Est (n=8)	0,36	0,03	0,15	0,01	0,02	0,01	0,05
pDAS5203	valor p	0,02	0,006	0,048	0,42	0,56	0,55	0,84

Constructo No.	Tipo de análisis	Sat totales	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
pDAS5227	Promedio (n=15)	6,09	0,56	1,39	0,06	0,39	0,15	0,16
pDAS5227	Desv. Est (n=15)	0,35	0,07	0,23	0,01	0,10	0,08	0,05
pDAS5227	valor 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	Desv. Est (n=5)	0,61	0,13	0,09	0,01	0,07	0,05	0,08
* Los números entre paréntesis muestran el número de eventos incluidos en el análisis. ** valor de p igual o menor de 0,05 se consideró estadísticamente significativo y se calculó con el software JMP.								

5 Los perfiles FA específicos de los eventos pDAS5203, pDAS5212 y pDAS5227 dieron como resultado una reducción del 21-25% en C18:0 en comparación con Nex710 (Fig. 15A). Todas estas diferencias fueron estadísticamente significativas en o por debajo de $p = 0,001$ (Tabla 16). Otros FA 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 da como resultado un aumento significativo en el contenido de C16:0/C18:0 en los tres eventos transgénicos del constructo (Fig. 165B).

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

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

15.2 Identificación de plantas para el análisis de ARNm

20 Dos eventos transgénicos, 5212 [1]-004 y 5227 [4]-012, que se produjeron por transformación con los constructos pDAS5212 y pDAS5227, respectivamente, se seleccionaron para el análisis de regulación por disminución transcripcional de *FatB*. Cincuenta a cien semillas T_1 se sembraron en el invernadero y las plantas se cultivaron. Se recogieron cuatro perforaciones de hoja de las plantas en la etapa de 2-3 hojas. Este material vegetal se analizó para determinar el número de copias de ZFP TF usando un ensayo qRT-PCR de *pat* (Ejemplo 6.2). Luego seleccionaron plantas aleatorias de ZFP nula y ZFP positiva para el avance hasta la madurez para recoger la semilla de T_2 (Tabla 17). Las vainas inmaduras de las mismas plantas se recogieron 25 días después de la floración (DAF) para el análisis de ARNm de *FatB*. Ambos eventos se segregaron como una sola copia. Estos eventos se etiquetaron como 5212-4 y 5227-12.

Tabla 17: Pruebas del evento transgénico para la identificación de plantas ZFP positivas y nulas.

Nombre evento	del Semilla/planta T_1	Planta Nº	Copia Nº	#Plantas a T_2
5212-004	Semilla plantada	50		
	Semilla germinada	49		

Nombre evento	del Semilla/planta T ₁	Planta N°	Copia N°	#Plantas a T ₂
	Pruebas para el número de copias de <i>pat</i>	17	Nula	5
		26	1 copia	3
		5	2 copia	4
		1	NT*	
	Prueba Chi	0,04		
5227-12	Semilla plantada	100		
	Semilla germinada	96		
	Pruebas para el número de copias de <i>pat</i>	19	Nula	4
		32	1 copia	3
		27	2 copia	4
		18	NT*	
	Prueba Chi	0,034		
*NT = No analizado				

15.3 Extracción de ARN y desarrollo del ensayo qRT-PCR

5 Se aisló ARN de semilla de *Brassica* inmadura con el kit de planta RNAEASY® obtenido de Qiagen (Valencia, CA). Las semillas se colocaron en tubos cluster que contiene tampón RLT (Qiagen)/β-mercaptoetanol (BME) y un microesfera de acero inoxidable. Las muestras se colocaron en una gradilla de tubos cluster y se homogenizaron a 500 golpes por minuto durante 5 minutos en un molino de microesferas (Kleco, Visalia, CA). La gradilla de tubos se giró 180 grados y se homogenizó durante 5 minutos adicionales. Las muestras se transfirieron luego a tubos Eppendorf de 1,5 ml y se centrifugaron durante 5 minutos a 20.000 x g y el sobrenadante se transfirió a un tubo nuevo. El ARN se aisló como se describe en el protocolo del fabricante (Qiagen, Valencia, CA). El ARN se eluyó de las columnas con 50 ul de agua libre de ARNasa y las muestras de ARN se cuantificaron con un NanoDrop 8000 por Thermo Scientific (Wilmington, DE).

15 Se trataron de diez a veinte microgramos del ARN con ADNasa libre ADN libre TURBO (número de catálogo AM1907; Applied Biosystems, Foster City, CA) ADNasa en tubos de 1,5 ml libres de ARNasa/ADNasa para eliminar el ADN genómico. Para cada muestra de ARN se preparó una reacción de cincuenta microlitros que contiene ARN, tampón 1X de ADNasa y 2 unidades de TURBO ADNasa. 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 20,000xg durante cinco minutos. Se extrajeron cuarenta microlitros (40 µl) del sobrenadante mientras se dejaba atrás la suspensión de inactivación de ADNasa. Cada muestra tratada con ADNasa se cuantificó con el Nanodrop de manera que se puedan usar cantidades equivalentes de ARN para la síntesis de ADNc.

25 Se preparó ADN complementario (ADNc) para cada muestra de ARN usando el kit de ADNc de alta capacidad (Applied Biosystems, Foster City, CA). Brevemente, se usaron 1,5 µg de ARN como molde en una reacción de cincuenta microlitros que contiene tampón de transcripción inversa 1x, 1X dNTPs, 1X oligómeros aleatorios y 125 unidades de transcriptasa inversa Multiscribe. Adicionalmente, 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, a 37 °C durante 2 horas, a 85 °C durante 5 minutos y a toda la noche a 4 °C.

Los cebadores y las sondas se diseñaron con el software Primer Express 3 de Applied Biosystems (Foster City, CA) para *FATB*, *KASII*, *ERF3*, *KRAB1* y *Tubulina*. Los cebadores se sintetizaron por Integrated DNA Technologies (Coralville, IA). Las siguientes son las secuencias para los cebadores:

SEQ ID NO:63:*FATB* fwd - 5' TCGCTGCCATAACAACCATTT 3';

5 SEQ ID NO:64:*FATB* rev - 5' CGCCTGGGTTTCCAGTCA 3';

SEQ ID NO:65:*KRAB1* fwd - 5' AAGGATGTGTTCTGGATTCA 3';

SEQ ID NO:66:*KRAB* rev - 5' CACAATCTGCTGTGCAGTATCAAG 3';

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

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

10 SEQ ID NO:69:*TUBULINA* fwd - 5' GAAGCTGAGAACAGCGATTGC 3';

SEQ NO:70:*TUBULINA* rev - 5' GTTCCTCCTCCCAACGAATG 3'.

Las sondas fueron 6FAM/MGB sintetizadas por Applied Biosystems (Foster City, CA):

SEQ ID NO:71:*FATB* sonda 6FAM TTTCTCAGCCGCCA;

SEQ ID NO:72:*KRAB1* sonda 6FAM TAGGGAAGAGTGAAGCT;

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

SEQ ID NO:74:*TUBULINA* sonda 6FAM TACAAGGTTTCCAAGTTT.

La expresión génica de *FatB*, *ERF3*, *KRAB1* y *tubulina* se evaluó mediante PCR en tiempo real con Applied Biosystem 7900HT (Foster City, CA). Se preparó una curva estándar usando alícuotas de varias muestras de ADNc dentro de los diferentes grupos para establecer el rango dinámico y la eficacia del ensayo. Se realizaron diluciones seriadas 1:5 y 1:4 posteriores 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 contiene mezcla maestra de Expresión Génica 1X (Applied Biosystems; Foster City, CA), 0,8 μ M de cebadores directos e inversos, 0,2 μ M de sonda 6FAM/MGB específica del gen y 4 microlitros de ADNc diluido. Todas las reacciones se realizaron bajo las siguientes condiciones en Applied Biosystem 7900HT con software SDS versión 2.4: 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. Los etapas 3 y 4 se repitieron durante 39 ciclos adicionales. La expresión del gen de *tubulina* se evaluó en cada placa, mientras que *FatB*, *ERF3* y *KRAB1* se corrieron por separado. Todas las reacciones se llevaron a cabo por triplicado.

30 Las curvas estándar se evaluaron para la regresión y los valores de pendiente y umbral se ajustaron según sea necesario. Las muestras noRT (menos controles de transcriptasa inversa que contenían ARN con todos los componentes de reacción de ADNc, excepto la transcriptasa inversa) se compararon con las muestras correspondientes para garantizar que existía una diferencia de Ct de 5 a 7. Los datos se importaron a Excel para su análisis. El valor de la cantidad promedio de *FatB*, *ERF3* y *KRAB1* se normalizó a la cantidad promedio de *tubulina* y se informó como relaciones de expresión de *FatB* y *tubulina*.

35 15.4 Análisis de regulación por disminución de ARNm de *FatB* específico de semilla en semillas inmaduras

Los resultados mostraron que la regulación por disminución del ARNm de *FatB* 25 días después de la floración (DAF) en la semilla inmadura dependía del estado de cigosidad de las plantas T_1 . Por ejemplo, no se observó regulación por disminución aparente del ARNm de *FatB* en las plantas heterocigotas (1 copia) de semilla inmadura 5227-12 (Fig. 17). Sin embargo, cuando la expresión de ZFP TF aumentó a 2 veces en las plantas homocigotas (2 copias), se observó una reducción de 21% en la expresión de ARNm de *FatB* ($p = 0,025$). Se obtuvieron resultados similares en otro evento, 5212-4, de un constructo diferente, pDAS5212. En este evento, nuevamente, no se observó una regulación por disminución de *FatB* aparente en las plantas heterocigotas (Fig. 18). Sin embargo, se observó una reducción de 15% en los transcritos de *FatB* en las plantas homocigotas cuando aumentó la expresión de TF ZFP.

45 Las plantas nulas, heterocigotas y homocigotas de ambos eventos avanzaron a madurez y se determinó el perfil de FA en una mezcla de 24 semillas de cada planta. Este proceso de segregación de linajes nulos de los linajes que contienen ZFP TF permitió diferencias de perfil de FA que reflejaban más estrechamente la presencia de ZFP TF.

15.4 Análisis del perfil de FA en semillas maduras T_2

50 Los perfiles de FA de la semilla madura T_2 variaron de acuerdo con la cigosidad de ZFP TF de la planta original T_1 para el evento 5227-12 (Tablas 18A y 18B). Aunque las plantas T_1 heterocigotas T_1 no exhibieron ninguna

5 regulación por disminución del ARNm de *FatB* aparente en 25 semillas DAF inmaduras, su perfil FA de semilla exhibió una reducción de 8% en el contenido de C18:0. No se observó una disminución significativa en el contenido de C16:0. Además, se observaron aumentos menores en C18:1 (oleico) y posteriores FA corriente abajo, lo que da da como resultado una reducción global de 5,5% en los FA saturados totales (Tablas 18A y 18B, Fig. 19). Cuando las plantas T₁ 5227-12 fueron homocigotas para ZFP TF, el ARNm de *FatB* se reguló por disminución significativamente. El perfil FA de estas semillas T₂ fue diferente. Se observó un aumento del 11% en el contenido de C16:1 que produce una reducción del 2,7% en la FA saturada total (Tablas 18A y 18B). Sin embargo, no se observó una reducción adicional en C18:0 en comparación con las plantas heterocigotas. Estos cambios coincidieron con el perfil FA de la semilla madura T₁ (Ejemplo 15.1).

10 **Tabla 18A: Perfil de FA indicado de semillas maduras T₂ del evento 5227-12**

Planta ID	Copia #	Sat totales	C14:0	C16:0	C16:1	C18:0	C18:1 (Oleico)	C18:1 (vaccénico)	C18 total:1	C18:2	C18:3
5227-12 Media (n=4)	0	7,21	0,05	4,00	0,28	2,06	71,47	5,27	76,74	10,92	3,22
5227-12 Desv. Est (n=4)	0	0,21	0,01	0,09	0,02	0,10	0,56	0,29	0,62	0,29	0,34
5227-12 Media (n=3)	1	6,84	0,05	3,82	0,25	1,88	71,99	4,88	76,87	11,03	3,33
5227-12 Desv. Est (n=3)	1	0,33	0,01	0,33	0,03	0,04	0,50	0,19	0,32	0,05	0,31
5227-12 Media (n=4)	2	7,00	0,05	3,96	0,30	1,91	70,99	5,09	76,08	11,62	3,30
5227-12 Desv. Est (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 FA de semillas maduras T₂ del evento 5227-12

Planta ID	Número de copia	Sat totales	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
5227-12 Media (n=4)	0	7,21	0,64	1,01	0,04	0,27	0,13	0,09
5227-12 Desv. Est (n=4)	0	0,21	0,03	0,02	0,01	0,01	0,03	0,01
5227-12 Media (n=3)	1	6,84	0,62	1,07	0,04	0,28	0,12	0,09
5227-12 Desv. Est (n=3)	1	0,33	0,02	0,11	0,00	0,03	0,02	0,01
5227-12 Media (n=4)	2	7,00	0,62	1,06	0,04	0,27	0,12	0,08
5227-12 Desv. Est (n=4)	2	0,29	0,09	0,03	0,01	0,03	0,03	0,01

15 Se obtuvieron resultados similares en los perfiles de FA de la semilla madura T₂ del evento 5212-4. En semillas de plantas heterocigotas, se observó una reducción del 5,5% en el contenido de C18:0, lo que produce una reducción de FA saturados totales de 3,2% (Tabla 18, Fig. 20). También se observó un aumento menor en C18:1; de lo contrario, las concentraciones de los FA restantes mostraron reducciones menores. Las plantas homocigotas no mostraron una reducción más alta en el contenido de C18:0 en comparación con la de las plantas hemocigotas a pesar de una regulación por disminución significativa del ARNm de *FatB*.

- 5 En resumen, en estos ejemplos se ha demostrado la regulación por disminución transcripcional del gen *FatB* con expresión específica de semilla de TF ZFP. Las ZFP TF fueron efectivas para conferir regulación por disminución transcripcional del gen *FatB* blanco en semillas homocigota T₂. Sin embargo, la reducción en C18:0 en semilla homocigota T₂ no fue más que la de la semilla segregante de T₂. Estos resultados indican que las copias ZFP TF homocigotas pueden desencadenar un mecanismo de retroalimentación para ajustar el perfil FA ya que los genes *FatB* son indispensables para el crecimiento y desarrollo de la planta (Bonaventure et al., The Plant Cell, 2003 15:1020-1033). Es digno de mención que la regulación por disminución de ARNm de *FatB* mediada por ZFP TF en la semilla inmadura está inversamente relacionada con la regulación por aumento de *FatB* para cambios en el contenido de C18:0.
- 10 Estos ejemplos muestran que los efectos dirigidos de ZFP TF sobre los genes *FatB* nativos alteraron específicamente la cantidad de ARNm de *FatB*, de este mdo se producen cambios del perfil de FA específicos y heredables en *B. napus*.

Tabla 19A: Perfil de FA de las semillas T₂ del evento 5212-4 en diferentes niveles de cigosidad

Planta N°	Copia N°	%Sat totales	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3
5212-4 Media (n=5)	0	7,02	0,05	3,94	0,32	1,88	76,03	11,17	3,50
5212-4 Desv. Est (n=5)	0	0,29	0,01	0,18	0,02	0,12	1,44	1,06	0,56
5212-4 Media (n=5)	1	6,79	0,04	3,83	0,32	1,81	76,86	10,74	3,36
5212-4 Desv. Est (n=5)	1	0,19	0,00	0,07	0,01	0,17	0,26	0,31	0,11
5212-4 Media (n=5)	2	6,96	0,04	3,89	0,32	1,89	76,38	11,00	3,38
5212-4 Desv. Est (n=5)	2	0,24	0,00	0,08	0,01	0,16	0,81	0,78	0,30

15 **Tabla 19B: Perfil de FA de las semillas T₂ del evento 5212-4 en diferentes niveles de cigosidad**

Planta N°.	Copia N°	%Sat totales	C20:0	C20:1	C20:2	C22:0	C24:0	C24:1
5212-4 Media (n=5)	0	7,02	0,64	1,16	0,04	0,33	0,20	0,11
5212-4 Desv. Est (n=5)	0	0,29	0,07	0,06	0,01	0,04	0,05	0,01
5212-4 Media (n=5)	1	6,79	0,62	1,11	0,04	0,31	0,18	0,10
5212-4 Desv. Est (n=5)	1	0,19	0,05	0,03	0,00	0,03	0,02	0,00
5212-4 Media (n=5)	2	6,96	0,63	1,12	0,05	0,31	0,21	0,10
5212-4 Desv. Est (n=5)	2	0,24	0,07	0,05	0,01	0,03	0,05	0,01

Lista de secuencias

- 20 <110> DOW AGROSCIENCES LLC et al.
- <120> PROTEÍNAS DE DEDOS DE CINCO MANIPULADAS GENÉTICAMENTE DIRIGIDAS A LOS GENES DE PLANTA INVOLUCRADOS EN LA BIOSÍNTESIS DE ÁCIDOS GRASOS
- 25 <130> 8325-4006.40
- <140> PCT/US2010/002817
- <141> 2010-10-22

<150> 61/279,528
 <151> 2009-10-22

5 <160> 96
 <170> PatentIn version 3.5

10 <210> 1
 <211> 38
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 1

20 cgcggatccg aacctgcgt ttgctggctt tgatgaaa 38

<210> 2
 <211> 35
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 2

30 ctcgagctgc tactgctagt tccggtggag gagcc 35

<210> 3
 <211> 736
 <212> ADN
 <213> Brassica napus

35 <400> 3

gaattcgccc ttcgcggatc cgaacctgc gtttgctggc tttgatgaaa atcttttgac 60

gtctccaogc acgagtttac gcatcccaga cgcctcotta gagagaagag agagatcgag 120

atcgatttcc tccctcttaa acctctctct ctctcgtgaa tctcatttcc ctttgccgct 180

agattctctc ttcacccttt tctcggcctt ccttcctctc ctccattactt ttttgtcgtc 240

ttctgctctc tetctctctc agcactcttc gccatggctg gtgctgctgc gtcttctctg 300

tacgcatctc cgctatgcac ctggttcgtc gccgcctgca tgtccgtctc ccacggcggc 360

40 ggagattccc gacaagccgt cgctctcaa tctaccgggc ggagtcgtcg aagcagacaa 420

cagctcacca aatgctctgg atccggtagc agcactactt cctttgggcc ttgcaatcac 480

tacaatgctt tgtcttctct cttcggatcg aactctgttt ctctcaatcg aaaccagagg 540

aggttgactc gagctgctac tgctagtctc ggtggaggag ccatggctgt tgcgatggat 600

atggaaaagg aagccaaggt tgacaacaaa cctcctacgg agcagcgccg ggttgtgtg 660

acaggcatgg gagttgaaac atcactaggt catgaccctg acaccttcta tgagaatctc 720

ctacaagggc gaattc 736

45 <210> 4
 <211> 732
 <212> ADN

ES 2 646 138 T3

<213> Brassica napus

<400> 4

```

gaattcgccc ttcgaggatc cgaacactgc gtttgctggc tttgatgaaa atgtatgtat      60
cttttgacgt ctccacgtac gagtttacgc atccagacgc ctcgttagag agaagagaga      120
gatcgagatc gagatcgaga tcgatttcct ccctctctct ctctggaatc tcatttcccc      180
ttaccgctag attctctctt cacccttttc tcgccccttc ttctctctcat tatttttttg      240
tcgtctctctg ctctctcaca gcaactctct ctttagctat ggtgggtggt gctgcgtctt      300
cttctctgta cgcactctccg ctatgcacct ggttcgtcgc tgcttgcacg tccgtctccc      360
acggcggcgg agattcccga caagccgtct ccctcaaatc taccgggcgg agtcgtcgaa      420
gcagacgaca gtcacccaaa tgcttggtct tttctggatc cggtagcgtt caggaggctc      480
tcgtcactac ttcttttggg ccttgcaatc actacaatgc ctgtctctct ctcttcggat      540
cgaactctgt ttctctcaat cgaaccaga ggaggttgaa tctggctgct gctagttccg      600
gtggaggagc catggctggt gcgatggata tgcaaaagga agccaagggt gacaacaaac      660
cccctacgga gcagcgccgt gttgtggtga caggcatggg agttgaaaca tcactaggtc      720
5 aaggcgcaat tc      732

```

<210> 5

<211> 7

<212> PRT

10 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Péptido sintético

15 <400> 5

```

Arg Ser Asp Asn Leu Ser Val
1           5

```

<210> 6

20 <211> 7

<212> PRT

<213> Secuencia artificial

<220>

25 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 6

```

Gln Lys Ile Asn Leu Gln Val
1           5

```

30 <210> 7

<211> 7

<212> PRT

<213> Secuencia artificial

35 <220>

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 7

40 <210> 7

```

Arg Ser Asp Thr Leu Ser Glu
1           5

```

<210> 8
 <211> 7
 <212> PRT
 5 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

 10 <400> 8

Thr Arg Ser Ser Arg Ile Asn
 1 5

 <210> 9
 15 <211> 7
 <212> PRT
 <213> Secuencia artificial

 <220>
 20 <223> Descripción de la secuencia artificial: Péptido sintético

 <400> 9

Arg Ser Asp Ala Leu Ala Arg
 1 5
 25
 <210> 10
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

 <400> 10
 35
Arg Ser Asp His Leu Ser Ala
 1 5

 <210> 11
 <211> 7
 40 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 45
 <400> 11

Thr Ser Ser Ser Arg Ile Asn
 1 5
 50
 <210> 12
 <211> 7
 <212> PRT
 <213> Secuencia artificial

 <220>
 55 <223> Descripción de la secuencia artificial: Péptido sintético

 <400> 12

Arg Ser Asp Asn Leu Ala Arg
 60 1 5

 <210> 13

<211> 7
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 13

Asp Arg Ser His Leu Ala Arg

10 1 5

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

15 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

20 <400> 14

Arg Ser Asp Asn Leu Ser Glu

1 5

<210> 15
 <211> 7
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

30 <400> 15

Arg Asn Ala His Arg Thr Thr

35 1 5

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

40 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

45 <400> 16

Gln Ser Gly Asn Leu Ala Arg

1 5

<210> 17
 <211> 7
 <212> PRT
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

55 <400> 17

Arg Ser Asp His Leu Ser Glu

1 5

60 <210> 18
 <211> 7
 <212> PRT

<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Péptido sintético

5 <400> 18

Gln Lys Ala Asn Arg Thr Lys
1 5

10 <210> 19
<211> 7
<212> PRT
<213> Secuencia artificial

15 <220>
<223> Descripción de la secuencia artificial: Péptido sintético

<400> 19

Arg Ser Asp Asp Leu Thr Arg
1 5

20 <210> 20
<211> 7
<212> PRT

25 <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: Péptido sintético

30 <400> 20

Thr Ser Ala Asn Leu Ser Arg
1 5

35 <210> 21
<211> 7
<212> PRT
<213> Secuencia artificial

<220>
40 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 21

Arg Ser Asp Asp Leu Ser Lys
1 5

45 <210> 22
<211> 7
<212> PRT
<213> Secuencia artificial

50 <220>
<223> Descripción de la secuencia artificial: Péptido sintético

<400> 22

55 <210> 23
<211> 7
<212> PRT
<213> Secuencia artificial

60

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 23

5 **Asp Lys Ser Asn Arg Lys Lys**
 1 5

<210> 24
 <211> 18
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

15 <400> 24

 cgtggagacg tcaaaaga 18

20 <210> 25
 <211> 20
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

 <400> 25

30 aaggaagggc gagaaaagg 20

 <210> 26
 <211> 17
 <212> ADN
 35 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

40 <400> 26

 agatgcgtaa caggaag 17

45 <210> 27
 <211> 17
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético

50 <400> 27

 ctaccgggcg gagtcgt 17

55 <210> 28
 <211> 20
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

 <400> 28

65

ES 2 646 138 T3

ttgactcgag ctgctactgc 20

<210> 29
 <211> 20
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Cebador sintético

10 <400> 29

ttccatattc catgcaaca 20

15 <210> 30
 <211> 447
 <212> ADN
 <213> Brassica napus

20 <400> 30

actcagcact ctctgccatg gtgggtgctg ctgcgtcttc ctggttacgca tctccgctat 60

gcacctgggtt cgtcgccgcc tgcattgctc tctcccacgg cggcggagat tcccgacaag 120

ccgtcgtctt caaatctacc gggcggagtc gtcgaagcag acaacagctc accaaatgct 180

ctggatccgg tagcagcact acttcctttg ggccttgcaa tcaactacaat gccttgtctt 240

ctctcttcgg atcgaactct gtttctctca atcgaaacca gaggagggtg actcgagctg 300

ctactgctag ttccgggtga ggagccatgg ctgttgcatg ggatatggaa aaggaagcca 360

aggttgacaa caaacctcct acggagcagc gccgggttgt tgtgacaggc atgggagttg 420

aaacatcact aggtcaaggg cgaattc 447

<210> 31
 25 <211> 19
 <212> ADN
 <213> Secuencia artificial

<220>
 30 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 31

acagcgattg cctacaagg 19

35 <210> 32
 <211> 22
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 32

45 agatggttaa gatcaccaa gg 22

<210> 33
 <211> 24
 50 <212> ADN
 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Cebador sintético
 <400> 33

5 aaggaagagg aaggagtcta acag 24
 <210> 34
 <211> 18
 <212> ADN
 10 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Cebador sintético

15 <400> 34

cttctgctct ccaccgta 18

<210> 35
 <211> 24
 <212> ADN
 20 <213> Secuencia artificial

<220>
 25 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 35

30 ggtaacgga tcaggatatt cttg 24
 <210> 36
 <211> 20
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 36

40 ccatgttggc aaaggcaacc 20
 <210> 37
 <211> 29
 45 <212> ADN
 <213> Secuencia artificial

<220>
 50 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 37

acaagagtgg attgatgatc tagagaggt 29

55 <210> 38
 <211> 29
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 38

65 ctttgatgcc tatgtgacac gtaaacagt 29

<210> 39
 <211> 29
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética
 <400> 39
 10
 ccagcgtaa g caataccagc cacaacacc 29
 <210> 40
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 20
 <400> 40
 cctctctacc accgtctcac atg 23
 25
 <210> 41
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 41
 35
 gatctggccg gactgttca 20
 <210> 42
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética
 45
 <400> 42
 cgctctcag ctaccacctc aacca 25
 <210> 43
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético
 <400> 43
 60
 tcgatcttga tatgttgga ga 22
 <210> 44
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 65
 <220>

<223> Descripción de la secuencia artificial: Cebador sintético

<400> 44

5 aggtgcagaa tcatgtggtg 20

<210> 45

<211> 16

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Sonda sintética

15 <400> 45

ftgatggtga agatgt 16

<210> 46

20 <211> 718

<212> ADN

<213> Brassica napus

<400> 46

25

gaattcgccc ttcgcggtac 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 ctwcakcwc tcwkcwctct tmgyatggt gggtgstgct 300

gcgtcttctt cctgttacgc atctccgcta tgcacctggt tcgctcgygc ytgcattgtcc 360

gtctcccacg gcggcggaga ttcccgacaa gccgtckcyc tcaaactctac cgggcggagt 420

cgctcgaagca gacaacagct caccaaagtc tctggatccg gtagcagcac tacttccttt 480

gggccttgca atcaactaaa tgccttgtct tctctcttcg gatcgaactc tgtttctctc 540

aatcgaacc agaggagggt gamtcgwgt gctrotgcta gttccggtgg aggagccatg 600

gctgttgca tggatatgca aaaggaagcc aaggttgaca acaaaccctc tacggagcar 660

cgccgtgttg tggtgacagg catgggagtt gaaacatcac taggtcaagg gcgaattc 718

<210> 47

<211> 28

30 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

35

<400> 47

aacgaaagga gatcgagaga ggagagag 28

40 <210> 48

<211> 28

<212> ADN

<213> Secuencia artificial

45 <220>

<223> Descripción de la secuencia artificial: Oligonucleótido sintético

ES 2 646 138 T3

<400> 48
5 cgaaggag atcgagag gcaccgca 28
<210> 49
<211> 28
<212> ADN
<213> Secuencia artificial
10 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético
<400> 49
15 aaggagaact ttaggttg gggagact 28
<210> 50
<211> 28
20 <212> ADN
<213> Secuencia artificial
<220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético
25 <400> 50
aaggagaatt ttaggttg gggagact 28
30 <210> 51
<211> 28
<212> ADN
<213> Secuencia artificial
35 <220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético
<400> 51
40 ctccaagag attggcgtaa cacttcgt 28
<210> 52
<211> 28
<212> ADN
45 <213> Secuencia artificial
<220>
<223> Descripción de la secuencia artificial: Oligonucleótido sintético
50 <400> 52
ctccaagag attggcgtaa ccttcatt 28
<210> 53
55 <211> 21
<212> ADN
<213> Secuencia artificial
<220>
60 <223> Descripción de la secuencia artificial: Cebador sintético
<400> 53
65 cttgaacgc ttatctcct c 21
<210> 54

<211> 20
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 54

10 ttccacaaca tctccccaag 20

<210> 55
 <211> 14
 <212> ADN
 15 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Sonda sintética

20 <400> 55

ctcaggctcc accc 14

<210> 56
 25 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>
 30 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 56

ctttgaaagc tcatcttct c 21

35 <210> 57
 <211> 20
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 57

45 ttccacaaca tctccccaag 20

<210> 58
 <211> 16
 50 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Sonda sintética

55 <400> 58

aaccttcac cctcca 16

60 <210> 59
 <211> 22
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 59

5 tcgatcttga tatgttgga ga 22

<210> 60
<211> 20
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 60

15 aggtgcagaa tcatgtggtg 20

<210> 61
<211> 21
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

25 <400> 61

acgagctacc tgacggacaa g 21

30 <210> 62
<211> 20
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 62

40 gagcgacggc tgaagagta 20

<210> 63
<211> 21
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

50 <400> 63

tcgctgcat aacaaccatt t 21

55 <210> 64
<211> 18
<212> ADN
<213> Secuencia artificial

60 <220>
<223> Descripción de la secuencia artificial: Cebador sintético

<400> 64

65 cgctgggtt tccagtca 18

<210> 65

<211> 22
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 65

10 aaggatgtgt tcgtggattt ca 22

<210> 66
 <211> 24
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

20 <400> 66

cacaatctgc tgtgcagtat caag 24

<210> 67
 <211> 17
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

30 <400> 67

gcgggcgggga gttgta 17

35 <210> 68
 <211> 18
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

45 <400> 68

ccccatcggc gttacatg 18

<210> 69
 <211> 21
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

55 <400> 69

gaagctgaga acagcgattg c 21

60 <210> 70
 <211> 20
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> Descripción de la secuencia artificial: Cebador sintético

<400> 70

5 gttcctcctc ccaacgaatg 20

<210> 71
 <211> 14
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética

15 <400> 71

tttctcagcc gccca 14

<210> 72
 <211> 18
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética

25 <400> 72

tagggaagag tggaagct 18

30 <210> 73
 <211> 14
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética

40 <400> 73

caggcctcag cctt 14

<210> 74
 <211> 18
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de la secuencia artificial: Sonda sintética

50 <400> 74

tacaaggttt ccaagttt 18

55 <210> 75
 <211> 7
 <212> PRT
 <213> Secuencia artificial

60 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 75

Arg Ser Asp Asn Leu Ser Ala
 1 5

65

<210> 76
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 <400> 76
 10
Gln Ser Ala His Arg Lys Thr
 1 5
 <210> 77
 <211> 7
 15 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 20 <400> 77
Gln Ser Ser His Arg Lys Thr
 1 5
 25 <210> 78
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 30 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 <400> 78
Arg Ser Asp His Leu Ser Val
 35 1 5
 <210> 79
 <211> 7
 <212> PRT
 40 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 45 <400> 79
Gln Asn Ala His Arg Ile Glu
 1 5
 <210> 80
 <211> 7
 50 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético
 55 <400> 80
Arg Ser Asp His Leu Ser Lys
 60 1 5
 <210> 81
 <211> 7

<212> PRT
 <213> Secuencia artificial

<220>
 5 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 81

Gln Asn Ala Asn Arg Ile Thr
 1 5

10 <210> 82
 <211> 7
 <212> PRT
 <213> Secuencia artificial

15 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 82

20 **Arg Ser Asp His Leu Ser Thr**
 1 5

<210> 83
 <211> 7
 25 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

30 <400> 83

His Ser Asn Thr Arg Lys Asn
 1 5

35 <210> 84
 <211> 7
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 84

Arg Ser Asp His Leu Ser Gln
 1 5

45 <210> 85
 <211> 7
 <212> PRT
 50 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

55 <400> 85

Asn Ser Ala Ser Arg Lys Asn
 1 5

60 <210> 86
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

5 <400> 86

Gln Ser Gly His Leu Ser Arg
 1 5

<210> 87
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

15 <400> 87

Asn Ser Asp Ser Leu Thr Glu
 1 5

20 <210> 88
 <211> 7
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

30 <400> 88

Arg Arg Ala Asp Leu Ser Arg
 1 5

35 <210> 89
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: Péptido sintético

40 <400> 89

Arg Ser Asp Ser Leu Ser Ala
 1 5

45 <210> 90
 <211> 7
 <212> PRT
 <213> Secuencia artificial

50 <220>
 <223> Descripción de la secuencia artificial: Péptido sintético

<400> 90

Gln Asn Ala His Arg Lys Thr
 1 5

55 <210> 91
 <211> 7
 <212> PRT
 <213> Secuencia artificial

60 <220>

ES 2 646 138 T3

<223> Descripción de la secuencia artificial: Péptido sintético

<400> 91

5 Arg Asn Ala Asp Arg Ile Thr
1 5

<210> 92

<211> 200

<212> PRT

10 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: Polipéptido sintético

15 <220>

<221> característica_misc

<222> (1)..(200)

<223> Esta región puede abarcar 5 a 200 restos "Gly"

20 <400> 92

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
1 5 10 15

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
20 25 30

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
35 40 45

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
50 55 60

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
65 70 75 80

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
85 90 95

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
100 105 110

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
115 120 125

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
130 135 140

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
145 150 155 160

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
165 170 175

ES 2 646 138 T3

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
 180 185 190

Gly Gly Gly Gly Gly Gly Gly Gly
 195 200

<210> 93
 <211> 1693
 <212> ADN
 <213> Brassica napus

5

<400> 93

ctcgagctgc tactgctagt tccggtggag gagccatggc tgttgcgatg gatatgghaa 60
 aggaagccaa ggttgacaac aaacctccta cggagcagcg ccgggttgtt gtgacaggca 120
 tgggagttga aacatcacta ggtcatgacc ctgacacctt ttatgagaat ctccctacaag 180
 gcaacagtgg tattagccag attgagaatt ttgattgttc tgcttttcct acgagaattg 240
 ctggagagat caagtcatc tcgactgaag ggtgggttgc tccaaaactc tcaaagagga 300
 tggacaagtt catgctctat cttctcactg ctggcaagaa ggctttggtt gatggtgggg 360
 taaccgaaga agtcatggca gagtttgaca aagcctaatg cggagtcttg attggctctg 420
 caatgggagg catgaaggtc tttcaagatg ctattgaagc tatgaagatc tcttacaaga 480
 agatgaatcc tttctgtgtg cttttcgcga caacgaacat gggttctgct atgcttgctt 540
 tggatctggg atggatgggg ccaactatt ctatctcaac tgcttctgca acaagcaact 600
 tttgcattct caattcagca aaccacatta tcaaaggaga agctgatgta atgctctgtg 660
 gtggctcgga ttcagttatt attccaatag ggttgggagg ttttgttgca tgccgtgctc 720
 tttctcaaag gaataatgat cccacaaaag cttcacgccc ttgggatagc aaccgagatg 780
 gtttctgtgat gggagaggga gctggagttt tgcttttggga agagcttgaa catgctaaga 840
 aaagaggagc aacaatctat gcagagttcc ttggtgggag tttcacatgt gatgcctatc 900
 acatgaccga gcctcgccct gatggtgctg gtgtgattct gtgtattgag agagcattgg 960
 ctgatgctgg gatttccaaa gaacagataa actatataaa tgcacatgca acctctacac 1020
 cagctggaga ccttaaggag taccaagccc ttgctcactg ctttggccaa aatcctgaga 1080
 taaaagttaa ttcacaaaa tctatgattg gacacttgct gggagctgct ggggcccgtg 1140
 aagctgtcgc aactgtgcag gccataagga ccggatgggt tcatccaaat atcaaccttg 1200
 agagtccaga caatggagtg gatacaaatt tgctggtggg tcctgagaag gagagattgg 1260
 acattaaagc agccttgtca aattcattcg ggtttggtgg ccacaactcc agcatcattt 1320
 ttgctcctta caagtgaaag cactcattgc ctgtactcca aacctggttg tgtaacttgc 1380
 tgtaagtgtt tacaagaact ccccatgtta tgttgttgcg ggaatcaaca cagtttgta 1440
 aactaccaag agctaageta agtttcctta ggatcaagat ccgtttgtgc cagagaactt 1500
 ggacaaaagag caaacgtagc agagtttggga tttagcttcc gtgtgatacc ttttgagtgg 1560
 aatctttgta gccttttctt tttttagtggt tttcatttct atttgttaat cattacaatc 1620
 tgaaaattgc caaaccaatt ctccgttaaa tttagtaact ctacacaaaa aaaaaaaaaa 1680
 aaaaaaaaaa aaa 1693

10

ES 2 646 138 T3

<210> 94
 <211> 1767
 <212> ADN
 <213> Arabidopsis thaliana

5

<400> 94

```

attctctctt cttctcttca cccatttctc gctttctcct ttgttctctc atctgggttc      60
ttctcaaagc ctcttccttt ttttgccatg gtgggtgCGT cttcctctta cgcatactccg      120
ttatgtacct ggtttggtgc tgcttgcatg tccgtctctc acggtggagg agatagccgt      180
caggctgttg ctcttcaatc tgggtggcgg agtcggcgaa ggaggcagct tagcaaatgc      240
tctgtcgtt ctggatccgc tagcattcag gctctcgtca cttcttgttt ggattttggt      300
ccttgtaact actacaacaa caacaatgca ttgtcttctc tctttggatc gaatagtgtt      360
tctttgaatc gaaaccagag gagattgaat cgtgctgcta gctccggtgg agccatggca      420
gtgatggaga tggaaaagga agctgcggtt aacaagaaac cacctacgga gcagcgtcga      480
gttgtagtga caggcatggg agttgaaaca tcattgggtc atgaccaca taccttctat      540
gagaatttgc tacaaggcaa cagtggattt agccagattg aaaattttga ttgttctgaa      600
tttctactgc gaattgCGG agagatcaaa agcttctcga ctgaaggatg ggttgctcca      660
aaactttcta aaaggatgga caaattcatg ctctatcttc tcacagctgg taagaaagct      720
ttggctgatg gtgggggttac 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 caataggggt gggaggtttt     1080
gttgcattgc gggctctttc acaaggaat aatgatccca 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 cgttagctag tgctgggatt tccaaggaac aaataaatta cataaatgca     1380
catgcaacct caacgcatgc tggagatatt aaggaatacc aagcccttgc tcaactgttt     1440
ggccaaaatc ctgagcttaa ggtaaattcc acaaaaacta tgattggaca cttgctggga     1500
gctgctgggg ccgtggaggc tgttgcaact gtgcaggcga tacggaccgg atgggttcat     1560
ccaaatatca acctcgagaa tccagacagt ggagtggata caaagctgct ggtgggtcct     1620
aagaaggaga gactggacat taaagcagcc ttgtcaaatt cattcgggtt tgggtgtcat     1680
aactccagca tcatttttgc tccttacaag tgaagcgaa agcagttgct tgtactccaa     1740
acctgattgt ataacttgct gtaaggt                                     1767
    
```

<210> 95
 <211> 893
 <212> ADN

15

ES 2 646 138 T3

<213> Brassica rapa

<400> 95

gacatggata ttttctgaga tttggaagtg tggatttgat ataagcaaag caaagcaagt 60
gctcgccttg atgtatcttt tgacgtctcc acgcacgagt ttaacgatcc agacgcctcg 120
ttagagagaa gagagagatc gagatcgatt cctccctctt aaacctctct ctctctcgtg 180
aatctcattt cccttaccg ctagattctc tcttcaccct tttctcgccc ttccttcctt 240
ctcctcatta cttttttgtc gtcttctgct ctctctctct ctacgactc ttcgctttag 300
ctatggtggg tgctgctgcg tcttctgct acgcatctcc gctatgcacc tggttcgtcg 360
ctgcctgcat gtcctgtctcc 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 tgttagtaaa gtttctcttt ttatgtgttt 720
atgttatttg gtcaaatgat tcaacttggg gatcaagttt aacgggaaag tatgctgatt 780
tggaatctca gcttaattgt caagtttcat tctttattgt gtgtgtaata tgcaaagcac 840
5 atttctgtga gtttgattca atattggact ctagataagg agtatgatga tac 893

<210> 96

<211> 812

<212> ADN

10 <213> Brassica oleracea

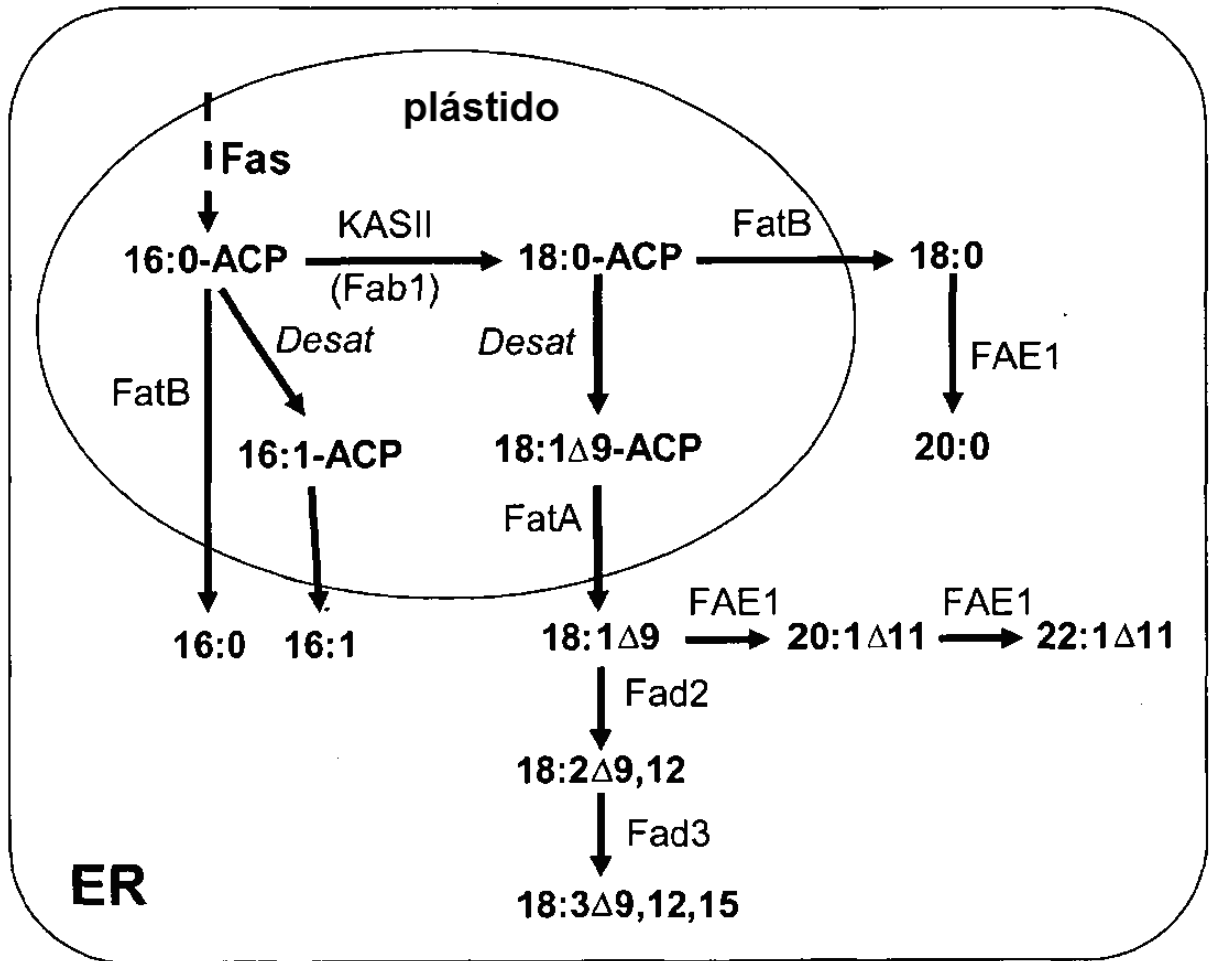
<400> 96

acggcacttt acgcaccaat atataccata tggtagccag ttaagggtac gaagtgattt 60
tccccctttc ttatgtatac agactccaat tccccaaatt tactctggtc atcgaacatc 120
gatacgtttg tatgtatgta cgtcttatct ctctctctgg ctatctctgt gtgttttggt 180
gttaataaca ggaggttttg ttttgtttc tttgacgtct ccaogcacga ttttaagcct 240
ccgttacgct ctgcacgcct cctgaaagag agagagagag agatcgaatc atcttaatta 300
15 aacctctctc gtgaatccgc tagattctct cttcacogtt ttctcgcctc tcctctcctc 360
tccattttgt tttcttctc ttcgtctctc tcttcttgcc attgccatgg tgggtgctgc 420
tgctctctcc tgttacgcat ctcccttatg caccttcttc gttgctgctt gcatgtccct 480
ctctcacggc ggcggtgata cccgtcaagc ctttggcggg agccgtcgaa ggagacaaca 540
gctcggcaaa tgctctggat ccggtagcat tctcgtgtct tcttgtttg agtttaagcc 600
ttgcagtcc tacaacaaga acaacaaagg caatgccttt cctctcctcg gatcgaatag 660
ccttctctg aatcgaaagc agaggaaact gaatcgagca acttcttctt cttccggtat 720
aaacattctt ctcccaaat ctcaagtctt ggtttctggg tatcgtttgc aagctgacaa 780
gctttgtttt ttgtgtataa agtttctcct tt 812

REIVINDICACIONES

- 5 **1.** Una proteína de fusión que comprende un dominio funcional, que es un dominio regulador transcripcional o un dominio de escisión y una proteína de dedos de cinc natural, que comprende las regiones de hélice de reconocimiento mostradas en una fila única de la Tabla 1, donde la proteína de fusión modula la expresión de al menos un gen de β -cetoacil-ACP sintetasa (KAS) endógeno.
- 2.** La proteína de fusión de la reivindicación 1, en la que el gen KAS está contenido dentro de una planta de *Brassica*.
- 3.** La proteína de fusión de la reivindicación 1 o 2, en la que la proteína de dedos de cinc se une a un sitio blanco como se muestra en la Tabla 2.
- 10 **4.** Un polinucleótido que codifica una o más proteínas de dedos de cinc o proteínas de fusión de cualquiera de las reivindicaciones 1 a 3.
- 5.** Un procedimiento de modificación de uno o más genes involucrados en la biosíntesis de ácidos grasos en una célula de planta, el procedimiento que comprende:
- 15 introducir, en la célula de planta, uno o más vector de expresión que comprende al menos un polinucleótido de acuerdo con la reivindicación 4, en condiciones tales que la una o más proteínas de dedos de cinc, que comprenden las regiones de la hélice de reconocimiento mostradas en una fila única de la Tabla 1, se expresan y se modifica la expresión de uno o más genes de KAS.
- 6.** El procedimiento de la reivindicación 5, en el que la modificación comprende la modulación de la expresión de al menos un gen KAS.
- 20 **7.** El procedimiento de la reivindicación 6, en el que se activa la expresión de al menos un gen KAS.
- 8.** El procedimiento de la reivindicación 6, en el que se reprime la expresión de al menos un gen KAS.
- 9.** El procedimiento de la reivindicación 5, en el que el polinucleótido codifica una nucleasa de dedos de cinc y al menos se escinde un gen de KAS.
- 10.** El procedimiento de la reivindicación 9, que además comprende la etapa de introducir un vector donante de modo que el vector donante se introduce en el sitio de escisión por recombinación homólogo.
- 25 **11.** Una célula de planta que comprende una o más proteínas de fusión de acuerdo con cualquiera de las reivindicaciones 1 a 3 o uno o más polinucleótidos de acuerdo con la reivindicación 4 y que comprende al menos un gen que se ha modificado de cualquiera de los procedimientos de cualquiera de las reivindicaciones 5 a 10.
- 12.** La célula de planta de la reivindicación 11, en la que se modifica el contenido de ácido graso dentro de la célula de planta.
- 30 **13.** Una planta que comprende al menos una célula de acuerdo con la reivindicación 11 o 12.
- 14.** La semilla o progenie de una planta de acuerdo con la reivindicación 13, que comprende una o más proteínas de fusión de acuerdo con cualquiera de las reivindicaciones 1 a 3 o uno o más polinucleótidos de acuerdo con la reivindicación 4.
- 35 **15.** Una planta que comprende al menos una célula que comprende la proteína de fusión de las reivindicaciones 1 a 3 o un polinucleótido de acuerdo con la reivindicación 4.
- 16.** La semilla o progenie de una planta de acuerdo con la reivindicación 15, que comprende una o más proteínas de fusión de acuerdo con cualquiera de las reivindicaciones 1 a 3 o uno o más polinucleótidos de acuerdo con la reivindicación 4.

40



Vía de biosíntesis de ácidos grasos

FIG. 1

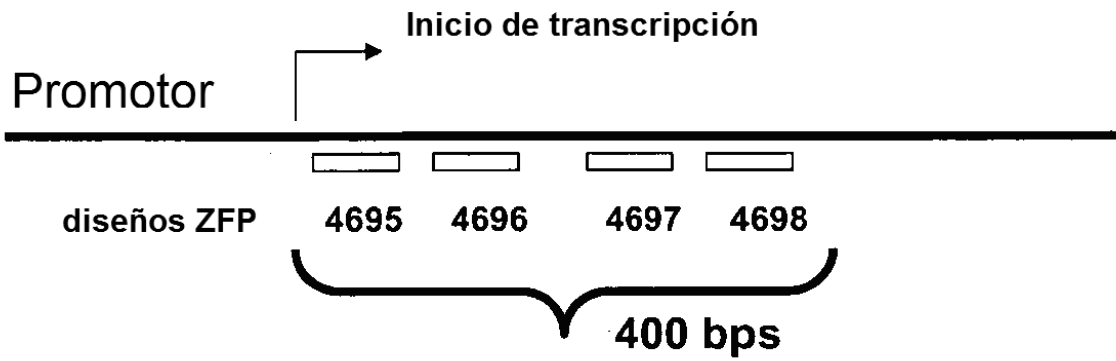


FIG. 2

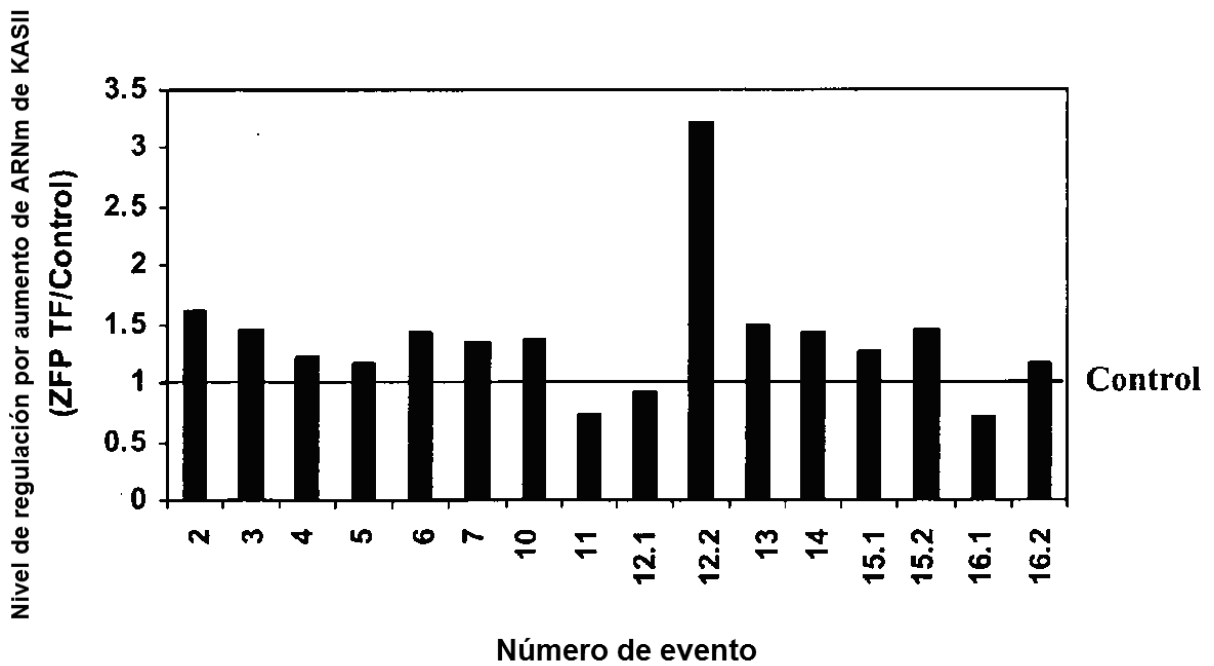


FIG. 3

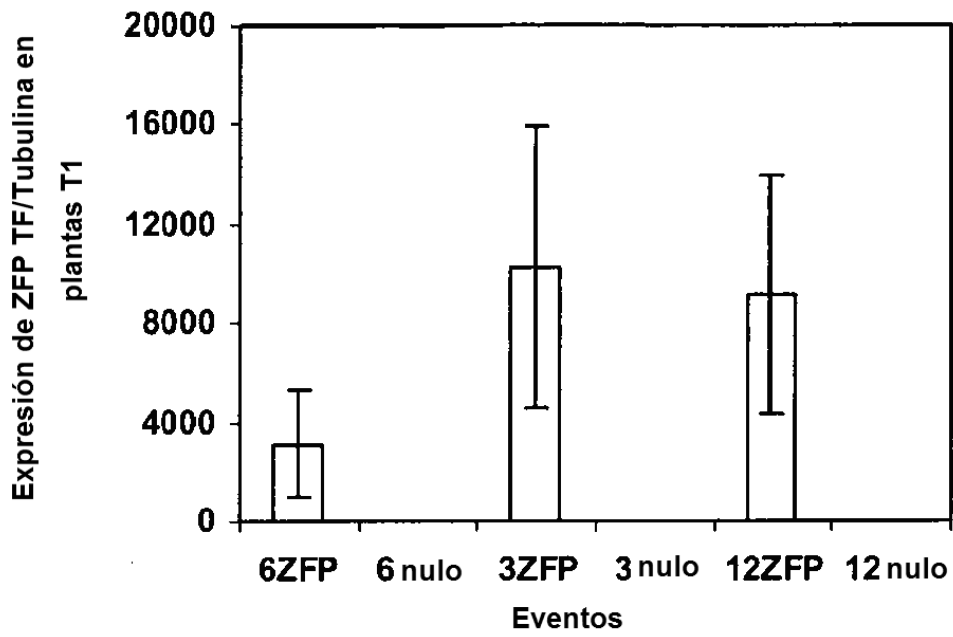


FIG. 4

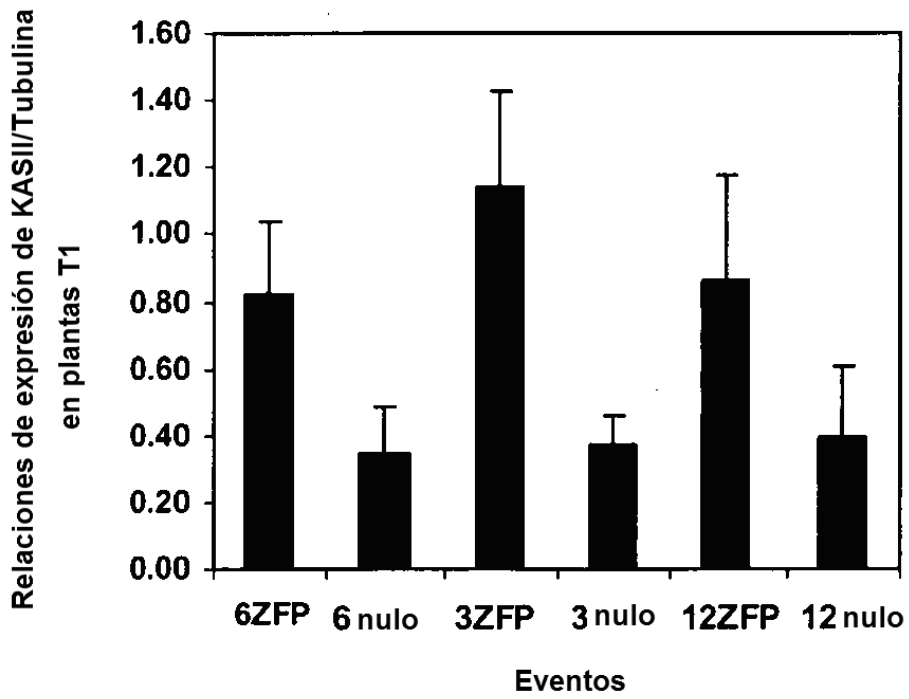


FIG. 5

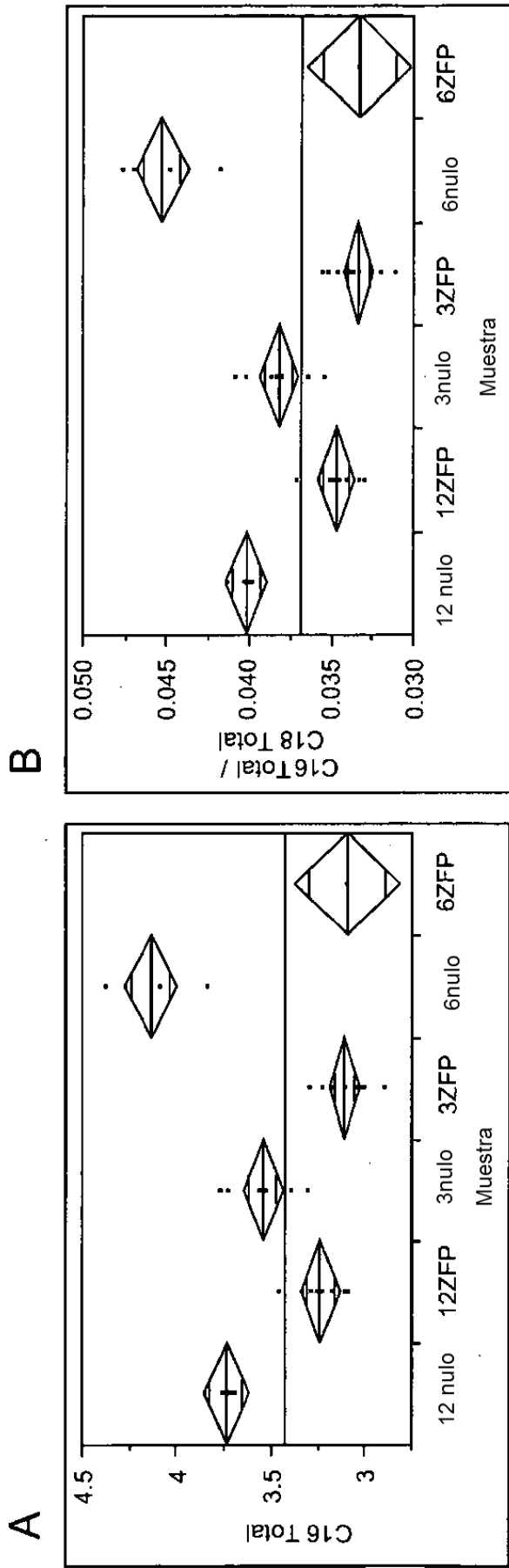


FIG. 6

ES 2 646 138 T3

```

1 50
AF244520 (1) -----
AF318307 (1) ATCTCTCTTCTTCTTCCACCCATTTCTCGCTTCTCCTTTGTTCTCTC
Consenso (1)

51 100
AF244520 (1) -----
AF318307 (51) ATCTGGGTTCTTCTCAAAGCCTCTCCCTTTTTTGCCATGGTGGGTGCGT
Consenso (51)

101 150
AF244520 (1) -----
AF318307 (101) CTTCCTCTTACGCATCTCCGTTATGTACCTGGTTTGTGCTGCTTGCATG
Consenso (101)

151 200
AF244520 (1) -----
AF318307 (151) TCCGTCTCTCACGGTGGAGGAGATAGCCGTGAGGCTGTGCTCTTCAATC
Consenso (151)

201 250
AF244520 (1) -----
AF318307 (201) TGGTGGGCGGAGTCCGCGAAGGAGGCAGCTTAGCAAATGCTCTGTCGCTT
Consenso (201)

251 300
AF244520 (1) -----
AF318307 (251) CTGGATCCGCTAGCATTGAGGCTCTCGTCACTTCTTGTGGATTTGGT
Consenso (251)

301 350
AF244520 (1) -----
AF318307 (301) CCTGTACTCACTACAACAACAATGCATTGTCTCTCTTTGGATC
Consenso (301)

351 400
AF244520 (1) -----CTTCGAGCTTCTACTGCT
AF318307 (351) GAATAGTGTCTTCTTGAATCGAAACCAGAGGAGATGATCGTCTGCT
Consenso (351) T GA G T CTGCT

401 450
AF244520 (18) AGTTCGGGTGGAGGAGCCATGGCTGTGCGATGGATATGGAAAAGGAAGC
AF318307 (400) AGCTCCGGTGGAG---GCCATGGCAGT---GATGGAGATGGAAAAGGAAGC
Consenso (401) AG TCCGGTGGAG GCCATGGC GT GATGGA ATGGAAAAGGAAGC

451 500
AF244520 (68) CAGGGTTGACAAACAACCTCTACGGAGCAGCGCCGGGTTGTTGTGACAG
AF318307 (444) TCGGGTTGACAAACAACCTCTACGGAGCAGCGCTCGAGTTGTAGTGACAG
Consenso (451) GGT ACAA AAACC CCTACGGAGCAGCG CG GTTGT GTGACAG

501 550
AF244520 (118) GCATGGGAGTTGAAACATCACTGGTCATGACCCGACACCTTTATATGAG
AF318307 (494) GCATGGGAGTTGAAACATCACTGGTCATGACCCGACACCTTTATATGAG
Consenso (501) GCATGGGAGTTGAAACATCA T GGTCATGACCC A ACCTT TATGAG

551 600
AF244520 (168) AATCTCTACAAGGCAACAGTGGTATTAGCCAGATTGAAATTTTGATTG
AF318307 (544) AATCTCTACAAGGCAACAGTGGTATTAGCCAGATTGAAATTTTGATTG
Consenso (551) AAT T CTACAAGGCAACAGTGGTATTAGCCAGATTGA AATTTTGATTG

601 700
AF244520 (218) TTCTGCTTTTCTACGGAATTGCTGGAGAGATCAAGTCAATCTCGACTG
AF318307 (594) TTCTGAAATTTCTACGGAATTGCTGGAGAGATCAAGTCAATCTCGACTG
Consenso (601) TTCTG TTTCTACG GAATTGC GGAGAGATCAA TTCTCGACTG

651 750

```

FIG. 7A

AF244520	(268)	AAGGATGGGTTGCTCCAAAAC	TC	AA	AGGATGGACAA	TT	CATGCTC	
AF318307	(644)	AAGGATGGGTTGCTCCAAAAC	TC	AA	AGGATGGACAA	TT	CATGCTC	
Consenso	(651)	AAGG	TGGGTTGCTCCAAAAC	TC	AA	AGGATGGACAA	TT	CATGCTC
			701					750
AF244520	(318)	TATCTTCTCAGCTGGCAAGAA	GCTTTGG	TGATGGTGGGGT	AA	CCGA		
AF318307	(694)	TATCTTCTCAGCTGGCAAGAA	GCTTTGG	TGATGGTGGGGT	AA	CCGA		
Consenso	(701)	TATCTTCTCAC	GCTGG	AAGAA	GCTTTGG	TGATGGTGGGGT	AC	GA
			751					800
AF244520	(368)	AGAAGTCATGGCAGAGTTGACAAA	CCAAATG	GGAGT	CT	TGATTGGCT		
AF318307	(744)	TGAAGTAATGGCAGAGTTGACAAA	CCAAATG	GGAGT	CT	TGATTGGCT		
Consenso	(751)	GAAGT	ATGGCAGAGTTGACAAA	CCAAATG	GGAGT	TTGATTGGCT		
			801					850
AF244520	(418)	CTGSCAATGGGAGGCATGAAGGTCTT	CAAGATGCTATTGAAGCT	TGAAG				
AF318307	(794)	CTGSCAATGGGAGGCATGAAGGTCTT	CAAGATGCTATTGAAGCT	TGAAG				
Consenso	(801)	C	GCAATGGGAGG	ATGAAGGTCTT	A	GATGCTATTGAAGCT	TGA	
			851					900
AF244520	(468)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	CCTTT	CGCACAAC	CGAA		
AF318307	(844)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	CCTTT	CGCACAAC	CGAA		
Consenso	(851)	ATCTCTTACAAGAAGATGAATCCTTT	TGTGT	CCTTT	CGCACAAC	AA		
			901					950
AF244520	(518)	CATGGGTTCTGCTATGCTTGC	TGGATCTGGGATGGATGGGGCCAAACT					
AF318307	(894)	CATGGGTTCTGCTATGCTTGC	TGGATCTGGGATGGATGGGGCCAAACT					
Consenso	(901)	CATGGGTTCTGCTATGCTTGC	TGGATCTGGGATGGATGGGGCCAAACT					
			951					1000
AF244520	(568)	ATTCTATCTCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	CAATTCA				
AF318307	(944)	ATTCTATCTCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	CAATTCA				
Consenso	(951)	ATTCTAT	TCAACTGCTTGTGC	ACAAGCAACTTTTGCATTCT	AATTCA			
			1001					1050
AF244520	(618)	GCAAACCACATTATCAAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC					
AF318307	(994)	GCAAACCACATTATCAAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC					
Consenso	(1001)	GCAAACCACATTAT	AAAGG	GAAGCTGATGTAATGCTCTGTGGTGGCTC				
			1051					1100
AF244520	(668)	GGATTCAGTTATTATTCCAATAGGGT	TGGGAGGTTTGTGTCATGCCG	TTG				
AF318307	(1044)	GGATTCAGTTATTATTCCAATAGGGT	TGGGAGGTTTGTGTCATGCCG	TTG				
Consenso	(1051)	GAT	CAGTTATTATTCCAATAGGGT	TGGGAGGTTTGTGTCATGCCG	G			
			1101					1150
AF244520	(718)	CTCTTTCCAAAGGAATAATGATCCCA	CAAAAGCTTCACG	CCTTGGGAT				
AF318307	(1094)	CTCTTTCCAAAGGAATAATGATCCCA	CAAAAGCTTCACG	CCTTGGGAT				
Consenso	(1101)	CTCTTTCCAAAGGAATAATGATCCCA	CAAAAGCTTCACG	CCTTGGGAT				
			1151					1200
AF244520	(768)	AGCAA	CGAGATGGTTTCGTGATGGGAGAGGAGCTGGAGTT	TTGCTTTT				
AF318307	(1144)	AGCAA	CGAGATGGTTTCGTGATGGGAGAGGAGCTGGAGTT	TTGCTTTT				
Consenso	(1151)	A	CAA	CGAGATGGTTTCGTGATGGGAGAGGAGCTGGAGTT	T	CTTTT		
			1201					1250
AF244520	(818)	GGAAGAGCTTCAACATGCTAAGAAA	AGAGG	GCAAC	ATCTA	CGAGAGT		
AF318307	(1194)	GGAAGAGCTTCAACATGCTAAGAAA	AGAGG	GCAAC	ATCTA	CGAGAGT		
Consenso	(1201)	GGAAGA	CT	GA	CATGCTAAGAAAAGAGG	GCAAC	ATCTA	CGAGAGT
			1251					1300
AF244520	(868)	TCCTTGGTGGGAGTTTACATGTGAT	GCCTATCACATGACCGAGCCTC					
AF318307	(1244)	TCCTTGGTGGGAGTTTACATGTGAT	GCCTATCACATGACCGAGCCTC					
Consenso	(1251)	TCCT	GGTGGGAGTTTACATGTGATGCCTATCACATGACCGAGCCTC	C				
			1301					1350

FIG. 7B

```

AF244520 (918) CCTGATGGGCTGGTGTATTCTGTGATTGAGAGAGCATTGGCTGATGC
AF318307 (1294) CCTGATGGGCTGGTGTATTCTGTGATTGAGAGAGCATTGGCTGATGC
Consenso (1301) CCTGATGG GCTGGTGT ATTCT TGTATTGAGAGAGC TT GCT TGC
1351 1400
AF244520 (968) TGGGATTTCCAAAGAACAATAAACTAATAAAATGCACATGCAACCTCA
AF318307 (1344) TGGGATTTCCAAAGAACAATAAACTAATAAAATGCACATGCAACCTCA
Consenso (1351) TGGGATTTCCAA GAACA ATAAA TA ATAAATGCACATGCAACCTC A
1401 1450
AF244520 (1018) CACCAGCTGGAGACCTTAAGGAGTACCAAGCCCTTGCTCACTGCTTTGGC
AF318307 (1394) CCGATGCTGGAGATTTAAGGAATACCAAGCCCTTGCTCACTGCTTTGGC
Consenso (1401) C C GCTGGAGA TTAAGGA TACCAAGCCCTTGCTCACTG TTTGGC
1451 1500
AF244520 (1068) CAAAATCCTGAGATAAAGTAAATCCACAAAATCTATGATTGGACACTT
AF318307 (1444) CAAAATCCTGAGCTTAAGTAAATCCACAAAATCTATGATTGGACACTT
Consenso (1451) CAAAATCCTGAG T AA GT AATCCACAAAATCTATGATTGGACACTT
1501 1550
AF244520 (1118) GCTGGGAGCTGCTGGGGCCGTGAGGCTGTGCAACTGTGCAGGCATAA
AF318307 (1494) GCTGGGAGCTGCTGGGGCCGTGAGGCTGTGCAACTGTGCAGGCATAA
Consenso (1501) GCTGGGAGCTGCTGGGGCCGT GA GCTGT GCAACTGTGCAGGC ATA
1551 1600
AF244520 (1168) GGACCGGATGGGTTTCATCCAAATATCAACCTGAGAGTCCAGACAATGGGA
AF318307 (1544) GGACCGGATGGGTTTCATCCAAATATCAACCTGAGAGTCCAGACAATGGGA
Consenso (1551) GGACCGGATGGGTTTCATCCAAATATCAACCT GAGA TCCAGACA TGGGA
1601 1650
AF244520 (1218) GTGGATACAAAATTGCTGGTGGGTCCAGAGGAGAGACTGGACATTAA
AF318307 (1594) GTGGATACAAAATTGCTGGTGGGTCCAGAGGAGAGACTGGACATTAA
Consenso (1601) GTGGATACAAA TGCTGGTGGGTCC AGAAGGAGAGA TGGACATTAA
1651 1700
AF244520 (1268) AGCAGCCTTGTCAAATTCATTCCGGTGGTGGTCAACTCCAGCATCA
AF318307 (1644) AGCAGCCTTGTCAAATTCATTCCGGTGGTGGTCAACTCCAGCATCA
Consenso (1651) AGCAGCCTTGTCAAATTCATTCCGGTGGTGGT CA AACTCCAGCATCA
1701 1750
AF244520 (1318) TTTTGTCTCCTTACAAGTGAAAGC -ACTCA-TTGGCTGTACTCCAAACC
AF318307 (1694) TTTTGTCTCCTTACAAGTGAAAGCGAAAGCAATTGGCTGTACTCCAAACC
Consenso (1701) TTTTGTCTCCTTACAAGTGAAAGC A CA TTGC TGTACTCCAAACC
1751 1800
AF244520 (1365) TGGTTGTGTAACCTGCTGTAAGTGTTTACAAGAACTCCCCATGTTATGTT
AF318307 (1744) TGAATTGTATAAAGTCTGTAAG-GT-----
Consenso (1751) TG TTGT TAACTTGCTGTAAG GT
1801 1850
AF244520 (1415) GTTGCGGAATCAACACAGTTTGTAAACTACCAAGAGCTAAGCTAAGTT
AF318307 (1768) -----
Consenso (1801) -----
1851 1900
AF244520 (1465) TCCTTAGGATCAAGATCCGTTTGTGCCAGAGAACTTGGACAAAGAGCAAA
AF318307 (1768) -----
Consenso (1851) -----
1901 1950
AF244520 (1515) CGTAGCAGAGTTGGATTTAGCTTCCGTGTGATACCTTTTGAGTGGAAATC
AF318307 (1768) -----
Consenso (1901) -----
1951 2000

```

FIG. 7C

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

FIG. 7D

ES 2 646 138 T3

```

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

51                               100
Seq ID 3 (1) -----
Seq ID 4 (1) -----
Seq ID 46 (1) -----
Seq ID 47 (1) -----
AC189461 (1) -----
BH723504 (51) GAAGTGATTTTCCCCTTTTCTTATGTATACAGACTCCAATTCCCCAAAATT
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) -----GACATTCGATAATT
BH723504 (101) TACTCTGGTCATCGAACATCGATACGTTTGTATGTATGTACCGTCTTATCTT
Consenso (101) -----GAATTCGCCCTT

151                               200
Seq ID 3 (13) CGCGGA--TCCGAACACTGCGTTTGGCTGGCTTTGATGAAA-----
Seq ID 4 (13) CGCGGA--TCCGAACACTGCGTTTGGCTGGCTTTGATGAAA-----
Seq ID 46 (13) CGCGGA--TCCGAACACTGCGTTTGGCTGGCTTTGATGAAA-----C
Seq ID 47 (1) -----
AC189461 (13) TTCTGAGATTTGGAAGTGTGGATTTGATATAAGCAAGCCAAAGCAAGTGC
BH723504 (151) CTCTCTC-TCGCTATGCTGCTGCTTTGTTGTTAATAACAGGAGG----
Consenso (151) CGCGGA TCCGAACACTGCGTTTGGCTGGCTTTGATGAAA

201                               250
Seq ID 3 (51) -----ATCTTTTACGCTCTCCACGCACGAGTTTACGCATCCA
Seq ID 4 (51) -----ATGTATGTATCTTTTACGCTCTCCACGTTACGAGTTTACGCATCCA-
Seq ID 46 (53) TTGTATGTATGTATCTTTTACGCTCTCCACGTTACGAGTTTACGCATCCA-
Seq ID 47 (1) -----
AC189461 (63) TCGCTTGTATGTATCTTTTACGCTCTCCACGCACGAGTTTACGCATCCA-
BH723504 (196) -TTTTGTTTGTGTTTCTTTTACGCTCTCCACGCACGAGTTTACGCATCCA
Consenso (201) -----TGATGTATCTTTTACGCTCTCCACGCACGAGTTTACGCATCCA

251                               300
Seq ID 3 (89) GACGCCTCGTTAGAGAGAAGAGAGATCGAGATCGATTTCTC-CTC
Seq ID 4 (96) GACGCCTCGTTAGAGAGAAGAGAGATCGAGATCGAGATCGAGATCGATCGAT
Seq ID 46 (102) GACGCCTCGTTAGAGAGAAGAGAGATCGAGATCGAGATCGAGATCGATCGAT
Seq ID 47 (1) -----
AC189461 (112) GACGCCTCGTTAGAGAGAAGAGAGATCGAGATCGATTTCTC-CTC
BH723504 (245) TACGCTCTGACGCTCTCTGAAAGAGAGAGAGAGATCGAA-TCATC
Consenso (251) GACGCCTCGTTAGAGAGAAGAGAGATCGAGATCGAGATCGA TC TC

301                               350

```

FIG. 8A

Seq ID 3	(137)	TTAAACCTCTCTCTCTCGTGAATCTCATTTC	CCCTTTGGCCGCT	-AGATT
Seq ID 4	(146)	TTCCCTCCCTCTCTCTCTCGTGAATCTCATTTC	CCCCCTTACC	CGCT-AGATT
Seq ID 46	(152)	TTCCCTCCCTCTCTCTCTCGTGAATCTCATTTC	CCCCCTTACC	CGCTA-GATT
Seq ID 47	(1)			
AC189461	(159)	TTAAACCTCTCTCTCTCGTGAATCTCATTTC	CCCCCTTACC	CGCT-AGATT
BH723504	(294)	TTAAATA--AACCTCTCTCGTGAATCT--	-----C-----CGCT-	AGATT
Consenso	(301)	TTAATCC CTCTCTCTCTCGTGAATCTCATTTC	CCCCCTTACC	CGCT AGATT
		351		400
Seq ID 3	(186)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	CT-TCCTCATTAC	TTTTTT
Seq ID 4	(194)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	-----TCCTCATTAT	TTTTTT
Seq ID 46	(200)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	-----TCCTCATAT	TTTTTT
Seq ID 47	(1)			
AC189461	(208)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	CTTCCTCATTAC	TTTTTT
BH723504	(327)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	-----CCTCTCCAT	TTTGT
Consenso	(351)	CTCTCTTACCCTTTTCTCGCCCTTCCTTC	TCCTCATTAT	TTTTTT
		401		450
Seq ID 3	(235)	GTCGTCTTCTGCTCTCTCTC-TCCTC-AGCACTC	-----TTGGCCATG	
Seq ID 4	(240)	GTCGTCTTCTGCTCTCTCTCAGCACTC-TCCTC	-----TTAGCTATG	
Seq ID 46	(246)	GTCGTCTTCTGCTCTCTCTCAGCACTC-TCCTC	-----TTAGCTATG	
Seq ID 47	(1)			
AC189461	(258)	GTCGTCTTCTGCTCTCTCTC-TCCTC-AGCACTC	TTAGCTATG	
BH723504	(372)	-TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	-----GCCATG	
Consenso	(401)	GTCGTCTTCTGCTCTCTCTC TC CTC AGCACTC	TT GCCATG	
		451		500
Seq ID 3	(277)	GTGGGTGCTGCTGCGTCTTC--CTGTTACGCATCT	CGCTATGCACCTG	
Seq ID 4	(282)	GTGGGTGCTGCTGCGTCTTCTTCTGTTACGCATCT	CGCTATGCACCTG	
Seq ID 46	(289)	GTGGGTGCTGCTGCGTCTTCTTCTGTTACGCATCT	CGCTATGCACCTG	
Seq ID 47	(21)	GTGGGTGCTGCTGCGTCTTC--CTGTTACGCATCT	CGCTATGCACCTG	
AC189461	(306)	GTGGGTGCTGCTGCGTCTTC--CTGTTACGCATCT	CGCTATGCACCTG	
BH723504	(410)	GTGGGTGCTGCTGCGTCTTC--CTGTTACGCATCT	CGCTATGCACCTG	
Consenso	(451)	GTGGGTGCTGCTGCGTCTTC CTGTTACGCATCT	CGCTATGCACCTG	
		501		550
Seq ID 3	(324)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
Seq ID 4	(332)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
Seq ID 46	(339)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
Seq ID 47	(68)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
AC189461	(353)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
BH723504	(457)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
Consenso	(501)	GTTCCGTGCGCTGCATGTCCTCTCCACGGCGGGG	GAGATTCCCGAC	
		551		600
Seq ID 3	(374)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
Seq ID 4	(382)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
Seq ID 46	(389)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
Seq ID 47	(118)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
AC189461	(403)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
BH723504	(507)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
Consenso	(551)	AAGCCGTCGCTCTCAAATCTACCGGGCGGAGTC	CGTCAAGAGACAAACAG	
		601		650

FIG. 8B


```

Seq ID 3 (424) CTCACCAATGCT-----CTGGATCCGGTAGCAG-----
Seq ID 4 (432) CTCACCAATGCTGGCTCTTTCTGGATCCGGTAGCSTT---CAGGAGGC
Seq ID 46 (439) CTCACCAATGCT-----CTGGATCCGGTAGCAG-----
Seq ID 47 (168) CTCACCAATGCT-----CTGGATCCGGTAGCAG-----
AC189461 (453) CTCACCAATGCT-----CTGGATCCGGTAGCAGCGTTCAGCAGGC
BH723504 (542) CTCGGCAATGCT-----CTGGATCCGGTAGCATTCTCGTGTCTTC
Consenso (601) CTCACCAATGCT-----CTGGATCCGGTAGCAG-----
651 700

Seq ID 3 (453) -----CACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
Seq ID 4 (479) TCTCGTCACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
Seq ID 46 (468) -----CACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
Seq ID 47 (197) -----CACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
AC189461 (494) TCTCGTCACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
BH723504 (583) -----TTGTTTGGAGTTTAAAGCCTTGCAATCACTACAACAAGAACACA
Consenso (651) CACTACTTCCCTTGGGCCTTGCAATCACTACAAT-----
701 750

Seq ID 3 (487) ---GCCTTGTCTT---CTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
Seq ID 4 (519) ---GCCTTGTCTT---CTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
Seq ID 46 (502) ---GCCTTGTCTT---CTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
Seq ID 47 (231) ---GCCTTGTCTT---CTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
AC189461 (534) ---GCCTTGTCTT---CTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
BH723504 (627) AAGCAATGCTTTCCTCTCTTCGGATCGAATAGCCTTCTCTCAATCGA
Consenso (701) GCCTTGTCTTCTCTCTTCGGATCGAATCTGTTTCTCTCAATCGA
751 FWD-----800

Seq ID 3 (532) AACCCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTCCGGTGGAGGAGC
Seq ID 4 (564) AACCCAGAGGAGGTTGAACTCGGCTG---CTGCTAGTCCGGTGGAGGAGC
Seq ID 46 (547) AACCCAGAGGAGGTTGAMTCGAGCTGCTACTGCTAGTCCGGTGGAGGAGC
Seq ID 47 (276) AACCCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTCCGGTGGAGGAGC
AC189461 (579) AACCCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTCCGGTAAATTTAAC
BH723504 (677) AACCCAGAGGAACTGAATCGAGCAACTTCTCTCTTCCGGTAAATTTAAC
Consenso (751) AACCCAGAGGAGGTTGACTCGAGCTGCTACTGCTAGTCCGGTGGAGGAGC
801 -----REV-----850

Seq ID 3 (582) -----C---ATGGCTGTTGCGATGGATA---TGGAAAAGCAAGCCAAGGTT
Seq ID 4 (611) -----C---ATGGCTGTTGCGATGGATA---TGGAAAAGCAAGCCAAGGTT
Seq ID 46 (597) -----C---ATGGCTGTTGCGATGGATA---TGGAAAAGCAAGCCAAGGTT
Seq ID 47 (326) -----C---ATGGCTGTTGCGATGGATA---TGGAAAAGCAAGCCAAGGTT
AC189461 (629) ATTTCACT---CAATCTGGTFTTTGATTTGTTGGTATTGTTGACAGGTT
BH723504 (727) TCTTCTCCCAAAATCTCAAGTCTTGGTTTCTGGTATCTGTTGCAAGCT
Consenso (801) C ATGGCTGTTGCGATGGATA TGGAAAAGCAAGCCAAGGTT
851 900

Seq ID 3 (622) GACAAACAAACCTCCTACG GAGCAGCG---CCGGTGTGTTGTGACAGGCCA
Seq ID 4 (651) GACAAACAAACCCCTACG GAGCAGCG---CCGTGTTGTGGTGACAGGCCA
Seq ID 46 (637) GACAAACAAACCCCTACG GAGCAGCG---CCGTGTTGTGGTGACAGGCCA
Seq ID 47 (366) GACAAACAAACCTCCTACG GAGCAGCG---CCGGTGTGTTGTGACAGGCCA
AC189461 (676) GACAAGCTTTGTTTGTGTAGTAAAGTTCTCTTTTATGTTGTTTATGT
BH723504 (776) GACAAGCTTTGTTTGTGTATAAAGTTCTCTTT-----
Consenso (851) GACAACAAACCTCCTACG GAGCAGCG CCGTGTGTTGTGACAGGCCA
901 950

```

FIG. 8C

```

Seq ID 3 (668) TGGGAGTTGAAACATCACTAGGTCATGACCTGACACCTTTTATGAGAAT
Seq ID 4 (697) TGGGAGTTGAAACATCACTAGGTCAGGGGCGAATC-----
Seq ID 46 (683) TGGGAGTTGAAACATCACTAGGTCAGGGGCGAATC-----
Seq ID 47 (412) TGGGAGTTGAAACATCACTAGGTCAGGGGCGAATC-----
AC189461 (725) TATTTCGTCAAAATGATTCACTTGGAGATCAAGTTTAAACGGGAAAGTATG
BH723504 (813) -----
Consenso (901) TGGGAGTTGAAACATCACTAGGTCAGGGGCGAATC
951 1000
Seq ID 3 (718) CTCCTACAAGGGGCGAATC-----
Seq ID 4 (733) -----
Seq ID 46 (719) -----
Seq ID 47 (448) -----
AC189461 (775) CTGATTGGAAATCTCAGCTTAATTGTCAAGTTTCATTCTTTATGTGTGT
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) ATAAGGAGTATGATGATAC
BH723504 (813) -----
Consenso (1051)

```

FIG. 8D

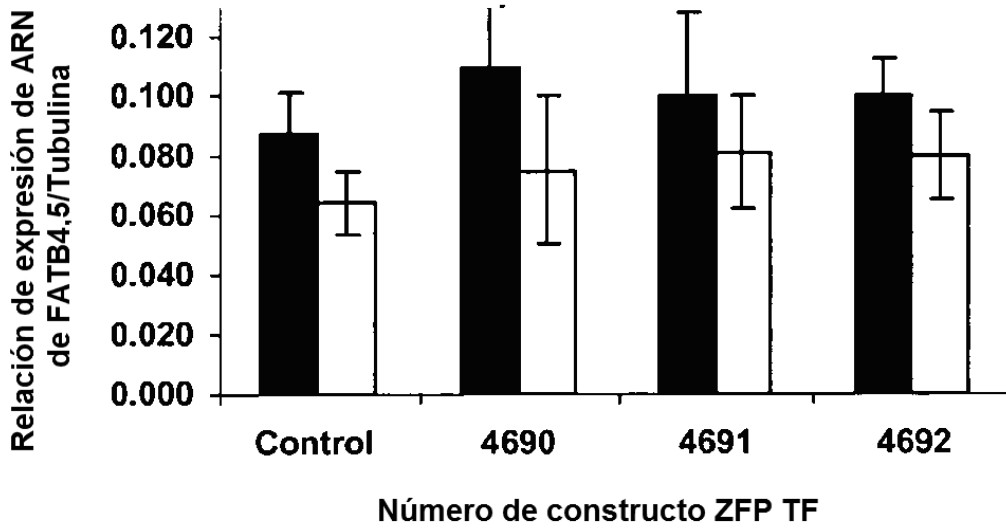


FIG. 9

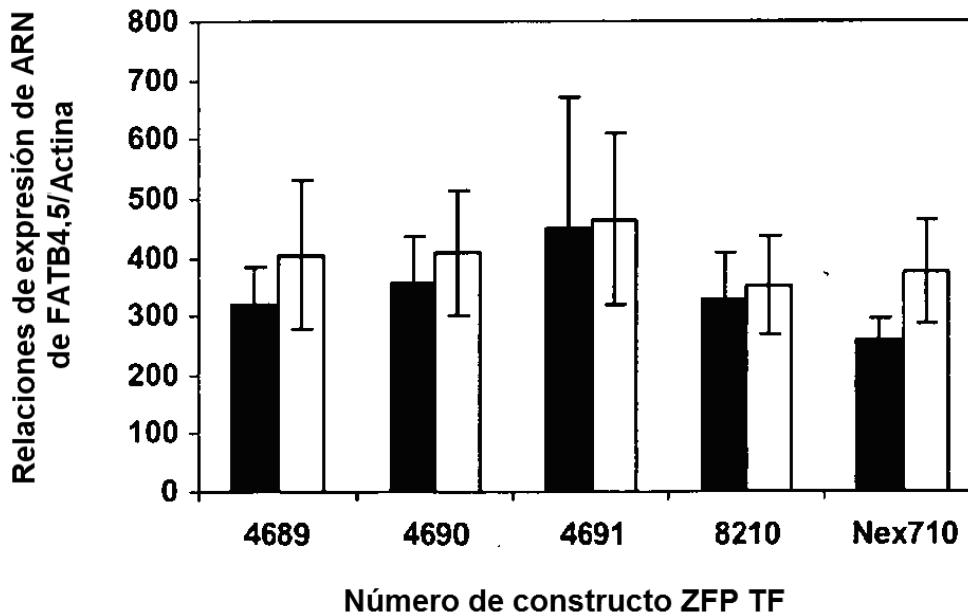


FIG. 10

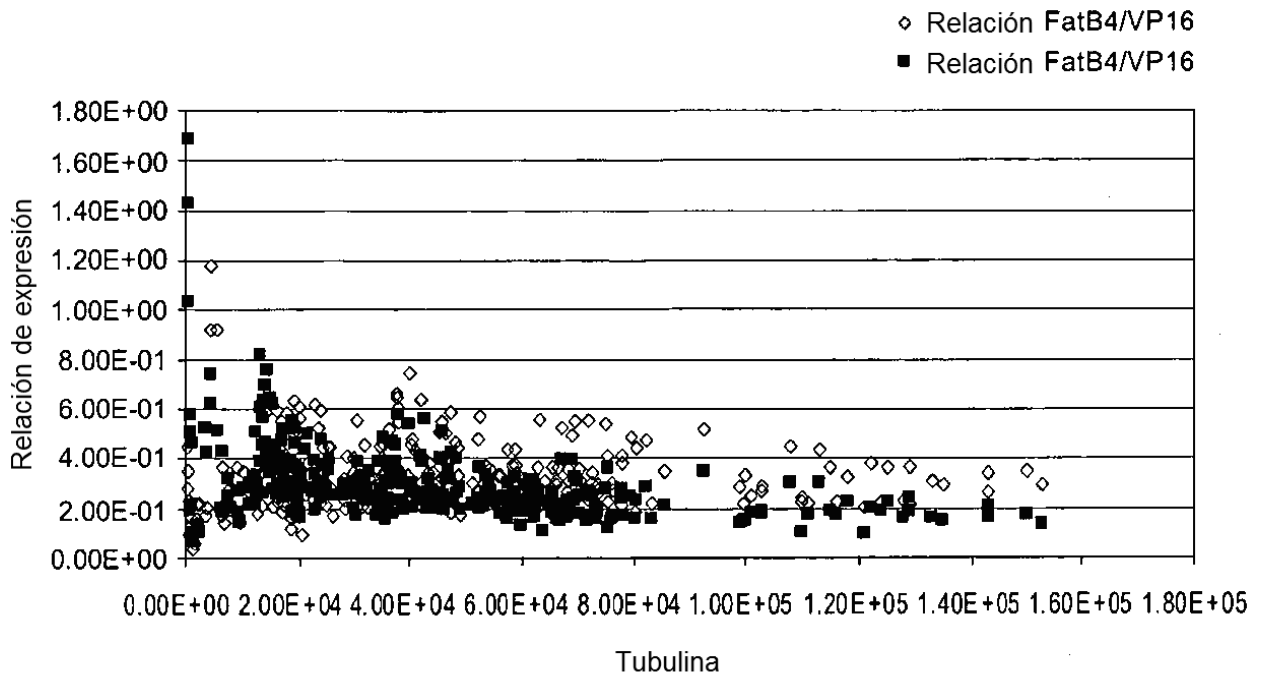


FIG. 11

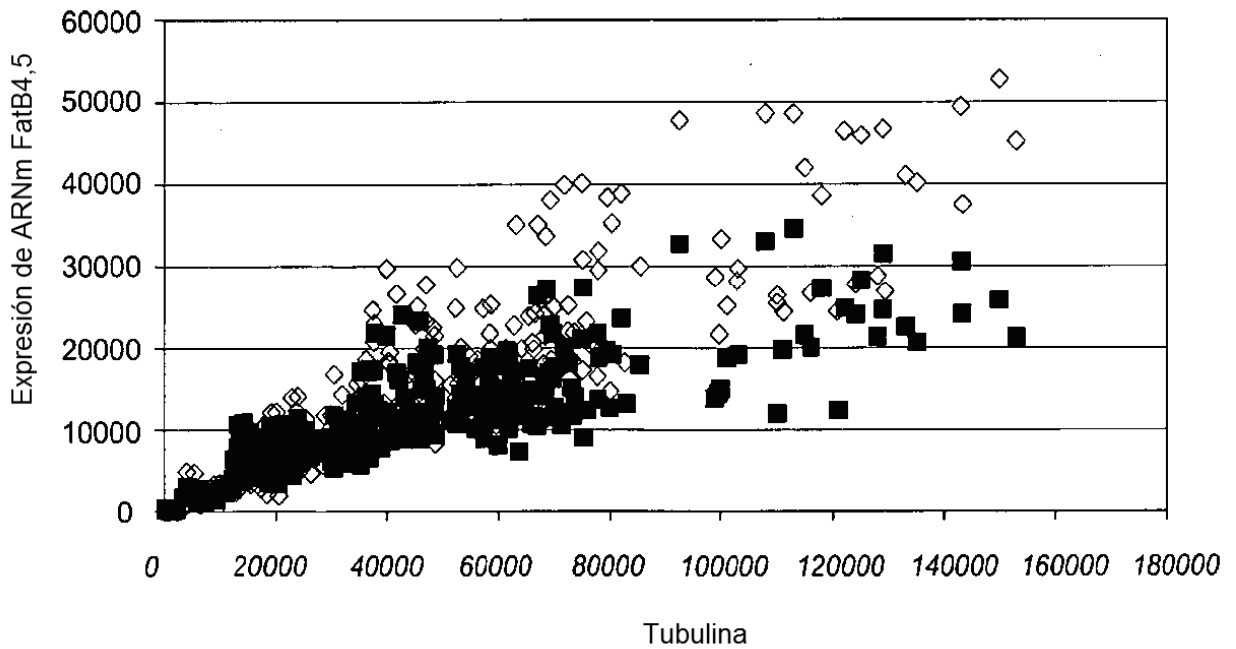


FIG. 12

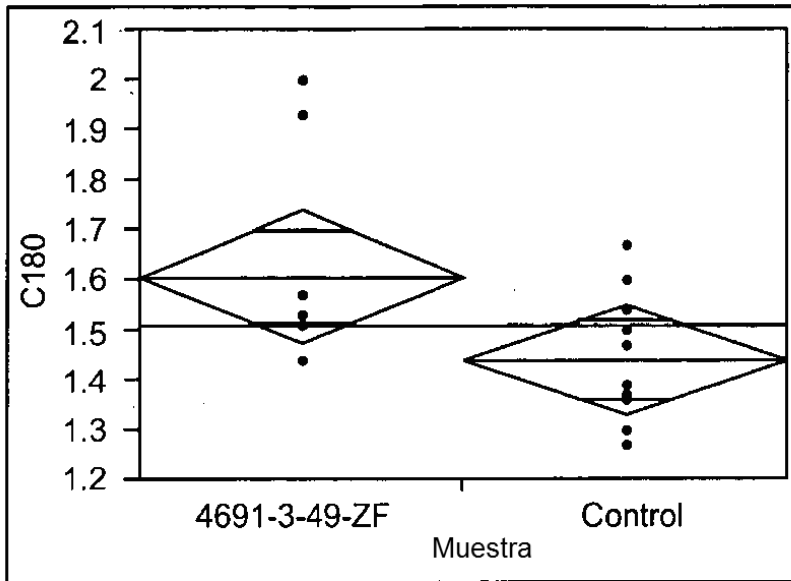


FIG. 13

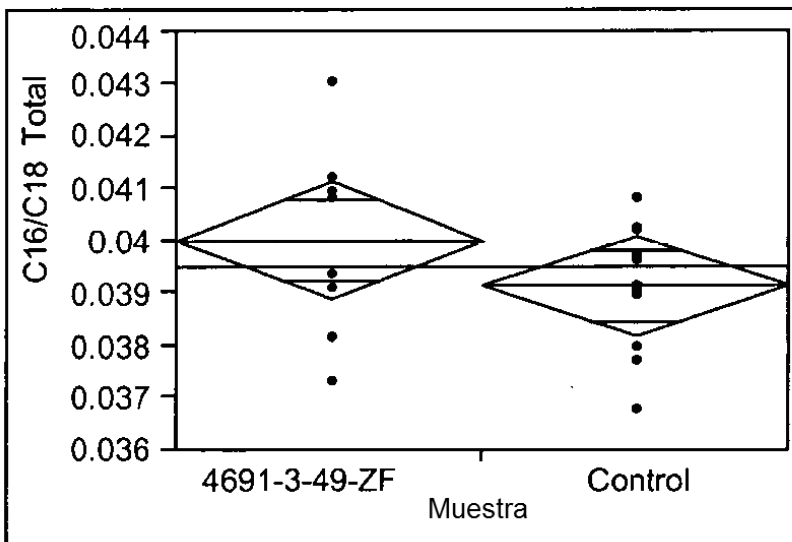


FIG. 14

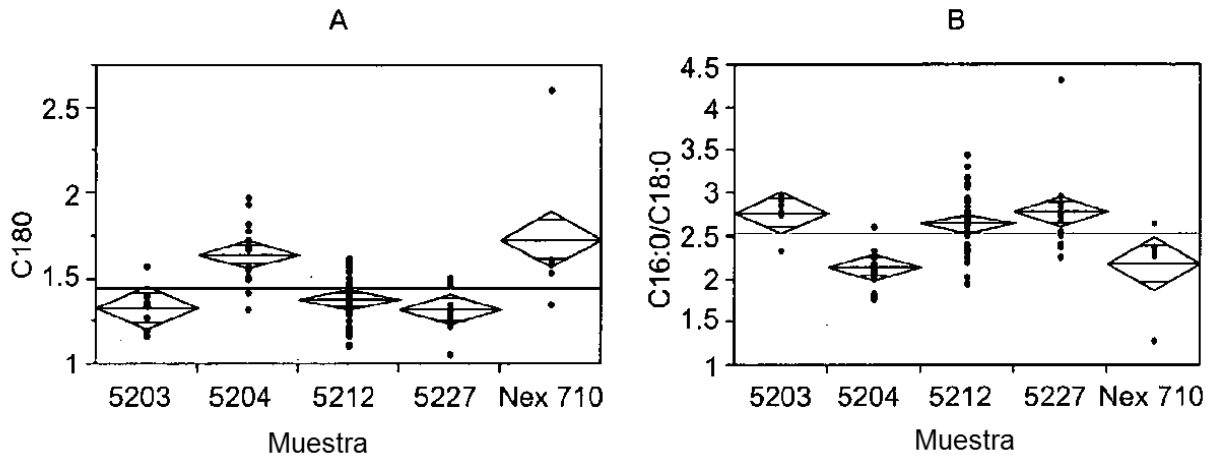


FIG. 15

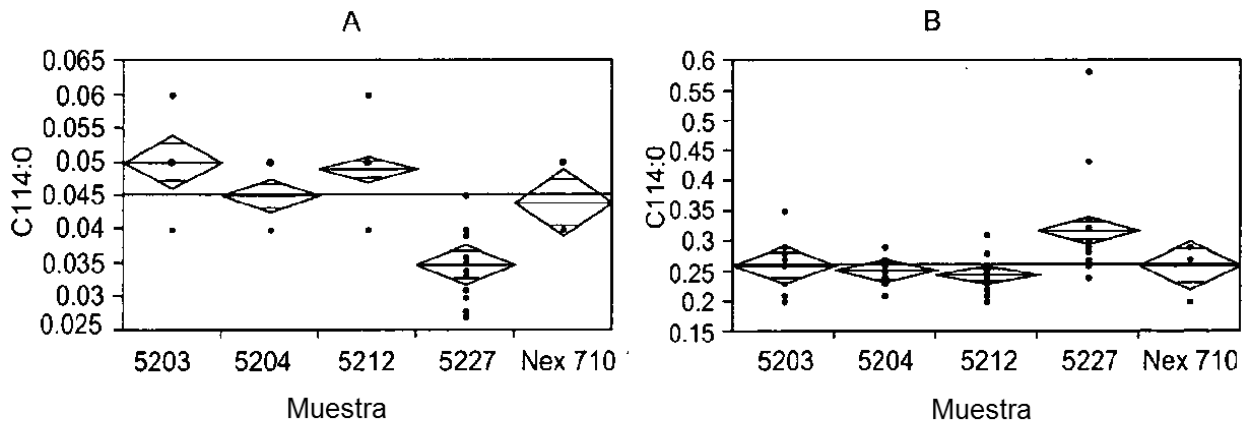


FIG. 16

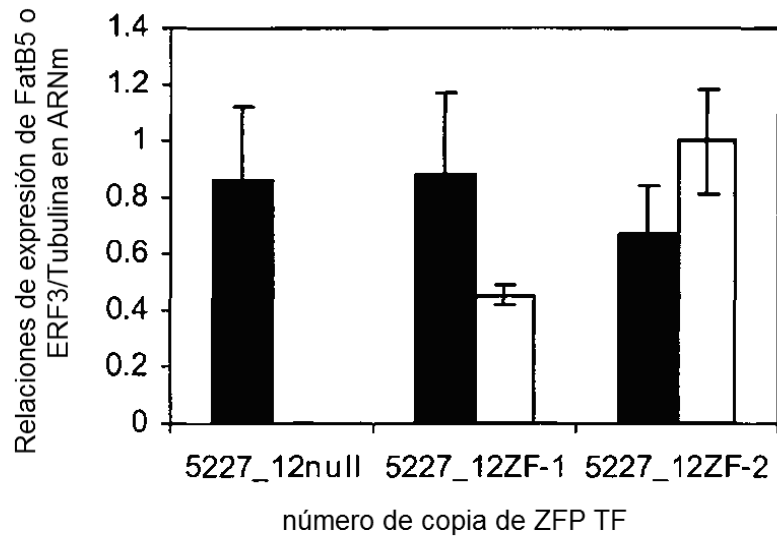


FIG. 17

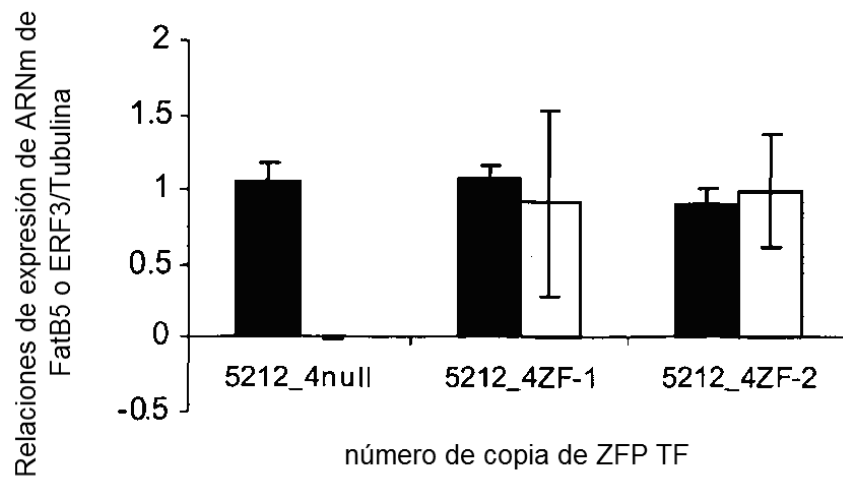


FIG. 18

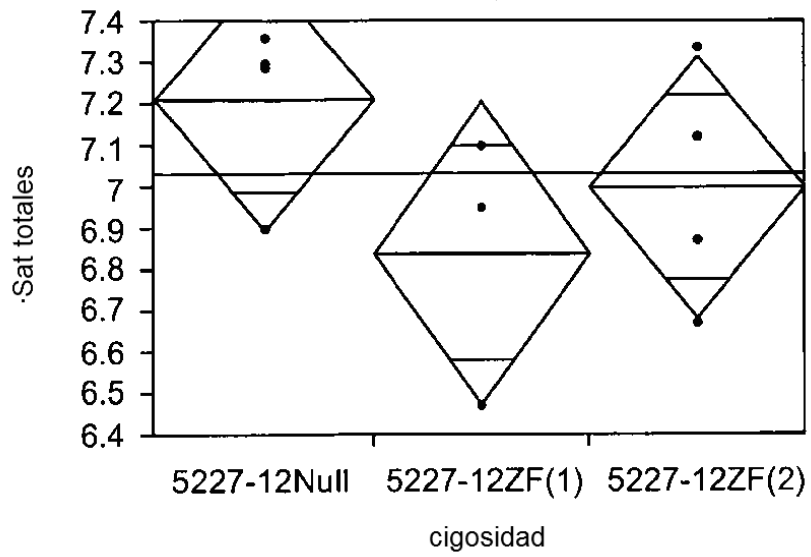


FIG. 19

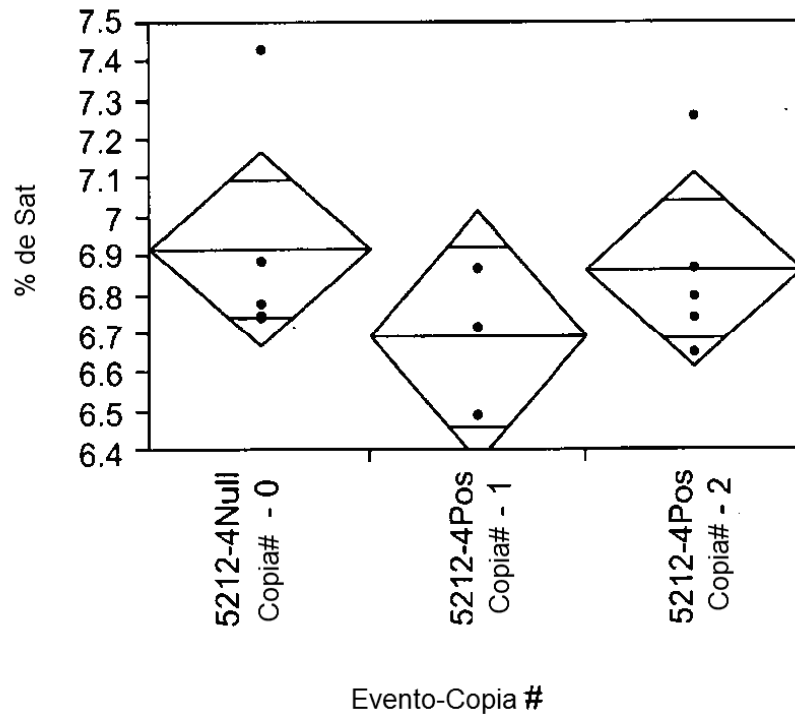


FIG. 20