

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 415 761**

51 Int. Cl.:

**C07K 14/415** (2006.01)

**C12N 15/82** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.11.2009 E 09759899 (9)**

97 Fecha y número de publicación de la concesión europea: **24.04.2013 EP 2370456**

54 Título: **Casetes de expresión para la expresión específica de semillas en plantas**

30 Prioridad:

**26.11.2008 EP 08356145**  
**12.12.2008 US 201648 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**26.07.2013**

73 Titular/es:

**BAYER CROPSCIENCE NV (50.0%)**  
**J.E. Mommaertslaan 14**  
**1831 Diegem, BE y**  
**CNRS CENTRE NATIONAL DE LA RECHERCHE**  
**SCIENTIFIQUE (50.0%)**

72 Inventor/es:

**PINOT, FRANCK;**  
**WERCK, DANIELE;**  
**EHLTING, JÜRGEN;**  
**OLRY, ALEXANDRE;**  
**DENOLF, PETER;**  
**VAN AUDENHOVE, KATRIEN;**  
**VAN DE PUTTE, LIEN y**  
**POSADA, ESMERALDA**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 415 761 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

Casetes de expresión para la expresión específica de semillas en plantas

Campo de la invención

5 La presente invención se refiere a métodos para la expresión de un gen de interés específicamente en semillas de plantas, incluso más específicamente en plantas oleaginosas. La invención proporciona además el uso de un casete de expresión para regular la expresión específica de semillas en plantas.

Introducción a la invención

10 Las plantas usan carbono fijado fotosintéticamente para respaldar el crecimiento y para acumular productos de reserva, tales como almidón o lípidos. El aceite de almacenamiento (triacilglicerol) es un producto vegetal importante con gran importancia económica en la nutrición humana y como una materia prima renovable para diversos productos industriales y biocombustibles. La producción a nivel mundial de aceite vegetal es aproximadamente 100 millones de toneladas métricas en total por año, que consiste principalmente en haba de soja, aceite de palma, aceite de colza y aceite de girasol. La producción de colza está aumentando a nivel mundial. La cosecha se usa principalmente para pienso y alimento, pero, cada vez más, se está usando en la producción de biocombustibles.

15 Como resultado de la gran importancia económica de los aceites vegetales y su uso ampliado como materia prima renovable, hay un considerable interés en la manipulación metabólica del contenido incrementado y/o modificado de oleaginosas. En semillas de colza (*Brassica napus* L.), en desarrollo, la sacarosa se descarga del floema y se metaboliza a intermedios glucolíticos, tales como fosfatos de hexosas, fosfoenolpiruvato y piruvato, que son importados subsiguientemente en el plastidio y usados para la síntesis de ácidos grasos. Los ácidos grasos libres son activados a ésteres de coenzima A (CoA), son exportados desde el plastidio y se usan para la acilación por etapas de la cadena principal de glicerol para sintetizar triacilglicerol en el retículo endoplásmico. En las dos primeras etapas del ensamblaje del triacilglicerol (TAG), glicerol-3-fosfato (Gly3P) se acila mediante Gly3P aciltransferasa (GPAT) hasta ácido lisofosfatídico, que entonces se acila posteriormente mediante ácido lisofosfatídico aciltransferasa (LPAT) hasta ácido fosfatídico. A esto le sigue la desfosforilación del ácido fosfatídico por ácido fosfatídico fosfohidrolasa, para liberar diacilglicerol (DAG), y la acilación final de diacilglicerol por DAG aciltransferasa (DAGAT). El almacenamiento final de triacilglicerol se produce en cuerpos grasos derivados del retículo endoplásmico.

20 La modificación de plantas que producen aceite, para alterar y/o mejorar las características fenotípicas (tales como productividad o calidad), requiere la sobreexpresión o disminución de genes endógenos, o la expresión de genes heterólogos en tejidos vegetales. Tal modificación genética se basa en la disponibilidad de un medio para llevar a cabo y controlar la expresión génica según se requiera. De hecho, la modificación genética se basa en la disponibilidad y uso de promotores adecuados, que son eficaces en plantas y que regulan la expresión génica para dar el efecto o efectos deseados en la planta transgénica. Para numerosas aplicaciones en biotecnología vegetal, es ventajoso un perfil de expresión específico de tejidos, puesto que efectos beneficiosos de la expresión en un tejido pueden tener desventajas en otros. Los promotores preferentes de semillas o específicos de semillas son útiles para expresar o disminuir genes, así como para producir grandes cantidades de proteína, y para producir aceites o proteínas de interés. Es ventajoso tener la elección de una variedad de diferentes promotores, de manera que se pueda seleccionar el promotor más adecuado para un gen particular, constructo, célula, tejido, planta o entorno. Además, el creciente interés en plantas cotransformantes con múltiples casetes de transcripción, y los problemas potenciales asociados con el uso de secuencias reguladoras comunes para estos fines, requiere una variedad de secuencias promotoras. Por lo tanto, existe una gran necesidad en la técnica de identificar nuevas secuencias que se puedan usar para la expresión de transgenes seleccionados en plantas económicamente importantes tales como plantas productoras de aceite. De este modo, es un objetivo de la presente invención proporcionar casetes de expresión nuevos y alternativos para la expresión específica de semillas de transgenes en plantas. Este objetivo se resuelve mediante la presente invención como se explica aquí adicionalmente.

25 Las citocromo P450 monooxigenasas, que catalizan etapas de oxigenación específicas del sustrato, regioespecíficas y estereoespecíficas en el metabolismo vegetal, han evolucionado hasta una enorme superfamilia de enzimas. Recientemente, las iniciativas de secuenciación del genoma vegetal revelaron más de 280 genes de longitud completa en *Arabidopsis thaliana*, 356 en arroz y 312 en *Populus trichocarpa*. Sin embargo, menos del 20% de las secuencias codificantes de las citocromo P450 monooxigenasas en el genoma de *A. thaliana* se han asociado con una función bioquímica específica.

Sumario de la invención

55 Durante nuestra investigación, nos interesamos en la función de las enzimas del citocromo P450 "huérfanas". Una familia de enzimas particular que se caracterizó fue la familia CYP704 de citocromo P450 monooxigenasas, que consiste en los dos miembros CYP704A1 y CYP704A2. De forma notable, se demostró que un casete de expresión que comprende una fusión quimérica de un promotor de CYP704A1 con ácido nucleico se expresó específicamente durante las etapas tardías del desarrollo de las semillas. La invención descrita aquí en las diferentes realizaciones, ejemplos, figuras y reivindicaciones proporciona el uso de promotores específicos de semillas y regiones promotoras

comprendidas en casetes de expresión para dirigir la expresión de genes heterólogos en semillas. En una realización de la invención, se proporciona un método para regular la expresión específica de semillas en plantas, que comprende (a) proporcionar un casete de expresión para regular la expresión específica de semillas en plantas, que comprende un promotor enlazado a un ácido nucleico que es heterólogo con relación a dicho promotor, y en el que dicho promotor tiene la secuencia nucleotídica de SEC ID NO: 1 o una secuencia nucleotídica que tiene al menos 95% de identidad de secuencias con SEC ID NO: 1, y (b) transformar una planta con dicho casete de expresión. SEC ID NO: 1 representa la secuencia nucleotídica del promotor del gen del citocromo P450 704A1 procedente de *A. thaliana*. Dicho casete de expresión puede dirigir la expresión de una proteína, un polipéptido, un péptido, un ARN antisentido, un ARN sentido o un ARN bicatenario.

En una realización particular, dicho método se aplica a una planta usada para la producción de aceite. Las plantas particulares son cánola, maíz, mostaza, haba de ricino, sésamo, algodón, linaza, haba de soja, *Arabidopsis*, *Phaseolus*, cacahuete, alfalfa, trigo, arroz, avena, sorgo, colza, centeno, caña de azúcar, alazor, palmas de aceite, lino, girasol, *Brassica campestris*, *Brassica napus*, *Brassica juncea*, *Crambe abyssinica*.

En una realización adicional, el método según la invención comprende además las etapas de (c) hacer crecer una planta que contiene dicho casete de expresión, en el que dicha planta produce semilla, y dicho ácido nucleico se transcribe en dicha semilla, y (d) aislar dicha semilla de dicha planta transformada.

En todavía otra realización, la invención proporciona el uso del casete de expresión como se describe para regular la expresión específica de semillas de un gen heterólogo en plantas.

#### Breve descripción de las figuras

Figura 1: Análisis de RT-PCR cuantitativa de la expresión del gen *CYP704A1* en diferentes órganos de *Arabidopsis thaliana*. Las cantidades de transcritos de *CYP704A1* ( $2^{-\Delta\Delta Ct}$ ) se normalizan con respecto a la referencia endógena *ACTIN2*. Los datos representan la media de 3 muestras con desviación estándar. El experimento se repitió dos veces con resultados similares. Representación en forma de gráfica de barras verticales de la evaluación mediante RT-PCR cuantitativa de la expresión de los transcritos de *CYP704A1* en diferentes órganos de *Arabidopsis thaliana*. Los (datos  $2^{-\Delta\Delta Ct}$ ) se dedujeron mediante el método Ct.

Figura 2 (para referencia solamente): RT-PCR cuantitativa de la expresión del gen *CYP704A2* en diferentes órganos de *Arabidopsis thaliana*. Las cantidades de transcritos de *CYP704A2* ( $2^{-\Delta\Delta Ct}$ ) se normalizan con respecto a la referencia endógena *ACTIN2*. Los datos representan la media de 3 muestras con desviación estándar. El experimento se repitió dos veces con resultados similares. Representación en forma de gráfica de barras verticales de la evaluación mediante qRT-PCR de la expresión de los transcritos de *CYP704A2* en diferentes órganos de *Arabidopsis thaliana*. Los (datos  $2^{-\Delta\Delta Ct}$ ) se dedujeron mediante el método Ct.

Figura 3 (para referencia solamente): RT-PCR cuantitativa de la expresión del gen *CYP704A2* en diferentes órganos de *Arabidopsis thaliana* con una sonda específica 5'-terminal. Las cantidades de transcritos de *CYP704A2* ( $2^{-\Delta\Delta Ct}$ ) se normalizan con respecto a la referencia endógena *ACTIN2*. Los datos representan la media de 3 muestras con desviación estándar. El experimento se repitió dos veces con resultados similares. Representación en forma de gráfica de barras verticales de la evaluación mediante qRT-PCR de la expresión de los transcritos de *CYP704A2* en diferentes órganos de *Arabidopsis thaliana* con un cebador específico 5'-terminal. Los (datos  $2^{-\Delta\Delta Ct}$ ) se dedujeron mediante el método Ct.

#### Descripción detallada de las realizaciones preferidas

La presente invención proporciona métodos para regular la expresión específica de semillas en plantas. La secuencia nucleotídica representada en SEC ID NO: 1 representa la secuencia nucleotídica del promotor del gen de la citocromo P450 monooxigenasa 704A1 (designado aquí en lo sucesivo como CYP704A1) de *Arabidopsis thaliana*. Este gen está situado en el cromosoma II, y se describe por el locus At2g44890 de *A. thaliana* de GenBank. De este modo, se describe un casete de expresión para regular la expresión específica de semillas en plantas, que comprende al menos una secuencia nucleotídica que regula la transcripción derivada del gen de *Arabidopsis thaliana* descrito por el locus At5g44890 del genoma de GenBank, y operablemente (o "funcionalmente") enlazado a ella al menos una secuencia de ácido nucleico que es heteróloga con relación a dicha secuencia nucleotídica que regula la transcripción.

También se describe un promotor específico de semillas que tiene la secuencia nucleotídica de SEC ID NO: 1 desde la posición nucleotídica 1 a la posición nucleotídica 1210, para uso en un casete de expresión según los métodos de la invención. Un promotor específico de semillas adecuado para uso en los métodos de la invención tiene la secuencia nucleotídica de SEC ID NO: 1 desde la posición nucleotídica 600 a la posición nucleotídica 1210. También es adecuado un promotor específico de semillas que tiene la secuencia nucleotídica de SEC ID NO: 1 desde la posición nucleotídica 800 a la posición nucleotídica 1210. SEC ID NO: 1 representa la región en dirección 5' (es decir, situada en dirección 5' de) a partir del codón que codifica el primer aminoácido de la proteína CYP704A1. Tal región promotora puede tener al menos alrededor de 300 a alrededor de 400 a alrededor de 500 pb, al menos

alrededor de 1000 pb, al menos alrededor de 1100 pb, al menos alrededor de 1200 pb en dirección 5' del codón de partida del gen CYP704A1.

Las frases "secuencia de ADN", "secuencia de ácido nucleico" y "molécula de ácido nucleico" se refieren a una estructura física que comprende una disposición ordenada de nucleótidos. La secuencia de ADN o secuencia nucleotídica puede estar contenida en una molécula nucleotídica más grande, en un vector, o similar. Además, la disposición ordenada de ácidos nucleicos en esta secuencia se puede representar en forma de un listado de secuencias, una figura, una tabla, un medio electrónico, o similar. El término "expresión" se refiere a la transcripción de un gen para producir el ARN correspondiente. En una realización particular, dicho ARN es ARNm, y la traducción de este ARNm produce el producto génico correspondiente (es decir, un péptido, polipéptido, o proteína). En otra realización particular, el ácido nucleico heterólogo, enlazado operablemente a los promotores de la invención, puede codificar también ARN antisentido, ARN sentido, ARN bicatenario o moléculas de microARN sintéticas, según reglas bien conocidas en la técnica, para disminuir la expresión de otros genes comprendidos en la semilla, o incluso de genes presentes en un patógeno o plaga que se alimenta de las semillas de la planta transgénica.

El término "heterólogo" se refiere a la relación entre dos o más secuencias de ácidos nucleicos o proteicas que derivan de diferentes fuentes. Por ejemplo, un promotor es heterólogo con respecto a una secuencia codificante si tal combinación no se encuentra normalmente en la naturaleza. Además, una secuencia particular puede ser "heteróloga" con respecto a una célula u organismo en el que se inserta (es decir, no aparece de forma natural en esa célula u organismo particular).

La expresión "gen quimérico" se refiere a cualquier gen que contiene: a) secuencias de ADN, incluyendo secuencias reguladoras y codificantes que no se encuentran juntas en la naturaleza, o b) secuencias que codifican partes de proteínas no unidas de forma natural, o c) partes de promotores que no están unidos de forma natural. En consecuencia, un gen quimérico puede comprender secuencias reguladoras y secuencias codificantes que derivan de diferentes fuentes, o comprende secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero dispuestas de manera diferente de la encontrada en la naturaleza. En la actual invención, un gen o polinucleótido o polipéptido "homólogo" se refiere a un gen o polinucleótido o polipéptido que comparte una similitud de secuencia con el gen o polinucleótido o polipéptido de interés.

Expresión "específica de semillas" (o "transcripción", que es equivalente), en el contexto de esta invención, significa transcripción de una secuencia de ácido nucleico mediante un promotor (o un elemento que regula la transcripción) de manera que la transcripción de dicha secuencia de ácido nucleico en semillas contribuye a más de 90%, preferiblemente más de 95%, más preferiblemente más de 99% de toda la cantidad del ARN transcrito desde dicha secuencia de ácido nucleico en toda la planta durante cualquiera de sus etapas de desarrollo.

Expresión "preferente de semillas" (o "transcripción", que es equivalente), en el contexto de esta invención, significa la transcripción de una secuencia de ácido nucleico mediante un elemento que regula la transcripción de manera que la transcripción de dicha secuencia de ácido nucleico en semillas contribuye a más de 50%, preferiblemente más de 60%, más preferiblemente más de 70%, incluso más preferiblemente más de 80% de toda la cantidad del ARN transcrito a partir de dicha secuencia de ácido nucleico en toda la planta durante cualquiera de sus etapas de desarrollo. La expresión "potenciada por semillas" es equivalente a la expresión "preferente de semillas".

Se encuentra que las secuencias de transcripción identificadas aquí median una expresión potente específicamente en semillas. Un casete de expresión que comprende SEC ID NO: 1 y sus fragmentos dirige la expresión específicamente en silicuas durante el desarrollo de las semillas y en semillas secas maduras. Una semilla representa un embrión en su cáscara, o el embrión es una parte de la semilla. En el contexto de esta invención, la expresión específica de semillas también significa por lo tanto específica de embriones. Las expresiones específica de semillas y específica de embriones se pueden usar aquí de forma intercambiable.

"Semilla" significa una semilla de una planta en cualquier etapa de su desarrollo, es decir, partiendo desde la fusión de polen y oocito, continuando a lo largo de la etapa de embrión y la etapa de la semilla latente, hasta la semilla germinante, terminando con órganos de plántulas tempranos, como por ejemplo cotiledones e hipocótilos. "Microspora" – en plantas de semillas – corresponde al grano de polen en desarrollo en la etapa uninucleada.

La frase "enlazada operablemente" se refiere a la disposición espacial funcional de dos o más regiones de ácidos nucleicos o secuencias de ácidos nucleicos. Por ejemplo, una región promotora puede estar situada con respecto a una secuencia de ácido nucleico de manera que la transcripción de una secuencia de ácido nucleico es dirigida por la región promotora. De este modo, una región promotora está "operablemente enlazada" a la secuencia de ácido nucleico. "Funcionalmente enlazada" es una expresión equivalente.

Como se usa aquí, "promotor" significa una región de secuencia de ADN que es esencial para el inicio de la transcripción del ADN, dando como resultado la generación de una molécula de ARN que es complementaria al ADN transcrito; esta región también se puede denominar como una "región reguladora de 5'". Los promotores están habitualmente localizados en dirección 5' de la secuencia codificante a transcribir, y tiene regiones que actúan como sitios de unión para ARN polimerasa II y otras proteínas, tales como factores de transcripción (factores proteicos que actúan en trans, que regulan la transcripción), para iniciar la transcripción de un gen enlazado operablemente. Los

promotores pueden contener en sí mismos subelementos (es decir, motivos de promotores) tales como elementos que actúan en cis o dominios potenciadores que regulan la transcripción de genes enlazados operablemente. Los promotores de esta invención se pueden alterar para que contengan "ADN potenciador" para ayudar a elevar la expresión génica. Como se conoce en la técnica, ciertos elementos de ADN se pueden usar para potenciar la transcripción del ADN. Estos potenciadores a menudo se encuentran en 5' con respecto al inicio de la transcripción en un promotor que funciona en células eucariotas, pero a menudo se pueden insertar en dirección 5' (5') o en dirección 3' (3') con respecto a la secuencia codificante. En algunos casos, estos elementos de ADN potenciadores en 5' son intrones. Entre los intrones que son útiles como ADN potenciador están los intrones en 5' procedentes del gen de actina 1 del arroz (véase el documento US5641876), el gen de actina 2 del arroz, el gen de alcohol deshidrogenasa del maíz, el gen de la proteína 70 de choque térmico del maíz (véase el documento US5593874), el gen shrunken 1 del maíz, el gen sensible a la luz 1 de *Solanum tuberosum*, y el gen de la proteína 70 de choque térmico de *Petunia hybrida* (véase el documento US5659122). De este modo, como se contempla aquí, un promotor o región promotora incluye variaciones de promotores derivadas insertando o suprimiendo regiones reguladoras, sometiendo al promotor a mutagénesis al azar o dirigida al sitio, etc. La actividad o fortaleza de un promotor se puede medir en términos de las cantidades de ARN que produce, o la cantidad de acumulación proteica en una célula o tejido, con respecto a un promotor cuya actividad transcripcional se ha evaluado previamente.

La confirmación de la actividad promotora para un fragmento promotor funcional en semillas se puede determinar por los expertos en la técnica, por ejemplo usando un constructo de promotor-informador que comprende la secuencia genómica enlazada operablemente a un gen informador de beta-glucuronidasa (GUS) como se explica aquí posteriormente. La capacidad de expresión preferencial de semillas de los fragmentos identificados o generados de los promotores de la invención se puede evaluar convenientemente enlazando de forma operable tales moléculas de ADN a una secuencia nucleotídica que codifica un marcador fácilmente puntuable, por ejemplo un gen de beta-glucuronidasa, introduciendo tal gen quimérico en una planta, y analizando el patrón de expresión del marcador en semillas en comparación con el patrón de expresión del marcador en otras partes de la planta. Otros candidatos para un marcador (o un gen informador) son cloranfenicol acetil transferasa (CAT) proteínas con propiedades fluorescentes, tales como proteína fluorescente verde (GFP) de *Aequora victoria*. Para definir una región promotora mínima, se elimina un segmento de ADN que representa la región promotora a partir de la región de 5' del gen de interés, y se enlaza operablemente a la secuencia codificante de un gen marcador (informador) mediante técnicas de ADN recombinante bien conocidas en la técnica. El gen informador está enlazado operablemente en dirección 3' del promotor, de manera que los transcritos que se inician en el promotor transcurren a través del gen informador. Los genes informadores codifican genéricamente proteínas, que se miden fácilmente, incluyendo, pero sin limitarse a, cloranfenicol acetil transferasa (CAT), beta-glucuronidasa (GUS), proteína fluorescente verde (GFP), beta-galactosidasa (beta-GAL), y luciferasa. El casete de expresión que contiene el gen informador bajo el control del promotor se puede introducir en un tipo celular apropiado mediante técnicas de transfección bien conocidas en la técnica. Para evaluar la proteína informadora, se preparan lisados celulares y se llevan a cabo ensayos apropiados, que son bien conocidos en la técnica, para la proteína informadora. Por ejemplo, si el gen información de elección fue CAT, los lisados procedentes de células transfectadas con constructos que contienen CAT bajo el control de un promotor bajo estudio se mezclan con cloranfenicol marcado isotópicamente y acetil-coenzima A (acetil-CoA). La enzima CAT transfiere el grupo acetilo desde acetil-CoA a la posición 2 ó 3 de cloranfenicol. La reacción se monitoriza mediante cromatografía de capa fina, que separa cloranfenicol acetilado del material sin reaccionar. Los productos de la reacción se visualizan entonces mediante autoradiografía. El nivel de actividad enzimática corresponde a la cantidad de enzima que se obtuvo, lo que a su vez revela el nivel de expresión y la funcionalidad específica de semillas a partir del promotor o fragmento promotor de interés. Este nivel de expresión también se puede comparar con otros promotores para determinar la fuerza relativa del promotor bajo estudio. Una vez que se confirman la actividad y funcionalidad, se pueden emplear análisis mutacionales y/o de supresión adicionales para determinar la región mínima y/o secuencias requeridas para iniciar la transcripción. De este modo, las secuencias se pueden suprimir en el extremo 5' de la región promotora y/o en el extremo 3' de la región promotora, y se pueden introducir sustituciones nucleotídicas. Estos constructos se introducen entonces nuevamente en células, y se determina su actividad y/o funcionalidad.

En lugar de medir la actividad de una enzima informadora, también se puede medir la actividad promotora transcripcional (y funcionalidad) midiendo el nivel de ARN que se produce. Este nivel de ARN, tal como ARNm, se puede medir en un único punto de tiempo o en múltiples puntos de tiempo, y como tal el incremento en número de veces puede ser un incremento en número de veces medio, o un valor extrapolado derivado a partir de valores medidos experimentalmente. Puesto que es una comparación de niveles, se puede usar cualquier método que mida niveles de ARNm. En un aspecto preferido, el tejido u órganos comparados son una semilla o tejido de semilla con una hoja o tejido de hoja. En otro aspecto preferido, se comparan múltiples tejidos u órganos. Una comparación múltiple preferida es una semilla o tejido de semilla comparado con 2, 3, 4 o más tejidos u órganos seleccionados del grupo que consiste en tejido floral, ápice floral, polen, hoja, embrión, brote, primordia de las hojas, ápice del brote, raíz, punta de la raíz, tejido vascular y cotiledón. Como se usa aquí, los ejemplos de órganos vegetales son semilla, hoja, raíz, etc., y los ejemplos de tejidos son primordia de las hojas, ápice del brote, tejido vascular, etc. La actividad o fuerza de un promotor se puede medir en términos de la cantidad de ARNm o acumulación proteica que produce específicamente, con respecto a la cantidad total de ARNm o proteína. El promotor expresa preferiblemente una secuencia de ácido nucleico enlazado operablemente en una cantidad mayor que alrededor de 1%, alrededor de 2%, más preferiblemente mayor que alrededor de 5, 6, 7, 8, o alrededor de 9%, incluso más preferiblemente mayor

que alrededor de 10, 11, 12, 13, 14, 15, 16, 17, 18, o alrededor de 19%, y lo más preferible mayor que alrededor de 20% del ARNm total. Como alternativa, la actividad o fuerza de un promotor se puede expresar con respecto a un promotor bien caracterizado (para el cual se evaluó previamente la actividad transcripcional).

5 Será claro aquí además que se pueden aislar de otras plantas promotores CYP704A1 equivalentes. Para este fin, se pueden aislar fragmentos de promotores ortólogos de otras plantas usando SEC ID NO: 1 o un fragmento funcional que tiene al menos 50 nucleótidos consecutivos de la misma como una sonda, y se pueden identificar secuencias nucleotídicas a partir de estas otras plantas que se hibridan bajo las condiciones de hibridación descritas aquí. A título de ejemplo, se puede usar un promotor de la invención para cribar una librería genómica de un cultivo o planta de interés para aislar secuencias promotoras correspondientes según técnicas bien conocidas en la técnica. De este modo, se puede usar una secuencia promotora de la invención como sonda para la hibridación con una librería genómica en condiciones de restricción media a alta. Como equivalente alternativo, se pueden aislar promotores usando la secuencia codificante de CYP704A1 para cribar una librería genómica (por ejemplo mediante hibridación o *in silico*) de un cultivo de interés. Cuando se obtiene una identidad suficiente entre las secuencias codificantes (como regla, identidad mayor que 85%), entonces se pueden aislar regiones promotoras en dirección 5' de los genes CYP704A1 ortólogos. La presente invención proporciona un ejemplo para clonar promotores ortólogos a partir de *Brassica napus* en el ejemplo 4.

El término "hibridación" se refiere a la capacidad de una primera hebra de ácido nucleico para unirse a una segunda hebra vía un emparejamiento de bases mediante enlace de hidrógeno cuando las dos hebras de ácido nucleico tienen identidad de secuencia suficiente. La hibridación se produce cuando las dos moléculas de ácido nucleico se recombinan entre sí en condiciones apropiadas. La hibridación de ácidos nucleicos es una técnica bien conocida por los expertos en la técnica de la manipulación de ADN. La propiedad de hibridación para un par dado de ácidos nucleicos es una indicación de su similitud o identidad. Otra indicación de que dos secuencias de ácidos nucleicos son sustancialmente idénticas es que las dos moléculas se hibridan entre sí en condiciones restrictivas. La frase "que se hibrida específicamente a" se refiere a la unión, formación de dúplex, o hibridación de una molécula sólo a una secuencia nucleotídica particular en condiciones restrictivas cuando esa secuencia está presente en una mezcla compleja de ADN o ARN (por ejemplo celular total). "Se une o unen sustancialmente" se refiere a una hibridación complementaria entre un ácido nucleico sonda y un ácido nucleico diana, y abarca desemparejamientos minoritarios que se pueden ajustar reduciendo la restricción de los medios de hibridación para lograr la detección deseada de la secuencia de ácido nucleico diana. "Condiciones de hibridación restrictivas" y "condiciones de lavado de hibridación restrictivas", en el contexto de experimentos de hibridación de ácidos nucleicos tales como hibridación Southern y Northern, dependen de las secuencias, y son diferentes bajo diferentes parámetros medioambientales. Un ejemplo de condiciones de lavado muy restrictivas es 0,15 M de NaCl a 72°C durante alrededor de 15 minutos. Un ejemplo de condiciones de lavado restrictivas es un lavado con 0,2 X SSC a 65°C durante 15 minutos. A menudo, un lavado de alta restricción va precedido por un lavado de baja restricción, para eliminar la señal de sonda de fondo. Un lavado de restricción media ejemplar para un dúplex de, por ejemplo, más de 100 nucleótidos, es 1 X SSC a 45°C durante 15 minutos. Un lavado de baja restricción ejemplar para un dúplex de, por ejemplo, más de 100 nucleótidos, es 4 a 6 X SSC a 40°C durante 15 minutos. Para sondas cortas (por ejemplo, alrededor de 10 a 50 nucleótidos), las condiciones restrictivas implican típicamente concentraciones salinas menores que alrededor de 1,5 M, más preferiblemente alrededor de 0,01 a 1,0 M, una concentración de ion Na (u otras sales), a pH 7,0 a 8,3, y la temperatura es típicamente al menos alrededor de 30°C, y al menos alrededor de 60°C para sondas largas (por ejemplo, >50 nucleótidos). Las condiciones restrictivas también se pueden lograr con la adición de agentes desestabilizantes, tal como formamida. En general, una relación de señal a ruido de 2 X (o mayor) que la observada para una sonda no relacionada en un ensayo de hibridación particular indica detección de una hibridación específica. Los ácidos nucleicos que no se hibridan entre sí en condiciones restrictivas todavía son sustancialmente idénticos si las proteínas que codifican son sustancialmente idénticas. Esto ocurre, por ejemplo, cuando se crea una copia de un ácido nucleico usando la degeneración de codones máxima permitida por el código genético. Las condiciones muy restrictivas se seleccionan para que sean iguales a la  $T_m$  para una sonda particular. Un ejemplo de condiciones restrictivas para hibridación de ácidos nucleicos complementarios que tienen más de 100 restos complementarios sobre un filtro en una transferencia Southern o Northern es formamida al 50%, por ejemplo hibridación en formamida al 50%, 1 M de NaCl, 1% de SDS a 37°C, y un lavado en 0,1 x SSC a 60 a 65°C. Las condiciones de baja restricción ejemplares incluyen hibridación con una disolución tampón de formamida al 30 a 35%, 1 M de NaCl, 1% de SDS (dodecilsulfato de sodio) a 37°C, y un lavado en 1 X a 2 X SSC (20 X SSC = 3,0 M de NaCl/0,3 M de citrato trisódico) a 50 a 55°C. Las condiciones de restricción moderadas ejemplares incluyen hibridación en formamida al 40 a 45%, 1,0 M de NaCl, 1% de SDS a 37°C, y un lavado en 0,5 X a 1 X SSC a 55 a 60°C. Lo siguiente son ejemplos de conjuntos de condiciones de hibridación/lavado que se pueden usar para clonar secuencias nucleotídicas ortólogas que son sustancialmente idénticas a secuencias nucleotídicas de referencia de la presente invención: una secuencia nucleotídica de referencia se hibrida preferiblemente a la secuencia nucleotídica de referencia en dodecilsulfato de sodio (SDS) al 7%, 0,5 M de  $\text{NaPO}_4$ , 1 mM de EDTA a 50°C, con un lavado en 2 X SSC, 0,1% SDS a 50°C, más deseablemente en dodecilsulfato de sodio (SDS) al 7%, 0,5 M de  $\text{NaPO}_4$ , 1 mM de EDTA a 50°C, con un lavado en 1 X SSC, 0,1% de SDS a 50°C, todavía más deseablemente en dodecilsulfato de sodio (SDS) al 7%, 0,5 M de  $\text{NaPO}_4$ , 1 mM de EDTA a 50°C, con lavado en 0,5 X SSC, 0,1% de SDS a 50°C, incluso todavía más deseablemente en dodecilsulfato de sodio (SDS) al 7%, 0,5 M de  $\text{NaPO}_4$ , 1 mM de EDTA a 50°C, con lavado en 0,1 X SSC, 0,1% de SDS a 50°C.

Los promotores específicos de semillas como se describen aquí pueden comprender una secuencia nucleotídica que tiene al menos 40%, al menos 50%, o al menos 60%, o al menos 70%, o al menos 80%, o al menos 90%, o al menos 95% de identidad de secuencia con los promotores y regiones promotoras descritos aquí. El término "variante", con respecto a la secuencia nucleotídica que regula la transcripción SEC ID NO: 1 de la invención, pretende significar secuencias sustancialmente similares. Las variantes alélicas de origen natural, tales como estas, se pueden identificar con el uso de técnicas de biología molecular bien conocidas, como, por ejemplo, con técnicas de reacción en cadena de la polimerasa (PCR) e hibridación, como se esquematizan aquí anteriormente. Las secuencias nucleotídicas variantes también incluyen secuencias nucleotídicas derivadas sintéticamente, tales como las generadas, por ejemplo, usando mutagénesis de SEC ID NO: 1 dirigida al sitio. Generalmente, las variantes de las secuencias nucleotídicas de la invención tendrán al menos 40%, 50%, 60%, hasta 70%, por ejemplo preferiblemente 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, hasta 79%, generalmente al menos 80%, por ejemplo 81% a 84%, al menos 85%, por ejemplo 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, hasta 98% y 99% de identidad de secuencia nucleotídicas con respecto a la secuencia nucleotídica nativa (de tipo salvaje o endógena). Los derivados de las moléculas de ADN descritas aquí pueden incluir, pero no se limitan a, supresiones de secuencia, mutaciones de un solo punto o de múltiples puntos, alteraciones en un sitio de enzima de restricción particular, adición de elementos opcionales, u otros medios de modificación molecular que pueden potenciar, o de otro modo alterar, la expresión del promotor. Las técnicas para obtener tales derivados son bien conocidas en la técnica (véase, por ejemplo, J. F. Sambrook, D. W. Russell, y N. Irwin (2000) *Molecular Cloning: A Laboratory Manual*, 3ª edición Volúmenes 1, 2, y 3. Cold Spring Harbor Laboratory Press). Por ejemplo, un experto en la técnica puede delimitar los elementos funcionales en los promotores descritos aquí y suprimir cualesquiera elementos no esenciales. Los elementos funcionales se pueden modificar o combinar para incrementar la utilidad o expresión de las secuencias de la invención para cualquier aplicación particular. Los expertos en la técnica están familiarizados con los materiales fuente estándar que describen condiciones específicas y procedimientos para la construcción, manipulación y aislamiento de macromoléculas (por ejemplo, moléculas de ADN, plásmidos, etc.), así como la generación de organismos recombinantes y el cribado y aislamiento de moléculas de ADN. Como se usa aquí, la expresión "porcentaje de identidad de secuencia" se refiere al porcentaje de nucleótidos idénticos entre los segmentos de una ventana de ADN óptimamente alineado. El alineamiento óptimo de secuencias para alinear una ventana de comparación es bien conocido por los expertos en la técnica, y se puede llevar a cabo mediante herramientas tales como el algoritmo de homología local de Smith y Waterman (Waterman, M. S. *Introduction to Computational Biology: Maps, sequences and genomes*. Chapman & Hall. Londres (1995), el algoritmo de alineamiento de homología de Needleman y Wunsch (*J. Mol. Biol.*, 48:443-453 (1970), la búsqueda de método de similitud de Pearson y Lipman (*Proc. Natl. Acad. Sci.*, 85:2444 (1988), y preferiblemente mediante implementaciones por ordenador de estos algoritmos tales como GAP, BESTFIT, FASTA, y TFASTA disponibles como parte del GCG (Marca Registrada), Wisconsin Package (Marca Registrada de Accelrys Inc., San Diego, Calif.). Una "fracción de identidad" para segmentos alineados de una secuencia de ensayo y una secuencia de referencia es el número de componentes idénticos que están compartidos por las dos secuencias alineadas, dividida entre el número total de componentes en el segmento de la secuencia de referencia, es decir, toda la secuencia de referencia o una parte definida más pequeña de la secuencia de referencia. El porcentaje de identidad de secuencia se representa como la fracción de identidad multiplicada por 100. La comparación de una o más secuencias de ADN puede ser con respecto a una secuencia de ADN de longitud completa o una porción de la misma, o con respecto a una secuencia de ADN más larga.

Los promotores de la presente invención pueden estar enlazados operablemente a una secuencia de ácido nucleico que es heteróloga con respecto al promotor. La secuencia de ácido nucleico puede ser generalmente cualquier secuencia de ácido nucleico para la que se desee un nivel incrementado o un nivel alterado (por ejemplo en un órgano diferente) de transcripción. Por ejemplo, la secuencia de ácido nucleico puede codificar un polipéptido que es adecuado para la incorporación en la dieta de un ser humano o un animal, o puede proporcionar cierto rasgo agrícola o industrial importante. Las secuencias de ácidos nucleicos heterólogas adecuadas incluyen, sin limitación, aquellas que codifican proteínas de almacenamiento de las semillas, enzimas de la ruta de ácidos grasos, epoxidasas, hidroxilasas, citocromo P450 monooxigenasas, desaturadas, enzimas biosintéticas de tocoferol, enzimas de la biosíntesis de carotenoides, enzimas biosintéticas de aminoácidos, enzimas de la ruta de esteroides, y enzimas de ramificación del almidón.

También se describe un vector, en particular un vector recombinante que comprende un casete de expresión de la invención. Un "vector recombinante" se refiere a cualquier agente tal como un plásmido, cósmido, virus, secuencia que se replica de forma autónoma, fago, o secuencia nucleotídica de ADN o ARN monocatenaria lineal, monocatenaria circular, bicatenaria lineal, o bicatenaria circular. El vector recombinante puede derivar de cualquier fuente, y es capaz de la integración genómica o la replicación autónoma. De este modo, en un vector recombinante se puede proporcionar cualquiera de los promotores y secuencias de ácidos nucleicos heterólogas descritos anteriormente. Un vector recombinante comprende típicamente, en una orientación desde 5' hasta 3': un promotor para dirigir la transcripción de una secuencia de ácido nucleico, y una secuencia de ácido nucleico. El vector recombinante puede comprender además un terminador transcripcional en 3', y una señal de poliadenilación en 3', otras secuencias de ácidos nucleicos no traducidas, secuencias de ácidos nucleicos de tránsito y seleccionadoras de dianas, marcadores seleccionables, potenciadores, y operadores, según se desee. La expresión "5' UTR" se refiere a la región no traducida de ADN en dirección 5', o 5' de la región codificante de un gen, y "3' UTR" se refiere a la región no traducida de ADN en dirección 3', o 3' de la región codificante de un gen. Los medios para preparar

vectores recombinantes son bien conocidos en la técnica. Los métodos para obtener vectores recombinantes particularmente adecuados para la transformación vegetal se describen en los documentos US4971908, US4940835, US4769061 y US4757011. Los vectores típicos, útiles para la expresión de ácidos nucleicos en plantas superiores, son bien conocidos en la técnica, e incluyen vectores derivados de un plásmido inductor de tumores (Ti) de *Agrobacterium tumefaciens*. En el vector recombinante, también se pueden proporcionar uno o más promotores adicionales. Estos promotores pueden estar enlazados operablemente, por ejemplo, sin limitación, a cualquiera de las secuencias de ácidos nucleicos descritas anteriormente. Como alternativa, los promotores pueden estar enlazados operablemente a otras secuencias de ácidos nucleicos, tales como aquellas que codifican péptidos de tránsito, proteínas marcadoras seleccionables, o secuencias antisentido. Estos promotores adicionales se pueden seleccionar en base al tipo celular en el que se insertará el vector. También, los promotores con función en bacterias, levaduras y plantas son todos bien conocidos en la técnica. Los promotores adicionales también se pueden seleccionar en base a sus rasgos reguladores. Los ejemplos de tales rasgos incluyen potenciación de la actividad transcripcional, inducibilidad, especificidad tisular, y especificidad por la etapa de desarrollo. Los promotores funcionales vegetales útiles para la expresión específica en semillas incluyen aquellos procedentes de proteínas de almacenamiento vegetales y procedentes de proteínas implicadas en la biosíntesis de ácidos grasos en oleaginosas. Los ejemplos de tales promotores incluyen las regiones reguladoras de 5' procedentes de secuencias de ácidos nucleicos estructurales tales como napina, faseolina, zeína, inhibidor de tripsina de haba de soja, ACP, estearoil-ACP desaturasa, y oleosina. La regulación específica de semillas se explica adicionalmente en el documento EP0255378. Los promotores adicionales particularmente preferidos en el vector recombinante incluyen los promotores de nopalina sintasa (nos), manopina sintasa (mas), y octopina sintasa (ocs), que son portados en plásmidos de *Agrobacterium tumefaciens* inductores de tumores; los promotores del virus del mosaico de la coliflor (CaMV) 19S y 35S; el promotor mejorado CaMV 35S; el promotor del virus del mosaico de la escrofularia (FMV) 35S; el promotor inducible por la luz procedente de la subunidad pequeña de ribulosa-1,5-bifosfato carboxilasa (ssRUBISCO); el promotor EIF-4A del tabaco. Un promotor adicional es preferiblemente selectivo de semillas, selectivo de tejidos, constitutivo, o inducible.

El vector recombinante también puede contener una o más secuencias de ácidos nucleicos adicionales. Estas secuencias de ácidos nucleicos adicionales pueden ser generalmente cualesquiera secuencias adecuadas para uso en un vector recombinante. Tales secuencias de ácidos nucleicos incluyen, sin limitación, cualquiera de las secuencias de ácidos nucleicos, y sus formas modificadas, descritas anteriormente. Las secuencias de ácidos nucleicos estructurales adicionales también pueden estar enlazadas operablemente a cualquiera de los promotores descritos anteriormente. La una o más secuencias de ácidos nucleicos estructurales pueden estar cada una enlazada operablemente a promotores distintos. Como alternativa, las secuencias de ácidos nucleicos estructurales pueden estar enlazadas operablemente a un único promotor (es decir, un único operón).

La presente invención también describe plantas transgénicas y células hospedantes transformadas que comprenden un promotor enlazado operablemente a una secuencia de ácido nucleico heteróloga. También se pueden introducir otras secuencias de ácidos nucleicos en la planta o célula hospedante, junto con el promotor y la secuencia de ácido nucleico estructural. Estas otras secuencias pueden incluir terminadores transcripcionales de 3', señales de poliadenilación en 3', otras secuencias de ácidos nucleicos no traducidas, secuencias de tránsito o seleccionadoras de dianas, marcadores seleccionables, potenciadores, y operadores. Las secuencias de ácidos nucleicos preferidas de la presente invención, que incluyen vectores recombinantes, secuencias de ácidos nucleicos estructurales, promotores, y otros elementos reguladores, se describen anteriormente.

El término "transformación" se refiere aquí a la introducción (o transferencia) de ácido nucleico a un hospedante receptor, tal como una planta o cualesquiera partes o tejidos de la planta, incluyendo células vegetales, protoplastos, callos, raíces, tubérculos, semillas, tallos, hojas, plántulas, embriones, polen y microsporas. Las plantas que contienen la secuencia de ácido nucleico transformada se denominan "plantas transgénicas". Transformada, transgénica y recombinante se refieren a un organismo hospedante, tal como una planta, en el que se ha introducido una molécula de ácido nucleico heteróloga (por ejemplo, un casete de expresión o un vector recombinante). El ácido nucleico se puede integrar de forma estable en el genoma de la planta.

Como se usa aquí, la frase "planta transgénica" se refiere a una planta que tiene un ácido nucleico introducido, introducido de forma estable en un genoma de la planta, por ejemplo los genomas nuclear o de plastidio.

Uno o más componentes de una planta, célula u organismo se puede comparar con una planta, célula u organismo que tiene un "antecedente genético similar". Un "antecedente genético similar" puede ser un antecedente en el que los organismos que se comparan comparten alrededor de 50% o más de su material genético nuclear. Un antecedente genético similar puede ser también un antecedente en el que los organismos que se comparan comparten alrededor de 75% o más, o alrededor de 90% o más de su material genético nuclear. Un antecedente genético similar puede ser también un antecedente en el que los organismos que se comparan son plantas, y las plantas son isogénicas excepto por cualquier material genético introducido originalmente usando técnicas de transformación de plantas.

Una célula hospedante transformada puede ser generalmente cualquier célula que sea compatible con la presente invención. Una planta o célula hospedante transformada puede ser o puede derivar de una planta monocotiledonea o una planta dicotiledonea, incluyendo, pero sin limitarse a, cánola, maíz, mostaza, haba de ricino, sésamo, algodón,



linaza, haba de soja, *Arabidopsis*, *Phaseolus*, cacahuete, alfalfa, trigo, arroz, avena, sorgo, colza, centeno, caña de azúcar, alazor, palmas de aceite, cáñamo, girasol, *Brassica campestris*, *Brassica napus*, *Brassica juncea*, *Crambe abyssinica*.

La planta o célula puede derivar de cánola. La planta o célula también puede derivar de *Brassica napus*.

- 5 A lo largo de la descripción y ejemplos, se hace referencia a las siguientes secuencias, representadas en el listado de secuencias:
- SEC ID NO: 1: secuencia nucleotídica del promotor del gen CYP704A1 de *Arabidopsis thaliana*
  - SEC ID NO: 2: secuencia nucleotídica del promotor del gen CYP704A2 de *Arabidopsis thaliana*
  - SEC ID NO: 3: secuencia de ARNm del gen CYP704A2 de *Arabidopsis thaliana*
  - 10 SEC ID NO: 4: secuencia de aminoácidos del gen CYP704A2 de *Arabidopsis thaliana*
  - SEC ID NO: 5: cebador directo para la amplificación de SEC ID NO: 1
  - SEC ID NO: 6: cebador inverso para la amplificación de SEC ID NO: 1
  - SEC ID NO: 7: primer cebador directo para la amplificación de SEC ID NO: 3
  - SEC ID NO: 8: primer cebador inverso para la amplificación de SEC ID NO: 3
  - 15 SEC ID NO: 9: segundo cebador directo para la amplificación de SEC ID NO: 3
  - SEC ID NO: 10: segundo cebador inverso para la amplificación de SEC ID NO: 3
  - SEC ID NO: 11: cebador directo para la amplificación de SEC ID NO: 2
  - SEC ID NO: 12: cebador inverso para la amplificación de SEC ID NO: 2
  - SEC ID NO: 13: promotor-3 de *Brassica napus* (CYP704Bn-3)
  - 20 SEC ID NO: 14: promotor-5 de *Brassica napus* (CYP704Bn-5)

## Ejemplos

### Materiales y métodos generales

25 Excepto que se indique de otro modo, los productos químicos y reactivos en los ejemplos se obtuvieron de Sigma Chemical Company, las endonucleasas de restricción procedieron de Fermentas o Roche-Boehringer, y otras enzimas modificantes o kits con respecto a ensayos bioquímicos o de biología molecular procedieron de Qiagen, Invitrogen y Q-BIOgene. Las cepas bacterianas procedieron de Invitrogen. Las etapas de clonación llevadas a cabo para los fines de la presente invención, tales como, por ejemplo, escisiones mediante restricción, electroforesis en gel de agarosa, purificación de fragmentos de ADN, enlazamiento de fragmentos de ADN, transformación de células de *E. coli*, crecimiento de bacterias, multiplicación de fagos, y análisis de secuencias de ADN recombinante, se

30 llevan a cabo como se describe por Sambrook (1989). La secuenciación de secuencias de ADN recombinante se lleva a cabo usando un secuenciador de ADN de fluorescencia por láser ABI, siguiendo el método de Sanger. Para aislar fragmentos de una molécula de ADN descrita aquí, se puede usar cualquier número de métodos bien conocidos por los expertos en la técnica. Por ejemplo, se puede usar la tecnología de la PCR (reacción en cadena de la polimerasa), para amplificar regiones de flanco a partir de una librería genómica de una planta usando la

35 información de secuencias públicamente disponible. Para amplificar secuencias de ADN desconocidas, adyacentes a una región central de secuencia conocida, se conoce por los expertos en la técnica un gran número de métodos. Los métodos incluyen, pero no se limitan a, PCR inversa, PCR vectorette, PCR con forma de Y, y enfoques de desplazamiento sobre el cromosoma. Los fragmentos de moléculas de ADN también se pueden obtener mediante otras técnicas, tales como sintetizando directamente el fragmento por medios químicos, como se realiza

40 habitualmente usando un sintetizador de oligonucleótidos automatizado. Para la presente invención, las moléculas de ADN se aislaron diseñando cebadores de PCR basados en información de secuencias disponible. Los materiales y métodos estándar para el trabajo molecular vegetal se describen en Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado juntamente por BIOS Scientific Publications Ltd. (UK) y Blackwell Scientific Publications.

Los productos químicos – NADPH procedió de Sigma (Saint Louis, MO).

45 Material vegetal y preparación microsómica – Tras la esterilización, se hicieron crecer semillas de *Arabidopsis* (ecotipo Col-0) sobre medio de Murashige y Skoog (medio MS 4,2 g/l, sacarosa 10 g/l, pastagar B 8 g/l, mioinositol 100 mg/l, tiamina 10 mg/l, ácido nicotínico 1 mg/l y piridoxina 1 mg/l - pH final 5,7) durante cinco semanas. Las plantas de *Arabidopsis* (aproximadamente 10 g) se homogeneizaron con un mortero y una mano de almirez en 50 ml de tampón de extracción (250 mM de tricina, 50 mM de NaHSO<sub>3</sub>, 5 g/l de seroalbúmina bovina, 2 mM de EDTA, 100

mM de ácido ascórbico y 2 mM de ditioneitol - pH final 8,2). El homogenado se filtró a través de una tela de filtración de nailon de 50 µm y se centrifugó durante 10 min. a 10.000 g. El sobrenadante resultante se centrifugó durante 1 h a 100.000 g. El sobrenadante (citosol) se almacenó directamente a -30°C, y el pelete microsomial se resuspendió en el tampón a pH 8,2 (50 mM de NaCl, 100 mM de tricina, 250 mM de sacarosa, 2 mM de EDTA y 2 mM de ditioneitol), con un homogeneizador Potter-Elvehjem, y se almacenó a -30°C. Todos los procedimientos para la preparación microsomial se llevaron a cabo a 0-4°C.

1. Cuantificación mediante PCR de la expresión de CYP704A1 (según la invención) y CYP704A2 (para referencia solamente) en *Arabidopsis thaliana*

Se extrajo ARN total de tallos, silicuas, hojas, raíces, semillas secas, flores maduras, pétalos, yemas de flores, semillas inmaduras y semillas durante germinación de plantas de *A. thaliana* de tipo salvaje Col-0, purificadas usando el kit "Total RNA" de productos Macherey-Nagel. La síntesis de ADNc mediante RT-PCR se llevó a cabo usando el kit de SuperScript™ III Reverse Transcriptase de Invitrogen™, siguiendo el protocolo de Invitrogen. De forma breve, se calentaron 13 µl de agua que contiene 2 µg de ARNm extraído total, 500 µM de mezcla de dNTP a 65°C durante 5 min. Después, en un volumen final de 20 µl, se añadieron 200 ng de cebadores al azar, 50 µM de DTT, 40 U de inhibidor de ARNasa recombinante RNase OUT™, 200 U de Superscript™ III RT, 1x de concentración final de tampón de primera hebra. La mezcla se incubó entonces durante 30-60 minutos a 50°C. Se llevó a cabo la Q-PCR usando pares de cebadores específicos, que permiten la hibridación específica a los ADNc de CYP704A1 (según la invención) y CYP704A2 (para referencia solamente). Las secuencias nucleotídicas de los cebadores son, para CYP704A1: Directo: 5'-GGTATACTCCTGTACACGCCACAA-3' (SEC ID NO: 5, desemparejamiento del 29% con la secuencia nucleotídica de CYP704A2), e Inverso: 5'-CCTGAAAGTAGGCTTTGTCCTC-3' (SEC ID NO: 6, desemparejamiento del 27% con la secuencia nucleotídica de CYP704A2). Para CYP704A2, se seleccionaron dos conjuntos de cebadores. El primero hibrida una región con homología a CYP704A1: directo: 5'-CAGTAGTAGAGAGAATATGGC-3' (SEC ID NO: 7, desemparejamiento del 28% con la secuencia nucleotídica de CYP704A1), e inverso: 5'-CTCTAGTAGAGAACTCAAAGC-3' 3' (SEC ID NO: 8, desemparejamiento del 7% con la secuencia nucleotídica de CYP704A1). El segundo conjunto de cebadores, usado como control, es más selectivo, puesto que el cebador directo se hibrida en los primeros 18 nucleótidos de CYP704A2. Esos primeros nucleótidos no tienen identidad nucleotídica en los dos genes. El cebador directo del conjunto 2 es 5'-ATGGAGATTTTGACGAGCATAGC-3' (SEC ID NO: 9, desemparejamiento del 78% con la secuencia nucleotídica de CYP704A1), y la secuencia inversa: 5'-CATAAGATAGATAGTGAAACACAAAACG-3' (SEC ID NO: 10, desemparejamiento del 25% con la secuencia nucleotídica de CYP704A1). El software usado para diseñar las secuencias oligonucleotídicas fue Primer Express. El programa de Q-PCR usado para estos experimentos fue 2 min. a 50°C, 10 min. a 95°C, y 40 ciclos de 15 s a 95°C, y 1 min. a 60°C. Las reacciones de q-PCR se llevaron a cabo al menos por triplicado con 1 µl de ADNc para cada muestra de órganos, y con CyberGreen como sonda fluorescente. Para la normalización de datos, se usó como control interno ACTIN 2 de *Arabidopsis thaliana* (locus At5g09810 de GenBank). Los niveles relativos de transcrito para cada gen se expresaron como relaciones con respecto a niveles de transcrito normalizados con respecto a ACTIN2 en cada órgano.

Debido a la fuerte identidad de secuencia entre CYP704A1 y CYP704A2, se diseñaron sondas de q-PCR para maximizar la especificidad de la hibridación. El análisis de las curvas de fusión obtenidas para las parejas seleccionadas de cebadores indicó una fuerte especificidad de la hibridación, como se indica por un único punto de fusión. Se detectó en semillas una fuerte expresión tanto de CYP704A1 como de CYP704A2 (respectivamente Fig. 1 y Fig. 2). Comparando los niveles de los transcritos en los diferentes órganos, CYP704A1 pareció estar expresado específicamente en las silicuas durante el desarrollo de las semillas y en las semillas secas maduras. La expresión en los otros órganos ensayados fue muy baja. CYP704A2 es específica de semillas, y la expresión más elevada de CYP704A2 se produce más bien al final del desarrollo de las semillas. Para confirmar estos datos, se repitió el experimento con el segundo conjunto de cebadores, más específicos para CYP704A2. Se obtuvieron resultados similares, como se puede ver en la Figura 3. De este modo, parece que los conjuntos de sondas son específicos para los genes CYP704A1 y CYP704A2.

CYP704A2 muestra la expresión más restringida en semillas maduras.

2. Generación de un casete de expresión que comprende un promotor de *Arabidopsis thaliana* específico de semillas (para referencia solamente)

Para aislar un fragmento promotor de CYP704A2 (representado en SEC ID NO: 2), se aisló ADN genómico de *Arabidopsis thaliana* (ecotipo Columbia), como se describe (Galbiati M et al (2000) *Funct. Integr. Genomics* 1(1):25-34). El ADN genómico aislado se empleó como ADN de matriz para una amplificación mediada por una reacción en cadena de la polimerasa (PCR) usando los cebadores oligonucleotídicos y protocolos indicados más abajo. Los cebadores usados para la amplificación comprendieron, respectivamente, sitios de restricción BamHI e HindIII.

La amplificación se lleva a cabo según lo siguiente:

10 ng de ADN genómico de *Arabidopsis thaliana*

1X tampón de PCR

1,5 mM de MgCl<sub>2</sub>,

200 μM de cada uno de dATP, dCTP, dGTP y dTTP

10 pmoles de cada uno de los cebadores oligonucleótidos:

■ Directo:

5 5'- CGAAGCTTAAGCTTGCAATCTCTCAGATACTTG-3' (SEC ID No: 11)

■ Inverso:

5'- GGGATCCACTTCGAAGTCAACGATAGTATC-3' (SEC ID No: 12)

3,5 Unidades de ADN polimerasa Isis (Invitrogen) en un volumen final de 50 μl.

Se empleó el siguiente programa de temperatura para las diversas amplificaciones (termociclador de BIORAD).

- 10 1. 96°C durante 10 min.  
 2. 58°C durante 2 min., seguido de 72°C durante 3 min. y 96°C durante 3 min. Repetido 30 veces.  
 3. 58°C durante 1 min., seguido de 72°C durante 10 min.  
 4. Almacenamiento a 4°C.

15 La verificación de las secuencias del producto de la PCR dio como resultado la secuencia nucleotídica como se representa en SEC ID NO: 2.

20 El producto de la PCR resultante se digirió con BamHI e HindIII, y se insertó en el vector pBi101 (Clontech), el cual también se digirió con las mismas enzimas de restricción (cuyos sitios de reconocimiento están presentes en el sitio de clonación múltiple del vector). El vector pBi101 contiene un gen GUS sin promotor en dirección 3' de un sitio de clonación múltiple. En el vector pBi101 recombinante resultante, el gen GUS se puso bajo el control transcripcional de un fragmento promotor de CYP704A2. De este modo, el constructo del fragmento promotor de CYP704A2 (909 pares de bases (pb) situados en dirección 3' del codón de partida de la secuencia de CYP704A2, es decir, SEC ID NO: 3) – GUS (un gen que codifica beta-glucuronidasa) es un casete de expresión que se clonó en el vector de transformación vegetal pBi101.

25 3. Perfil de expresión del constructo de promotor CYP704A2:GUS en plantas de *Arabidopsis thaliana* transformadas de forma estable (para referencia solamente)

30 En la siguiente etapa, el vector recombinante que comprende el casete de expresión del ejemplo 2 se usó para transformar de forma estable *Arabidopsis thaliana*. El protocolo para la transformación de *A. thaliana* mediada por *Agrobacterium* fue según el método de inmersión floral descrito por Clough SJ y Bent AF (1998) Plant J 16(6): 735-43. La actividad de β-glucuronidasa se monitorizó en plantas con el sustrato cromogénico X-Gluc (ácido 5-bromo-4-cloro-3-indolil-β-D-glucurónico) durante los ensayos de actividad correspondientes (Jefferson RA et al (1987) EMBO J. 20;6(13):3901-7). Para la determinación de una actividad promotora y especificidad tisular, el tejido vegetal se disecó, se embebió, se tiñó y se analizó como se describe (por ejemplo, Baumlein H et al (1991) Mol. Gen. Genetics 225(3):459-67). De este modo, la actividad de beta-glucuronidasa en las plantas transformadas se atestiguó por la presencia del color azul debido al metabolismo enzimático del sustrato X-Gluc.

35 La tinción de GUS se llevó a cabo sobre semillas maduras disecadas después de 2 horas de imbibición de las semillas. Se observó que la expresión estaba localizada esencialmente en los cotiledones, y también en el meristemo de la raíz apical. El perfil de expresión confirma los experimentos de q-PCR del ejemplo 1.

4. Aislamiento de promotores ortólogos a partir de *Brassica napus* (para referencia solamente)

40 Se identificaron cinco marcos de lectura abiertos (ORF1, -2, -3, -4 y -5) en nuestro banco de datos de secuencias génicas de *Brassica napus* patentada, que poseen un grado elevado de identidad con la secuencia de ARNm de CYP704A2 de *A. thaliana* representada en SEC ID NO: 3. Los promotores correspondientes, que dirigen la expresión de ORF1, -2, -3, -4 y -5, se identificaron basándose en el desplazamiento genómico *in silico* en la secuencia genómica de *B. napus*, seguido del aislamiento de estos cinco promotores con clonación mediante PCR. Se obtuvieron secuencias promotoras de aproximadamente 2-3 kb (es decir, alrededor de 2000-3000 pares de bases antes del codón de partida), dando como resultado un promotor de CYP704Bn 1, 2, 3, 4 y 5. Los promotores CYP704Bn-3 (representado en SEC ID NO: 13) y CYP704Bn-5 (representado en SEC ID NO: 14) se retuvieron basándose en la activación específica de estos promotores en semillas como se atestigua mediante la aparición específica de embriones de los ORF-3 y ORF-5. De hecho, a partir de una búsqueda BLAST en las bases de datos de transcriptomas de embriones de *B. napus*, se concluyó que sólo ORF-3 y ORF-5 son expresados en el embrión de *B. napus* PPS02 (línea de ensayo de Bayer). Por lo tanto, se consideraron los fragmentos promotores

correspondientes para dirigir una expresión específica de semillas.

5 Se construyeron cuatro genes quiméricos diferentes, en los que SEC ID NO: 13 (promotor CYP704Bn-3) y SEC ID NO: 14 (promotor CYP704Bn-5) están genéticamente fusionados a respectivamente el marco de lectura abierto de GUS (beta-glucuronidasa), en el vector pTOF13, y respectivamente el marco de lectura abierto de EGFP (proteína fluorescente verde mejorada), en el vector pCO274.

En la etapa siguiente, *B. napus* se transforma genéticamente con vectores de transformación vegetal que comprenden los genes quiméricos anteriores usando métodos conocidos en la técnica. La expresión específica de semillas de los genes quiméricos se observa en semillas de plantas de *B. napus* transformadas de forma estable.

5. Caracterización de los promotores específicos de semillas de *B. napus* (para referencia solamente)

10 DoOP (<http://doop.abc.hu>, bases de datos de promotores ortólogos, Barta E. et al (2005) Nucleic Acids Research Vol. 33, D86-D90) es una herramienta bioinformática para identificar los bloques de secuencias conservados (secuencias de consenso o motivos de consenso) entre secuencias promotoras eucariotas ortólogas. Los dos promotores de *A. thaliana* (CYP704A1, CYP704A2) y los cinco promotores de *B. napus* (CYP704Bn-1, -2, -3, -4 y -5) se analizaron en busca de la presencia de secuencias de consenso promotoras (es decir, motivos de secuencias). El análisis muestra que los promotores de *B. napus* CYP704Bn-3 (SEC ID NO: 13) y CYP704Bn-5 (SEC ID NO: 14) comparten el número más elevado de motivos de secuencias con el promotor CYP704A2 de *A. thaliana* (SEC ID NO: 2). Se concluye que los promotores CYP704Bn-3 y CYP704Bn-5 de *B. napus* representan los promotores ortólogos del promotor CYP70A2 de *A. thaliana*.

15

Tabla 1: La tabla muestra la presencia de motivos de DOOP en secuencias del promotor CYP704. CYP704A1 (SEC ID NO: 1) y CYP704A2 (SEC ID NO: 2) son promotores de *A. thaliana*. CYP704Bn-1, CYP704Bn-2, CYP704Bn-3, CYP704Bn-4 y CYP704Bn-5 son secuencias promotoras de *B. napus*. El análisis muestra que los promotores CYP704Bn-3 y CYP704Bn-5 de *B. napus* comparten un número significativo de motivos de secuencias de consenso con el promotor CYP704A2 procedente de *A. thaliana*.

ID del motivo	Longitud del motivo	Consenso	CYP704Bn-3	CYP704Bn-5	CYP704A1	CYP704A2	CYP704Bn-1	CYP704Bn-2	CYP704Bn-4
m1	15	CTCTCAGATACTTGA (SEC ID NO: 15)	m1	m1		m1			
m2	7	TTTGCCA	m2	m2	m2	m2		m2	m2
m3	13	GAAATCTTCTCTC (SEC ID NO: 16)	m3	m3		m3			
m4	17	GATCTCACCTTCTCCTT (SEC ID NO: 17)	m4	m4		m4			
m6	9	GTAGCTTCA	m6			m6			
m7	10	TCATGATCCT (SEC ID NO: 18)	m7	m7		m7			
m9	11	TAATGAAGAAG (SEC ID NO: 19)	m9	m9		m9			
m10	29	TCATGTAAACCCACCTTCTGCTCTAAC (SEC ID NO: 20)	m10			m10			
m13	11	TCCTCTTTCGC (SEC ID NO: 21)	m13	m13		m13			
m14	17	AAATCGAAGTATCCTTT (SEC ID NO: 22)	m14	m14		m14			
m15	11	AGCTTATACGC (SEC ID NO: 23)	m15	m15		m15			
m16	9	CGTTCGTTA	m16	m16		m16			
m17	8	TAACCGGA	m17	m17		m17		m17	
m18	11	CTCCATCGATA (SEC ID NO: 24)	m18	m18		m18			
m19	8	TTCATGGA	m19	m19		m19	m19		
m21	7	TAGGATC	m21			m21		m21	
m22	9	TTAGATTCT	m22	m22		m22			
m23	28	GAGCTAAACGAATCAATTATACCTCTGA (SEC ID NO: 25)				m23			

(continuación)

ID del motivo	Longitud del motivo	Consenso	CYP704Bn-3	CYP704Bn-5	CYP704A1	CYP704A2	CYP704Bn-1	CYP704Bn-2	CYP704Bn-4
m24	14	GAAACTCATATCTA (SEC ID NO: 26)	m24			m24			CYP704Bn-4
m25	10	AAAAAACAAA (SEC ID NO: 27)	m25	m25		m25			
m26	18	GGTTTCACGCACCATGTT (SEC ID NO: 28)	m26			m26			
m27	13	AGAAAAGAAGGTTT (SEC ID NO: 29)	m27			m27			
m28	10	TATACATATA (SEC ID NO: 30)	m28			m28			
m29	10	GATCATCTTC (SEC ID NO:31)	m29	m29		m29	m29	m29	m29

## 6. Construcción del vector para sobreexpresión y experimentos de “supresión” génica

Se diseñan vectores usados para la expresión de “ácidos nucleicos candidatos” de longitud completa de interés en plantas para (sobre)expresar la proteína de interés de una manera preferida de semillas, y son de dos tipos generales, biolístico y binario, dependiendo del método de transformación vegetal a usar. Para la transformación biolística (con los denominados vectores biolísticos), los requisitos son los siguientes.

5

Una cadena principal con un marcador seleccionable bacteriano (típicamente un gen de resistencia a antibiótico) y origen de replicación funcional en *Escherichia coli* (por ejemplo ColEI), y (2) una porción específica de la planta que consiste en: a. un casete de expresión génica que consiste en el promotor representado en SEC ID NO: 1, o un fragmento funcional de al menos 50 bases consecutivas de SEC ID NO: 1, o una secuencia nucleotídica que tiene al menos 40% de identidad con SEC ID NO: 1, o un fragmento funcional de al menos 40% de identidad con un fragmento de al menos 50 bases consecutivas de SEC ID NO: 1 – el gen de interés (típicamente un ADNc de longitud completa) y un terminador transcripcional (por ejemplo, el terminador nos de *Agrobacterium tumefaciens*); b. un casete marcador seleccionable vegetal, que consiste en un promotor adecuado, un gen marcador seleccionable y un terminador transcripcional (por ejemplo el terminador nos).

10

15

Los vectores diseñados para la transformación mediante *Agrobacterium tumefaciens* (los denominados vectores binarios) consisten en una cadena principal con un marcador seleccionable bacteriano, funcional tanto en *E. coli* como en *A. tumefaciens* (por ejemplo resistencia a espectinomomicina mediada por el gen *aadA*), y dos orígenes de replicación, funcionales en cada uno de los hospedantes bacterianos mencionados anteriormente, más el gen *virG* de *A. tumefaciens*, (2) una porción específica de la planta como se describe para vectores biolísticos anteriores, excepto que en este caso esta porción está flanqueada por secuencias frontera derecha e izquierda de *A. tumefaciens* que median la transferencia del ADN flanqueado por estas dos secuencias a la planta.

20

25

Los vectores diseñados para reducir o eliminar la expresión de un único gen o de una familia de genes relacionados (tales como vectores de silenciamiento génico) específicamente en semillas son también de dos tipos generales, que corresponden a la metodología usada para disminuir la expresión génica: interferencia de ARN antisentido o bicatenario (dsRNAi). Para vectores antisentido, se puede usar un fragmento génico de longitud total o parcial (típicamente, una porción del ADNc) en los mismos vectores descritos para la expresión de longitud total, como parte del casete de expresión génica. Para la reducción de la expresión génica mediada por antisentido, la región codificante del gen o fragmento génico estará en la orientación opuesta con respecto al promotor, con lo que se obtendrá ARNm a partir de la hebra no codificante (antisentido) de una manera preferida de semillas en plantas.

30

35

Para vectores de dsRNAi, se usa un fragmento génico parcial en el casete de expresión génica, y se expresa tanto en orientaciones sentido como antisentido, separado por una región espaciadora (típicamente un intrón vegetal o un marcador seleccionable). Se diseñan vectores de este tipo para formar un tallo de ARNm bicatenario, que resulta del emparejamiento de bases de los dos fragmentos génicos complementarios de una manera preferida de semillas en plantas. Los vectores biolísticos o binarios diseñados para la sobreexpresión o eliminación pueden variar de muchas maneras, incluyendo, por ejemplo, los marcadores seleccionables usados en plantas y bacterias, los terminadores transcripcionales usados en la expresión génica y casetes marcadores seleccionables vegetales, y las metodologías usadas para la clonación en gen o fragmentos génicos de interés (típicamente, clonación mediada por enzimas de restricción convencional o a base de Gateway™ recombinasa).

## LISTADO DE SECUENCIAS

40

<110> Bayer BioScience N.V. Le Centre national de la Recherche Scientifique

<120> Casetes de expresión para la expresión específica de semillas en plantas

<130> BCS 08-2016

<160>31

<170> PatentIn version 3.3

45

<210> 1

<211> 1210

<212> ADN

<213> *Arabidopsis thaliana*

<400> 1

ES 2 415 761 T3

ggtgctgttt cttctatctc ctatcttggt catcgttcct ttttggctct gatcatttct 60  
 tgtgccaaga tctattgacg atattaccct ccgctctggt attcccttag ataaatcata 120  
 tgtgttgtaa attaggccca atgacatgct catgaaaatg gaaaactaaa cagagcttct 180  
 gtcacttgac gtttcagggt ttcaaaagag aaacattaat acagagcgga gggtaatfff 240  
 gtgtcgattc atatfffcca cagagcaaaa agtttccttc agcaagaatc aattcatatc 300  
 agaaatacag ataataatga gaaaacacca tttttgaaga agccgataca taaagttgaa 360  
 taaatcaatt cgatttcatt tgtcaagaac cggaactgga cccatatcaa accggtffta 420  
 ttgatcttcc tttgtcttfc cgagagcatg gggagtgcag aaaatgatga catggacata 480  
 gagtgcagaa aggagataat atctactaca tagcaatgtg taggatgact tacatttagg 540  
 gtcaagatga ggaagatttc aagccagaga gatggcttaa ggatggcgtg ttccaacctt 600  
 aagaaataca tttctaataa aacctgagat gttttcttfc tttgactatg tatgtaggct 660  
 ggtccaagaa tctgtcttgg caaagatttc gcataccggc aaatgaagac acttattcac 720  
 ttctttcgct tcaaaatggc tgatgagaag actactgtat gttactagag aatgcttact 780  
 cttcgtgtcc aaggaggaca aacacctcta aaaaactfff ttttcatatc aaataatata 840  
 tggaagtata atcaattaaa aaaaacacat ttttcccca tatcttctgt tttgtcttta 900  
 tataatttga ttacgtggta caaaagaaag ttgaggccct aattcatatt gtgcgacttt 960  
 ttcaccaca attactacat gagaaaacat caaaacgaca aaatcggtta cactcaacgt 1020  
 taaccttca ggtaaaagag agggttcaga atgtttcggt ggatgcgtgt aaaatccgtc 1080  
 ttgactfff cgtgattgtg aatcgtctat atatgacttc tttagcggaa gattgaaaaa 1140  
 atcagacaac aagacgaaac aacttccaag ccaaaaaaaa aaaagtgaat caatggagat 1200  
 tttgacgagc 1210

<210> 2

<211> 920

<212> ADN

5 <213> *Arabidopsis thaliana*

<400> 2

aagcttgcaa tctctcagat acttgatfff gccaatftga aatcttctct cggtacgatc 60



ES 2 415 761 T3

tcaccttctc cttttcacta ccatatcatt caacaaatca tcatcactac aactttat	120
ttcaacataa catatataag tcaaaacaaa cttatcaaaa cctggaacta atatacgaag	180
gagaaggagt agaagtagct tcattcatga tcctttgagc atttctataa tgaagaagag	240
catcatcatg taaacccac tctttctgctc taaccccttt agcgatttcc tctttcgcta	300
aatcgaagta tcctttaagc ttatacgcaa ttcgttcggt agtaaccgga actccatcga	360
taccattcat ggactcagaa gaagttgatg atgatgaaac actaggatct ttcttagatt	420
cttcgttgag aatagagcta aacgaatcaa ttatacctct gaggaaactc atatctaagc	480
aataaacaat aaaatcgatc atcacttccg agtttaatca aacaaaaaac aaaaaacat	540
caagaaggca aaattggttt cacgcaccat gttgtgctcg tatatataac tcatgtcgag	600
aagattcgaa tcaaaatcga aatcaaatcc agcgaattga agaaatttga tgaaatctta	660
tgaactggga aatacgcaa cgatgaacaa accaaacaac acacggtaa gtctttaacc	720
ctctacgtag aaagaaggtt ctaatgggct ttacaaggcc tatacatata taatcttagg	780
cccaatgata ttttatcggg ggcccagttt tgttttgttt tcgatcatct tcttcgccgt	840
taaatatfff attctctacg gtaccttttg tctctttgta ctttcagata ctatcgttga	900
cttcgaagta agaaaagaaa	920

<210> 3

<211> 1775

<212> ADN

5 <213> *Arabidopsis thaliana*

<220>

<221 > CDS

<222> (12)..(1544)

<400> 3

ES 2 415 761 T3

aagaaaagaa a atg gag att ttg acg agc ata gct att aca gta gca aca	50
Met Glu Ile Leu Thr Ser Ile Ala Ile Thr Val Ala Thr	
1 5 10	
acg atc ttc atc gtt ttg tgt ttc act atc tat ctt atg atc aga atc	98
Thr Ile Phe Ile Val Leu Cys Phe Thr Ile Tyr Leu Met Ile Arg Ile	
15 20 25	
ttt acc ggg aaa tca aga aac gac aag aga tat gct cca gtg cat gcc	146
Phe Thr Gly Lys Ser Arg Asn Asp Lys Arg Tyr Ala Pro Val His Ala	
30 35 40 45	
acg gtc ttt gat ctt ctc ttc cac agc gac gag tta tac gat tac gag	194
Thr Val Phe Asp Leu Leu Phe His Ser Asp Glu Leu Tyr Asp Tyr Glu	
50 55 60	
aca gag atc gcg aga gag aag ccg act tac agg ttt ttg agt cca gga	242
Thr Glu Ile Ala Arg Glu Lys Pro Thr Tyr Arg Phe Leu Ser Pro Gly	
65 70 75	
cag agc gag ata tta act gca gat cct cgt aac gtg gag cat att ctc	290
Gln Ser Glu Ile Leu Thr Ala Asp Pro Arg Asn Val Glu His Ile Leu	
80 85 90	
aag aca aga ttc gat aac tat agt aaa gga cac agt agt aga gag aat	338
Lys Thr Arg Phe Asp Asn Tyr Ser Lys Gly His Ser Ser Arg Glu Asn	
95 100 105	
atg gcg gat ctt cta gga cat ggg atc ttt gct gtt gat gga gag aaa	386
Met Ala Asp Leu Leu Gly His Gly Ile Phe Ala Val Asp Gly Glu Lys	
110 115 120 125	
tgg aga caa cag agg aag ctt tcg agc ttt gag ttc tct act aga gtt	434
Trp Arg Gln Gln Arg Lys Leu Ser Ser Phe Glu Phe Ser Thr Arg Val	
130 135 140	
tta aga gat ttt agc tgc tct gtt ttt agg aga aat gca tct aag ctt	482
Leu Arg Asp Phe Ser Cys Ser Val Phe Arg Arg Asn Ala Ser Lys Leu	
145 150 155	
gtt ggt ttt gtc tcg gag ttt gct ctc tct gga aaa gct ttt gat gct	530
Val Gly Phe Val Ser Glu Phe Ala Leu Ser Gly Lys Ala Phe Asp Ala	
160 165 170	
caa gat ttg ttg atg aga tgt aca ctg gac tcc atc ttt aaa gtt ggg	578
Gln Asp Leu Leu Met Arg Cys Thr Leu Asp Ser Ile Phe Lys Val Gly	
175 180 185	
ttt ggt gtg gag tta aaa tgt ttg gat ggg ttt agc aaa gaa ggg caa	626
Phe Gly Val Glu Leu Lys Cys Leu Asp Gly Phe Ser Lys Glu Gly Gln	
190 195 200 205	
gag ttt atg gaa gct ttt gat gaa ggt aac gtt gca act agt tcc aga	674
Glu Phe Met Glu Ala Phe Asp Glu Gly Asn Val Ala Thr Ser Ser Arg	
210 215 220	
ttc atc gat cct ctt tgg aag ctg aaa tgg ttt ttc aac att gga tca	722
Phe Ile Asp Pro Leu Trp Lys Leu Lys Trp Phe Phe Asn Ile Gly Ser	
225 230 235	
caa tct aag ctc aag aag agc att gct act ata gat aaa ttt gtc tat	770
Gln Ser Lys Leu Lys Lys Ser Ile Ala Thr Ile Asp Lys Phe Val Tyr	
240 245 250	
agt ctc att acc act aaa agg aaa gag ctt gct aag gaa cag aac act	818
Ser Leu Ile Thr Thr Lys Arg Lys Glu Leu Ala Lys Glu Gln Asn Thr	
255 260 265	

ES 2 415 761 T3

ggt gtt aga gag gac ata cta tcg aga ttc cta gtg gag agt gag aaa Val Val Arg Glu Asp Ile Leu Ser Arg Phe Leu Val Glu Ser Glu Lys 270 275 280 285	866
gat ccg gag aac atg aat gat aag tac ctg agg gat ata att ctg aac Asp Pro Glu Asn Met Asn Asp Lys Tyr Leu Arg Asp Ile Ile Leu Asn 290 295 300	914
ttc atg att gct ggt aag gac aca acc gct gca cta ctc tct tgg ttc Phe Met Ile Ala Gly Lys Asp Thr Thr Ala Ala Leu Leu Ser Trp Phe 305 310 315	962
ttg tac atg ctc tgc aaa aac cca ctt gtt cag gag aaa atc gta caa Leu Tyr Met Leu Cys Lys Asn Pro Leu Val Gln Glu Lys Ile Val Gln 320 325 330	1010
gaa atc aga gat gtg aca ttt agt cac gag aaa acg acc gat gta aat Glu Ile Arg Asp Val Thr Phe Ser His Glu Lys Thr Thr Asp Val Asn 335 340 345	1058
ggt ttc gtt gaa agt att aac gaa gag gct ctt gat gag atg cac tac Gly Phe Val Glu Ser Ile Asn Glu Glu Ala Leu Asp Glu Met His Tyr 350 355 360 365	1106
ctc cat gca gcc ttg tct gag acc ttg agg ctc tac cct cct gtg cct Leu His Ala Ala Leu Ser Glu Thr Leu Arg Leu Tyr Pro Pro Val Pro 370 375 380	1154
gtg gac atg agg tgt gca gaa aat gat gac gtt ctt cca gat gga cat Val Asp Met Arg Cys Ala Glu Asn Asp Asp Val Leu Pro Asp Gly His 385 390 395	1202
aga gtg agc aaa ggg gat aat atc tac tac ata gcc tat gca atg ggt Arg Val Ser Lys Gly Asp Asn Ile Tyr Tyr Ile Ala Tyr Ala Met Gly 400 405 410	1250
agg atg act tac att tgg ggt caa gat gct gaa gaa ttc aag cca gag Arg Met Thr Tyr Ile Trp Gly Gln Asp Ala Glu Glu Phe Lys Pro Glu 415 420 425	1298
aga tgg ctc aag gac ggc tta ttc caa ccc gaa tca cca ttc aaa ttc Arg Trp Leu Lys Asp Gly Leu Phe Gln Pro Glu Ser Pro Phe Lys Phe 430 435 440 445	1346
ata agc ttt cat gct ggt cca aga atc tgt ctt ggc aag gat ttc gca Ile Ser Phe His Ala Gly Pro Arg Ile Cys Leu Gly Lys Asp Phe Ala 450 455 460	1394
tac cgg cag atg aag ata gta tca atg gca ctt ctt cac ttc ttt cgc Tyr Arg Gln Met Lys Ile Val Ser Met Ala Leu Leu His Phe Phe Arg 465 470 475	1442
ttc aaa atg gct gat gag aac agc aag gtg tat tac aag aga atg ctt Phe Lys Met Ala Asp Glu Asn Ser Lys Val Tyr Tyr Lys Arg Met Leu 480 485 490	1490
act ctt cat gtc gat gga gga ctc cat ctc tgt gca atc ccg agg aca Thr Leu His Val Asp Gly Gly Leu His Leu Cys Ala Ile Pro Arg Thr 495 500 505	1538
agc act tgagtactta accagaatac tttccatact tttatagtgc ctccatgagt Ser Thr 510	1594
tttgtaaca ttctcaataa aggtggaaca agacaccaat aacgttccat gtaataaata	1654
atctgccttt atagatcttg ggatgactca actatcacat atattagaat ccgatgaaca	1714

ES 2 415 761 T3

cgacacgata ttgtataatt ttacagtcac ctagtacaat aaagcagata agtcgcactg 1774

t 1775

<210> 4

<211> 511

<212> PRT

5 <213> *Arabidopsis thaliana*

<400> 4

Met Glu Ile Leu Thr Ser Ile Ala Ile Thr Val Ala Thr Thr Ile Phe  
1 5 10 15

Ile Val Leu Cys Phe Thr Ile Tyr Leu Met Ile Arg Ile Phe Thr Gly  
20 25 30

Lys Ser Arg Asn Asp Lys Arg Tyr Ala Pro Val His Ala Thr Val Phe  
35 40 45

Asp Leu Leu Phe His Ser Asp Glu Leu Tyr Asp Tyr Glu Thr Glu Ile  
50 55 60

Ala Arg Glu Lys Pro Thr Tyr Arg Phe Leu Ser Pro Gly Gln Ser Glu  
65 70 75 80

Ile Leu Thr Ala Asp Pro Arg Asn Val Glu His Ile Leu Lys Thr Arg  
85 90 95

Phe Asp Asn Tyr Ser Lys Gly His Ser Ser Arg Glu Asn Met Ala Asp  
100 105 110

Leu Leu Gly His Gly Ile Phe Ala Val Asp Gly Glu Lys Trp Arg Gln  
115 120 125

Gln Arg Lys Leu Ser Ser Phe Glu Phe Ser Thr Arg Val Leu Arg Asp  
130 135 140

Phe Ser Cys Ser Val Phe Arg Arg Asn Ala Ser Lys Leu Val Gly Phe  
145 150 155 160

Val Ser Glu Phe Ala Leu Ser Gly Lys Ala Phe Asp Ala Gln Asp Leu  
165 170 175

Leu Met Arg Cys Thr Leu Asp Ser Ile Phe Lys Val Gly Phe Gly Val  
180 185 190

Glu Leu Lys Cys Leu Asp Gly Phe Ser Lys Glu Gly Gln Glu Phe Met  
195 200 205

ES 2 415 761 T3

Glu Ala Phe Asp Glu Gly Asn Val Ala Thr Ser Ser Arg Phe Ile Asp  
 210 215 220

Pro Leu Trp Lys Leu Lys Trp Phe Phe Asn Ile Gly Ser Gln Ser Lys  
 225 230 235 240

Leu Lys Lys Ser Ile Ala Thr Ile Asp Lys Phe Val Tyr Ser Leu Ile  
 245 250 255

Thr Thr Lys Arg Lys Glu Leu Ala Lys Glu Gln Asn Thr Val Val Arg  
 260 265 270

Glu Asp Ile Leu Ser Arg Phe Leu Val Glu Ser Glu Lys Asp Pro Glu  
 275 280 285

Asn Met Asn Asp Lys Tyr Leu Arg Asp Ile Ile Leu Asn Phe Met Ile  
 290 295 300

Ala Gly Lys Asp Thr Thr Ala Ala Leu Leu Ser Trp Phe Leu Tyr Met  
 305 310 315 320

Leu Cys Lys Asn Pro Leu Val Gln Glu Lys Ile Val Gln Glu Ile Arg  
 325 330 335

Asp Val Thr Phe Ser His Glu Lys Thr Thr Asp Val Asn Gly Phe Val  
 340 345 350

Glu Ser Ile Asn Glu Glu Ala Leu Asp Glu Met His Tyr Leu His Ala  
 355 360 365

Ala Leu Ser Glu Thr Leu Arg Leu Tyr Pro Pro Val Pro Val Asp Met  
 370 375 380

Arg Cys Ala Glu Asn Asp Asp Val Leu Pro Asp Gly His Arg Val Ser  
 385 390 395 400

Lys Gly Asp Asn Ile Tyr Tyr Ile Ala Tyr Ala Met Gly Arg Met Thr  
 405 410 415

Tyr Ile Trp Gly Gln Asp Ala Glu Glu Phe Lys Pro Glu Arg Trp Leu  
 420 425 430

Lys Asp Gly Leu Phe Gln Pro Glu Ser Pro Phe Lys Phe Ile Ser Phe  
 435 440 445

His Ala Gly Pro Arg Ile Cys Leu Gly Lys Asp Phe Ala Tyr Arg Gln  
 450 455 460

Met Lys Ile Val Ser Met Ala Leu Leu His Phe Phe Arg Phe Lys Met  
 465 470 475 480

Ala Asp Glu Asn Ser Lys Val Tyr Tyr Lys Arg Met Leu Thr Leu His  
 485 490 495

ES 2 415 761 T3

Val Asp Gly Gly Leu His Leu Cys Ala Ile Pro Arg Thr Ser Thr  
500 505 510

<210> 5  
<211> 24  
<212> ADN  
5 <213> cebador artificial  
<220>  
<223> Cebador de PCR  
<400> 5  
ggtatactcc tgtacagcc acaa 24  
10 <210> 6  
<211> 22  
<212> ADN  
<213> cebador artificial  
<220>  
15 <223> Cebador de PCR  
<400> 6  
cctgaaagta ggcttgtcc tc 22  
<210> 7  
<211>21  
20 <212> ADN  
<213> cebador artificial  
<220>  
<223> Cebador de PCR  
<400> 7  
25 cagtagtaga gagaatatgg c 21  
<210> 8  
<211>21  
<212> ADN  
<213> cebador artificial  
30 <220>  
<223> Cebador de PCR  
<400> 8  
ctctagtaga gaactcaaag c 21  
<210> 9  
35 <211> 23

<212> ADN  
<213> cebador artificial  
<220>  
<223> Cebador de PCR  
5 <400> 9  
atggagattt tgacgagcat agc 23  
<210> 10  
<211> 28  
<212> ADN  
10 <213> cebador artificial  
<220>  
<223> Cebador de PCR  
<400> 10  
cataagatag atagtgaaac acaaaacg 28  
15 <210> 11  
<211> 33  
<212> ADN  
<213> cebador artificial  
<220>  
20 <223> Cebador de PCR  
<400> 11  
cgaagcttaa gcttgcaatc tctcagatac ttg 33  
<210> 12  
<211> 30  
25 <212> ADN  
<213> cebador artificial  
<220>  
<223> Cebador de PCR  
<400> 12  
30 gggatccact tcgaagtcaa cgatagtatc 30  
<210> 13  
<211> 1953  
<212> ADN  
<213> *Brassica napus*  
35 <400> 13

ES 2 415 761 T3

ttgatgtcag aaacaaagca aaatgtgaac tcacttgtgc attaataaaa aaatcacctg 60  
 cacgtttacc aagagcttgc aacctctcag atacttgact ttgccattta gaaatcttct 120  
 ctctataaga tctcaccttc tccttctcac ttaccatata cacaagtcaa caccacaaat 180  
 aaaaaccaat gttcaagaaa ctatacaaga gtagattaat tatacattta tataaaatat 240  
 tttagttggt aggtttaatt ataaattatt atgctgagtt agttatatac aaaacaaaat 300  
 aaattagaaa aatttgtgaa tttgaacaaa acattattgc ttaatttaac aaaattagac 360  
 taatttagat cagattttaa ggaaaatggt tagtataaat ctgaaagaac attacttggga 420  
 actgatgtaa gaaggagaag gcgtggaggt agcttcactc atgatcctct gggcgtttct 480  
 gtaatgaaga agcgcgtcgt catgtaaacc ccactcttct gctctaacac ctttggcgat 540  
 ctctcttttc gccaaatcga agtatccttt gagcttatac gcgacacggt cgttactaac 600  
 cggagctcca tcgataacgt tcatggagga ggcagatgaa gaagaagatg atacggttagg 660  
 atcgtgatta gattctacgg tgaagattga gctacacgaa tcaattatat ctctgaagaa 720  
 actcatatct aaacaaaaaa aaaacaaaga ttccgaattt aattgaacag tggaatatac 780  
 tgaatacaaa atcgaaaatg acagaacagg tttcacgcac catgttttgg tcggatgaat 840  
 tcatgtcggga gaggaatcaa tcaaatcgaa tcagaattac cagtgaatta aaaataatcg 900  
 aaggttgaat aatattcagt ttataagcct ctatgtcggga aagaaggttt ccaactaata 960  
 tcagaagccc ttggttttgg gccttttaga gtgaagatca aattgaatac cattcatgcc 1020  
 tatatatatc cctatagaag aaagaaggtt cgaactcgtt tcatcctgcc tactagtacc 1080  
 agtactgggc attcggattt cggttccggt ccgagtgtat cggataaatc attctaccca 1140  
 ataagtattt attaaatttc ggtttggttt cgattcagtt ttgatttgggt atcaattcga 1200  
 tttcgagtaa cagtttaata tatgaatcaa atatatatac atataaaata tgtaactaat 1260  
 ctaagtagta tattggttat acacttacat attcgttcga gttattaaag tgctaatttg 1320  
 taagtattta ctatattata atataaaata ccaaatatat atatgtataa tataagagaa 1380  
 tacaatttac taccttatat tagaaagttg aaatttgttt ttcttttcga atatttgagt 1440  
 gtatattcgg ttcccggttc tggtttcggt tcgggttaatc ggatgtaaaa aaatatgaac 1500  
 cattcgagta tctgaaggtt tcggtttcca aataattcag ttccgttccg gttaagtttt 1560  
 tcggttttgg tttttttggc ccaaggctac tgcctacatt ttaatacgat tcttaatttt 1620  
 aggccttaaa tagaggaatc atatcaaaca aaaaatttca gtaaattaa aattaatgac 1680  
 atcgagtgaa tgagctggaa aaaaaagaa aaaaactctc ttcttcttct tttgaaacac 1740  
 cacgaagaaa actcaaaaag ttgaatataa tctatattca taagttaata accctctaac 1800  
 taggtgacat ttgttttaa atgggcctta catatttggc ctagttagat catcttcatc 1860  
 atggcgccgt gacatttgtt ttctacctac tgctgttgac ttctcaacgg aaagactaaa 1920  
 acagagaaga tctccacagc agaaacaaaa aaa 1953



ES 2 415 761 T3

<211> 3135

<212> ADN

<213> *Brassica napus*

<400> 14

```

aaaagaaagt aaacagaaca aggctaagca cagaccatt tagatgtcaa tgaagatgct      60
gagacgttga agaatgttgc ctgtgactcg gaagcaactg ccttggaag cattgttttt      120
ccattgccag gtggcccaa gagaagcaaa cctaacataa agaagaactt tcaattaagt      180
aagtcacagg agaagaaca gaagaaacca atcttacctc tagctggctt tcggagacca      240
gtgaacaaat ctcttcgttt cgctggtaaa ataaccatct ccagtaaagc ttgtttggct      300
ccatcaagtc cagctgcaag atagttcagg cataagtgtg aatcacactg atatagctca      360
taaacaatgg cactaactca ccaacatcat cccacttcac agatgggctt ctgtccacta      420
ttgtttgtgt gatcatctct tctaacttat catcgtacc attcccagat tctttcgcag      480
gttttggttt tgtagtagta gcatctctag ggctcctcgc cttcccaact ccacctctgg      540
gaagtgaagt cttttgtgtt gaaactcttc tattagacgc cgttgaggaa gctgaagctg      600
aagaaggaga tgatggtaca gttctctaca taatcaagaa ttcaaaaaaa gatgattcat      660
catatgaatc aatccccact acaaacatca cattctccta aagacaaata ctttattctc      720
agacactcca acacctaata caagaacgcc attgatgtca gaaacaaagc aaaccaatca      780
agtgtcaatg tgaactcact ggtgcattaa taaaaaaatc acctgcacgc ttaccaagag      840
cttgcaacct ctcaagatac tgactttgcc atttagaaat cttctctctg taagatctca      900
ccttctcctt ctcaactatac catatacaca agtcaacacc acaataaaaa accagtgttc      960
caagaaatta tacaagagta gattaattat acatttatat aatatatttt agttggtagg     1020
ttaattata aattattatg atgaaaatta tcagatttag atatacaaaa caaaataaat     1080
tagacaaatt tgtgaatttg aacaaacatt actgcttaat ttaacaaat tagactaatt     1140
tagatcaatt ttaaattact gagagaacat tacttggaac tgatgtaaga aggagaagga     1200
gtggaggtag cttcgttcat gatcctctga gcgtttctgt aatgaagaag cgcgtcgttg     1260
tgtaaaccct actcctctgc tctaaccact ttggcgatct cctctttcgc caaatcgaag     1320
tacccttga gcttatacgc aactcgttcg ttactaaccg gaactccatc gataacgttc     1380
atggaggagc cagatgaaga agatgatgat acggtatgat cgtgattaga ttcttccgtg     1440
aagattgagc taaacgaatc cgattatata tctgaagaaa ctctttcta acaaaaaaaaa     1500
acaagattc cgaatttaat tgaacagtgg tgatactcaa aatcaaaaac gacagaacag     1560
gtttcacgca ccatgaattg gtcggatgaa ttcatgtcgg agaggaatca aatcgaatca     1620
gaaataatcc aagtgactgg aaacgaagaa aaagaaaaga cgcaaggttg aataatattc     1680
atagtttata agcctctagg tcggaaagag ggtttccaac tgatatcaga aggccttggt     1740

```

ES 2 415 761 T3

tttgggcctt cgagagaggg atcaaattgg tttggttcgg ttccggataa gaaccattcg 1800  
 gtttcgagta tatcggataa atcattctta tacccaatag gtacttatga aatttcgatc 1860  
 cggtttcgat tcagtttttg atttggtttc gatcacttac atattcgtca atccaacat 1920  
 agttcgagtt actaaaatgc taatttgtaa gtatttacta tattttaata tacaatatat 1980  
 atatatatat atatatatat atatatatat atatatatat ataaataata taagagagta 2040  
 caatttacta tcctatatcc atgttcaaaa ttttgctagg cgctacgctg gcggttgatc 2100  
 tgggcctagc gccgaatcaa ttaatcggag attaatcgac gattaatcgg agattagttt 2160  
 tctattatth tattttacata tatgcacata tattaataa ttacaacata aacgtaaadc 2220  
 ttagttggta cagtggata cccatagta ttaaaagttc ccaacgtggg ttcgaagccc 2280  
 atcttttgat ttttttgctt taaatcgttt tttgtttaat gagttggaaa ttaccaatth 2340  
 acccccgctt tttaccttt taacctgca atttcagacc taattttcag cgtaaatgct 2400  
 atttttcatt gattatcatt gttcttctt gctattcacg tagctttcaa tcacttatga 2460  
 tagatcaatc actaaacatt cgattttgaa cactgaaaat tcaaaaaagg tggagctccg 2520  
 agtttttggc cgaatatacc aaaatcgcc cggtactca ccggatagcg tctttccgcc 2580  
 gcgtactcgg ccgcctgagc cgcgttttcg aacacggccc tatattagaa agttgaaatt 2640  
 tgtttttttg gatatttgag tacatgttcg gttttcggtt cggttaatca tatataaaaa 2700  
 tataggaacc atcgtgtatt tgtgggttc gatctagatc ttgtttccgg ataattcgg 2760  
 tcggtttttc ggtttcggtt ttttactgc ctacatttta atatgattct taattttagg 2820  
 cctatttctt aaacagagga atcatatcaa acaaaaaaat tcagtaaatt aaaaattaat 2880  
 gacatcgagt gaatcaactg gaaaagaaga aaactctctt cttctttttt ggaacaccac 2940  
 gaagaaaacc caaaagttga atatcatcta ttcatatggt aatataaacc ctctaactag 3000  
 gtgacatttg tttataatgg gccttaatat tcggcctaatt tagatcatct tcatcatggc 3060  
 gccgtgacat ttgttttcta cctactgagg ttgacttctc aacagaaaga ctaaaaacag 3120  
 agaaaatctc aaaaa 3135

<210> 15

<211> 15

<212> ADN

5 <213> secuencia artificial

<220>

<223> motivo m1

<400> 15

ctctcagata cttga 15

10 <210> 16

<211> 13

<212> ADN

<213> secuencia artificial  
<220>  
<223> motivo m3  
<400> 16  
5 gaaatcttct ctc 13  
<210> 17  
<211> 17  
<212> ADN  
<213> secuencia artificial  
10 <220>  
<223> motivo m4  
<400> 17  
gatctcacct tctcctt 17  
<210> 18  
15 <211> 10  
<212> ADN  
<213> secuencia artificial  
<220>  
<223> motivo m7  
20 <400> 18  
tcatgatcct 10  
<210> 19  
<211> 11  
<212> ADN  
25 <213> secuencia artificial  
<220>  
<223> motivo m9  
<400> 19  
taatgaagaa g 11  
30 <210> 20  
<211> 29  
<212> ADN  
<213> secuencia artificial  
<220>  
35 <223> motivo m10  
<400> 20  
tcatgtaaac ccactcttc tgctctaac 29

<210>21  
<211> 11  
<212> ADN  
<213> secuencia artificial  
5 <220>  
<223> motivo m13  
<400> 21  
tcctctttcg c 11  
<210> 22  
10 <211> 17  
<212> ADN  
<213> secuencia artificial  
<220>  
<223> motivo m14  
15 <400> 22  
aaatcgaagt atcctt 17  
<210> 23  
<211> 11  
<212> ADN  
20 <213> secuencia artificial  
<220>  
<223> motivo m15  
<400> 23  
agcttatacg c 11  
25 <210> 24  
<211> 11  
<212> ADN  
<213> secuencia artificial  
<220>  
30 <223> motivo m18  
<400> 24  
ctccatcgat a 11  
<210> 25  
<211> 28  
35 <212> ADN  
<213> secuencia artificial  
<220>

<223> motivo m23  
<400> 25  
gagctaaacg aatcaattat acctctga 28  
<210> 26  
5 <211> 14  
<212> ADN  
<213> secuencia artificial  
<220>  
<223> motivo m24  
10 <400> 26  
gaaactcata tcta 14  
<210> 27  
<211> 10  
<212> ADN  
15 <213> secuencia artificial  
<220>  
<223> motivo m25  
<400> 27  
aaaaa caaa 10  
20 <210> 28  
<211> 18  
<212> ADN  
<213> secuencia artificial  
<220>  
25 <223> motivo m26  
<400> 28  
ggttcacgc accatgtt 18  
<210> 29  
<211> 13  
30 <212> ADN  
<213> secuencia artificial  
<220>  
<223> motivo m27  
<400> 29  
35 agaaagaagg ttc 13  
<210> 30  
<211> 10

# ES 2 415 761 T3

<212> ADN

<213> secuencia artificial

<220>

<223> motivo m28

5 <400> 30

tatacatata 10

<210>31

<211> 10

<212> ADN

10 <213> secuencia artificial

<220>

<223> motivo m29

<400> 31

gatcatcttc 10

**REIVINDICACIONES**

- 5 1. Un método para regular la expresión específica de semillas en plantas, comprendiendo dicho método (a) proporcionar un casete de expresión que comprende un promotor enlazado operablemente a un ácido nucleico que es heterólogo con relación a dicho promotor, en el que dicho promotor tiene la secuencia nucleotídica de SEC ID NO: 1 o una secuencia nucleotídica que tiene al menos 95% de identidad de secuencia con SEC ID NO: 1, y (b) transformar una planta con dicho casete de expresión.
2. El método de la reivindicación 1, en el que la expresión del ácido nucleico da como resultado la expresión de una proteína, o expresión de un ARN antisentido, ARN sentido o ARN bicatenario.
3. El método de la reivindicación 1 ó 2, en el que dicha planta es una planta usada para la producción de aceite.
- 10 4. El método de la reivindicación 3, en el que dicha planta se selecciona del grupo que consiste en cáñamo, maíz, mostaza, haba de ricino, sésamo, algodón, linaza, haba de soja, *Arabidopsis*, *Phaseolus*, cacahuete, alfalfa, trigo, arroz, avena, sorgo, colza, centeno, caña de azúcar, alazor, palmas de aceite, cáñamo, girasol, *Brassica campestris*, *Brassica napus*, *Brassica juncea*, *Crambe abyssinica*.
- 15 5. El método según una cualquiera de las reivindicaciones 1-4, que comprende además las etapas de (c) hacer crecer una planta que contiene dicho casete de expresión, en el que dicha planta produce semillas, y dicho ácido nucleico es transcrito en dichas semillas, y (d) aislar dichas semillas de dicha planta transformada.
6. Uso del casete de expresión como se describe en la reivindicación 1 para regular la expresión específica de semillas de un gen heterólogo en plantas.

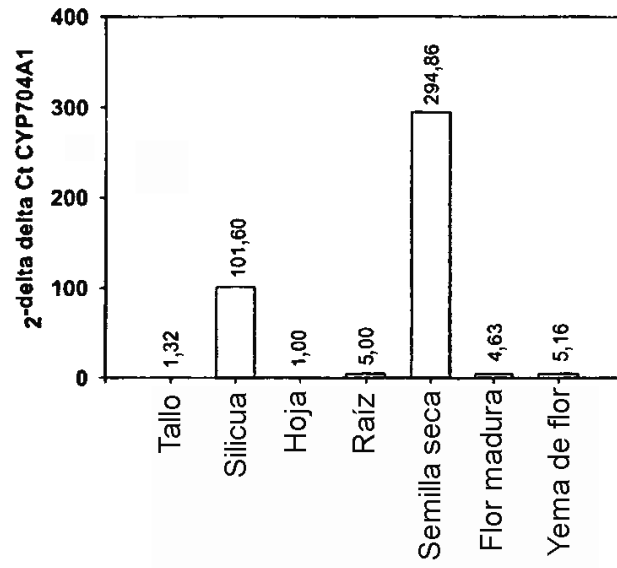


Figura 1

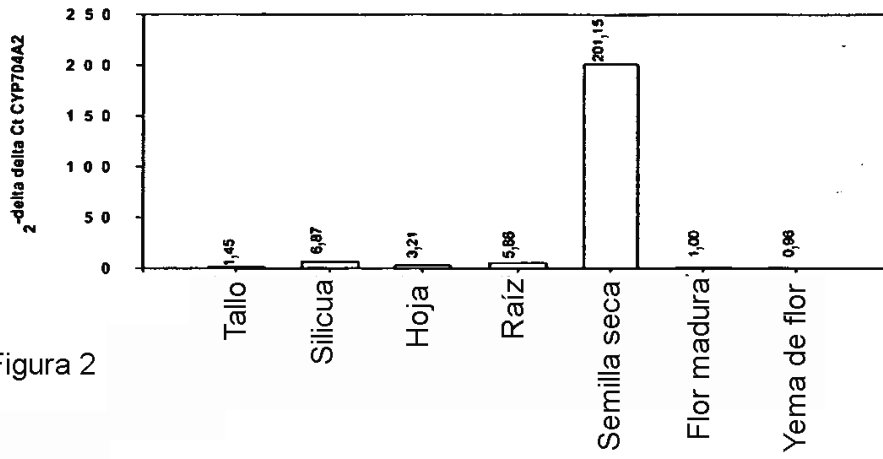


Figura 2



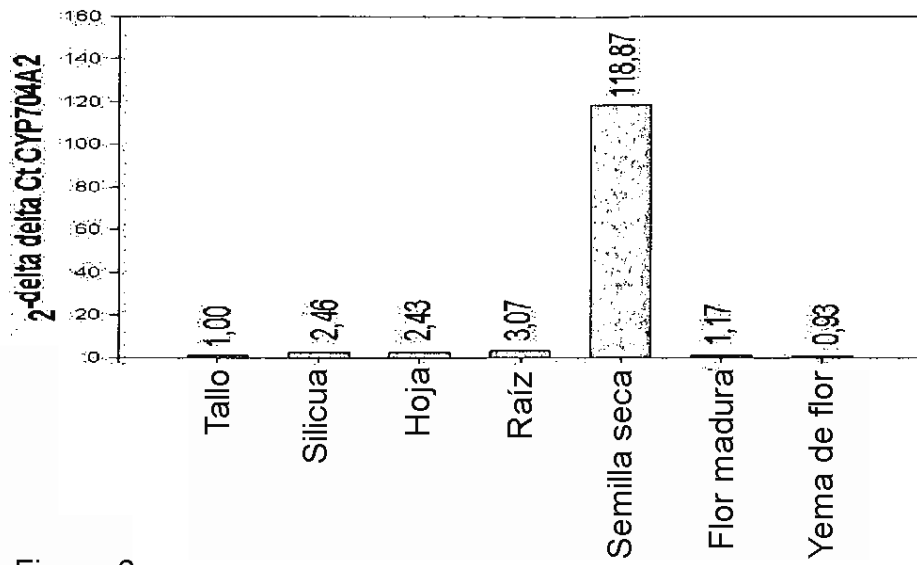


Figura 3