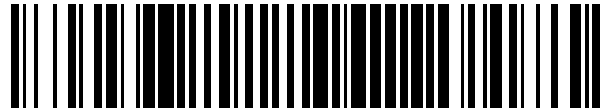


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 400 907**

51 Int. Cl.:

**A01H 5/00** (2006.01)

**C12N 15/82** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.12.2003 E 03811675 (2)**

97 Fecha y número de publicación de la concesión europea: **05.12.2012 EP 1571901**

54 Título: **Generación de plantas con contenido alterado de aceite**

30 Prioridad:

**18.12.2002 US 434763 P**

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

**15.04.2013**

73 Titular/es:

**AGRIGENETICS, INC. (100.0%)  
9330 ZIONSVILLE ROAD  
INDIANAPOLIS, IN 46268-1053, US**

72 Inventor/es:

**LIGHTNER, JONATHAN y  
CLENDENNEN, STEPHANIE, K.**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 400 907 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

Generación de plantas con contenido alterado de aceite.

## Antecedentes de la invención

5 La capacidad para manipular la composición de semillas para cultivo, particularmente el contenido y la composición de los aceites de semillas, presenta importantes aplicaciones en las industrias agrícolas, tanto en cuanto a los aceites alimenticios procesados como en cuanto a los aceites para alimentación de animales. Las semillas de cultivos agrícolas contienen una diversidad de componentes valiosos, incluyendo aceite, proteína y almidón. El procesamiento industrial permite separar algunos de estos componentes, o todos, para la venta individual en aplicaciones específicas. Por ejemplo, casi el 60% de la cosecha de soja de EE.UU. es machacada por la industria de procesamiento de soja. El procesamiento de la soja produce aceite purificado, que es vendido a precio elevado, mientras que el resto es vendido principalmente para pienso para ganado a menor precio (US Soybean Board, 2001 Soy Stats). La semilla de canola es machacada para producir aceite y la harina de canola como coproducto (Canola Council of Canada). Casi el 20% de la cosecha 1999/2000 de maíz de EE.UU. fue industrialmente refinada, principalmente para la producción de almidón, etanol y aceite (Corn Refiners Association). Por lo tanto, a menudo es deseable maximizar el contenido de aceite de las semillas. Por ejemplo, para semillas oleaginosas procesadas tales como las de soja y canola, un aumento del contenido absoluto de aceite de la semilla aumentará el precio de dichos granos. Para el maíz procesado, se puede desear aumentar o disminuir el contenido de aceite dependiendo de la utilización de otros componentes principales. La disminución de aceite puede mejorar la calidad del almidón aislado al reducirse los indeseados sabores asociados con la oxidación del aceite. Alternativamente, en la producción de etanol, donde el sabor no importa, el aumento del contenido de aceite puede aumentar el valor global. En muchos granos para pienso, tales como los de maíz y trigo, es deseable aumentar el contenido de aceite de las semillas porque el aceite tiene mayor contenido energético que otros componentes de las semillas, tales como los hidratos de carbono. El procesamiento de semillas oleaginosas, como la mayoría de los negocios de procesamiento de granos, es un negocio que requiere grandes inversiones y recursos; por ello, pequeños cambios en la distribución de los productos, de componentes de bajo valor a componente oleoso de alto valor, pueden producir sustanciales impactos económicos en las empresas procesadoras de granos.

La manipulación biotecnológica de aceites puede proporcionar una alteración de la composición y una mejora de la producción de aceite. Las alteraciones de la composición incluyen, entre otras, aceites de soja y maíz con alto contenido oleico (Patentes de EE.UU. números 6.229.033 y 6.248.939) y semillas que contienen laurato (Patente de EE.UU. n° 5.639.790). El trabajo en la alteración de la composición se ha centrado predominantemente en las semillas oleaginosas procesadas pero se ha extendido rápidamente a cultivos de semillas no oleaginosas, incluyendo el maíz. Aunque hay un considerable interés en aumentar el contenido de aceite, la única biotecnología actualmente puesta en práctica en este área es la tecnología High-Oil Corn (HOC) (DuPont, Patente de EE.UU. n° 5.704.160). En la HOC se emplean polinizadores de alto contenido de aceite desarrollados mediante cría selectiva clásica junto con hembras híbridas (de parte masculina estéril) selectas en un sistema de producción al que se hace referencia como TopCross. El sistema TopCross High Oil eleva el contenido de aceite de los granos recolectados de maíz de ~3,5% a ~7%, mejorando el contenido energético del grano.

Aunque ha sido fructífero, el sistema de producción HOC presenta limitaciones inherentes. En primer lugar, el sistema de tener un bajo porcentaje de polinizadores responsables del conjunto de semillas de un campo completo contiene riesgos inherentes, particularmente en años de sequía. En segundo lugar, los contenidos de aceite de los actuales campos de HOC han alcanzado una meseta de aproximadamente 9% de aceite. Finalmente, el maíz con alto contenido de aceite no es esencialmente un cambio bioquímico, sino más bien un mutante anatómico (tamaño de embrión aumentado) que presenta el resultado indirecto de un aumento del contenido de aceite. Por estas razones, sería especialmente valiosa una estrategia alternativa para alto contenido de aceite, particularmente una que derivara de una producción bioquímica alterada.

Los cultivos diana más obvios para el mercado del aceite procesado son la soja y la semilla de colza, y una gran recopilación de trabajo comercial (por ejemplo, la Patente de EE.UU. n° 5.952.544 y la solicitud PCT WO9411516) demuestra que *Arabidopsis* es un excelente modelo para el metabolismo del aceite en estos cultivos. Exploraciones bioquímicas en la composición del aceite de semillas han permitido identificar genes de *Arabidopsis* para muchas enzimas biosintéticas críticas y han conducido a la identificación de genes ortólogos agrónomicamente importantes. Por ejemplo, exploraciones en que se usan poblaciones sometidas a mutagénesis química han permitido identificar mutantes lipídicos cuyas semillas presentan una composición de ácidos grasos alterada (B. Lemieux et al., 1990, Theor. Appl. Genet. 80, 234-240; D. W. James y H. K. Dooner, 1990, Theor. Appl. Genet. 80, 241-245). Exploraciones de mutagénesis por DNA de transferencia (DNA-T) (Feldmann et al., Science 243, 1351-1354, 1989) que permitían detectar una composición de ácidos grasos alterada permitieron identificar los genes de omega-3 desaturasa (*FAD3*) y delta-12 desaturasa (*FAD2*) (Patente de EE.UU. n° 5952544; N. S. Yadav et al., 1993, Plant Physiol. 103, 467-476; Okuley et al., Plant Cell, enero de 1994, 6 (1): 147-58). En una exploración que se centraba en el contenido de aceite en vez de en la calidad del aceite, se analizaban mutantes químicamente inducidos en cuanto a semillas arrugadas o densidad de semillas alterada, a partir de lo cual se infería un contenido alterado de aceite de semillas (N. Focks y C. Benning, Plant Physiol. 118: 91-101, 1998). Otra exploración, diseñada para identificar enzimas implicadas en la producción de ácidos grasos de cadena muy larga, permitió identificar una

mutación en el gen que codifica una diacilglicerol aciltransferasa (DGAT) como responsable de una acumulación reducida de triacilglicerol en semillas (V. Katavic et al., *Plant Physiol.*, mayo de 1995, 108 (1): 399-409). Se mostró además que la sobreexpresión, específica de semillas, del cDNA de DGAT estaba asociada con un contenido aumentado de aceite de semillas (Jako et al., *Plant Physiol.*, junio de 2001, 126 (2): 861-74).

- 5 En las plantas, el etiquetado por activación se refiere a un método para generar mutaciones aleatorias mediante la inserción de una construcción de ácido nucleico heteróloga que comprende secuencias reguladoras (por ejemplo, un potenciador) en el genoma de una planta. Las secuencias reguladoras pueden actuar para potenciar la transcripción de uno o más genes nativos de la planta; en consecuencia, el etiquetado por activación es un método fructífero para generar "ganancia de función", generalmente mutantes dominantes [véanse, por ejemplo, Hayashi et al., *Science* (1992) 258: 1350-1353; y Weigel et al., *Plant Physiology* (2000) 122: 1003-1013]. La construcción insertada proporciona una etiqueta molecular para la rápida identificación de la planta nativa cuya expresión incorrecta causa el fenotipo mutante. El etiquetado por activación puede también causar fenotipos con "pérdida de función". La inserción puede dar lugar a la alteración de un gen nativo de la planta, en cuyo caso el fenotipo es generalmente recesivo.
- 10
- 15 El etiquetado por activación ha sido usado en diversas especies, incluyendo el tabaco y *Arabidopsis*, para identificar muchas clases diferentes de fenotipos mutantes y los genes asociados con estos fenotipos [Wilson et al., *Plant Cell* (1996) 8: 659-671; Schaffer et al., *Cell* (1998) 93: 1219-1229; Fridborg et al., *Plant Cell* (1999) 11: 1019-1032; Kardalsky et al., *Science* (1999) 286: 1962-1965; S. Christensen et al., 9<sup>th</sup> International Conference on *Arabidopsis* Research, Univ. of Wisconsin-Madison, 24-28 de junio de 1998, Resumen 165].

## 20 **Compendio de la invención**

La invención proporciona una planta transgénica como la expuesta en la Reivindicación 1 adjunta. En realizaciones preferidas, la planta transgénica es seleccionada del grupo que consiste en semilla de colza, soja, maíz, girasol, algodón, cacao, cártamo, palma de aceite, palma de coco, lino, ricino y cacahuete. La invención proporciona además un método para producir aceite, que comprende cultivar la planta transgénica y recuperar el aceite de dicha planta.

- 25 La planta transgénica de la invención se produce mediante un método que comprende introducir, en células progenitoras de la planta, un vector de transformación de plantas que comprende una secuencia de nucleótidos que codifica el polipéptido HIO30 que comprende la secuencia de aminoácidos ID. SEC. n° 2 y cultivar las células progenitoras transformadas para producir una planta transgénica, en donde la secuencia polinucleotídica HIO30 se expresa causando un fenotipo con un porcentaje aumentado de masa de semilla que es aceite.

## 30 **Descripción detallada de la invención**

### Definiciones

A menos que se indique otra cosa, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que tendrían para un experto en la técnica de la presente invención. Para las definiciones y términos de la técnica, los profesionales son particularmente dirigidos a Sambrook et al., "Molecular Cloning: A Laboratory Manual" (segunda edición), Cold Spring Harbor Press, Plainview, New York, 1989, y a F. M. Ausubel et al., "Current Protocols in Molecular Biology", John Wiley & Sons, New York, N.Y., 1993. Se ha de entender que esta invención no se limita a la metodología, los protocolos y los reactivos particulares descritos, ya que estos pueden variar.

- 35
- 40 Como se usa en esta memoria, el término "vector" se refiere a una construcción de ácido nucleico diseñada para transferencia entre células huésped diferentes. Un "vector de expresión" se refiere a un vector que tiene la capacidad para incorporar fragmentos de DNA heterólogo a una célula extraña y expresarlos en ella. Se dispone comercialmente de muchos vectores de expresión procarióticos y eucarióticos. La selección de vectores de expresión apropiados está dentro del conocimiento de quienes tienen experiencia en la técnica.

- 45 Una construcción o secuencia de ácido nucleico "heteróloga" tiene una porción de la secuencia que no es nativa para la célula vegetal en que se expresa. Con respecto a una secuencia de control, "heteróloga" se refiere a una secuencia de control (es decir, un promotor o potenciador) que no actúa en la naturaleza para regular el mismo gen cuya expresión está actualmente regulando. En general, las secuencias de ácido nucleico heterólogas no son endógenas con respecto a la célula o parte del genoma en que están presentes, y han sido añadidas a la célula por infección, transfección, microinyección, electroporación o similar. Una construcción de ácido nucleico "heteróloga" puede contener una combinación de secuencia de control/secuencia de codificación de DNA que sea igual o diferente que una combinación de secuencia de control/secuencia de codificación de DNA hallada en la planta nativa.

- 50
- 55 Como se usa en esta memoria, el término "gen" significa el segmento de DNA implicado en la producción de una cadena polipeptídica, el cual puede incluir o no regiones que precedan a la región de codificación, tales como, por ejemplo, secuencias 5' no traducidas [5'-UTR (del inglés, 5'-untranslated region)] o "líder" y secuencias 3'-UTR o "tráiler", así como secuencias intermedias (intrones) entre segmentos de codificación individuales (exones) y una secuencia reguladora no transcrita.

5 Como se usa en esta memoria, "recombinante" incluye la referencia a una célula o vector que ha sido modificado por la introducción de una secuencia de ácido nucleico heteróloga, o a la célula que procede de una célula así modificada. De esta manera, por ejemplo, las células recombinantes expresan genes que no se encuentran en idéntica forma en la forma nativa (no recombinante) de la célula o expresan genes nativos que, en cualquier caso, se expresan anormalmente, se infraexpresan o no se expresan en absoluto como resultado de una deliberada intervención humana.

10 Como se usa en esta memoria, la locución "expresión génica" se refiere al proceso mediante el cual se produce un polipéptido basándose en la secuencia de ácido nucleico de un gen. El proceso incluye tanto transcripción como traducción; en consecuencia, "expresión" puede referirse a una secuencia polinucleotídica o polipeptídica, o a ambas. A veces, la expresión de una secuencia polinucleotídica no conduce a la traducción en proteína. "Sobreexpresión" se refiere a una expresión aumentada de una secuencia polinucleotídica y/o polipeptídica con respecto a su expresión en una planta de tipo silvestre [u otra referencia (por ejemplo, no transgénica)] y puede tener que ver con una secuencia presente en la naturaleza o no presente en la naturaleza. "Expresión ectópica" se refiere a la expresión en un momento, en un lugar y/o con un nivel aumentado, que no tiene naturalmente lugar en la planta no alterada o de tipo silvestre. "Infraexpresión" se refiere a una expresión disminuida de una secuencia polinucleotídica y/o polipeptídica, generalmente de un gen endógeno, con respecto a su expresión en una planta de tipo silvestre. Las expresiones "expresión incorrecta" y "expresión alterada" abarcan sobreexpresión, infraexpresión y expresión ectópica.

20 En el contexto de la inserción de una secuencia de ácido nucleico en una célula, el término "introducida" significa "transfección" o "transformación" o "transducción" e incluye la referencia a la incorporación de una secuencia de ácido nucleico a una célula eucariótica o procariótica, donde la secuencia de ácido nucleico se puede incorporar al genoma de la célula (por ejemplo, cromosoma, plásmido, plasto o DNA mitocondrial), se puede convertir en un replicón autónomo o se puede expresar transitoriamente (por ejemplo, mRNA transfectado).

25 Como se usa en esta memoria, una "célula vegetal" se refiere a cualquier célula procedente de una planta, incluyendo células de tejido indiferenciado (por ejemplo, un callo) así como semillas, polen propágulos y embriones de plantas.

Como se usan en esta memoria, las expresiones "nativo" y "tipo silvestre" relativas a un rasgo o fenotipo de planta dado, se refieren a la forma en que ese rasgo o fenotipo se encuentra en la misma variedad de planta en la naturaleza.

30 Como se utiliza en esta memoria, el término "modificado" relativo a un rasgo de una planta se refiere a un cambio en el fenotipo de una planta transgénica con respecto al de la planta no transgénica similar. Un "fenotipo con contenido alterado de aceite" se refiere al fenotipo mensurable de una planta genéticamente modificada, donde la planta presenta un aumento o disminución estadísticamente significativo en el contenido global de aceite (es decir, el porcentaje de masa de semilla que es aceite) en comparación con el de la planta similar pero no modificada. Un fenotipo con alto contenido de aceite se refiere a un aumento en el contenido global de aceite.

35 Como se utiliza en esta memoria, una secuencia polinucleotídica o gen "mutante" difiere de la correspondiente secuencia polinucleotídica o gen de tipo silvestre en términos de secuencia o expresión, donde la diferencia contribuye a un fenotipo o rasgo vegetal modificado. Con respecto a una planta o línea vegetal, el término "mutante" se refiere a una planta o línea vegetal que tiene un fenotipo o rasgo vegetal modificado, donde el fenotipo o rasgo modificado está asociado con la expresión modificada de una secuencia polinucleotídica o gen de tipo silvestre.

40 Como se utiliza en esta memoria, el término "T1" se refiere a la generación de plantas a partir de la semilla de plantas T0. La generación T1 es el primer conjunto de plantas transformadas que puede ser seleccionado por aplicación de un agente de selección, por ejemplo, un antibiótico o un herbicida, para el que la planta transgénica contiene el correspondiente gen de resistencia. El término "T2" se refiere a la generación de plantas por autofertilización de las flores de plantas T1, previamente seleccionadas como transgénicas. Las plantas T3 se generan a partir de las plantas T2, etc. Como se usa en esta memoria, la "progenie directa" de una planta dada procede de la semilla (o, a veces, de otro tejido) de esa planta y está en la generación inmediatamente subsiguiente; por ejemplo, para un linaje dado, una planta T2 es la progenie directa de una planta T1. La "progenie indirecta" de una planta dada procede de la semilla (u otro tejido) de la progenie directa de esa planta, o de la semilla (u otro tejido) de generaciones subsiguientes de ese linaje; por ejemplo, una planta T3 es la progenie indirecta de una planta T1.

45 Como se usa en esta memoria, la expresión "parte vegetal" incluye cualquier órgano o tejido vegetal, incluyendo, sin limitación, semillas, embriones, regiones meristemáticas, tejido de callo, hojas, raíces, brotes, gametofitos, esporofitos, polen y microesporas. Se pueden obtener células vegetales a partir de cualquier órgano o tejido vegetal y de cultivos preparados a partir de los mismos. La clase de plantas que se puede utilizar en los métodos de la presente invención es, en general, tan amplia como la clase de plantas superiores sensibles a técnicas de transformación, incluyendo plantas tanto monocotiledóneas como dicotiledóneas.

50 Como se utiliza en esta memoria, "planta transgénica" incluye una planta que comprende un polinucleótido

heterólogo en su genoma. El polinucleótido heterólogo puede estar establemente integrado en el genoma o puede ser extracromosómico. Preferiblemente, el polinucleótido de la presente invención está establemente integrado en el genoma para que el polinucleótido sea hecho pasar a las generaciones sucesivas. Una célula, tejido u órgano vegetal o una planta en que se han introducido los polinucleótidos heterólogos es considerada "transformada", "transfectada" o "transgénica". También se consideran transgénicas las progenies directa e indirecta de plantas o células vegetales transformadas que también contienen el polinucleótido heterólogo.

#### Identificación de plantas con un fenotipo con contenido alterado de aceite

Se usó una exploración mediante etiquetado por activación en *Arabidopsis* para identificar la asociación entre el gen que denominamos "HIO30" (At3g52260; GI#22331747:54-938) y un fenotipo con contenido alterado de aceite (específicamente, un fenotipo con alto contenido de aceite). En resumen, y como se describe adicionalmente en los Ejemplos, se mutó un gran número de plantas de *Arabidopsis* con el vector pSKI015, que comprende un DNA-T del plásmido Ti de *Agrobacterium tumefaciens*, un elemento potenciador vírico, y un gen marcador seleccionable (Weigel et al., *supra*). Cuando el DNA-T se inserta en el genoma de plantas transformadas, el elemento potenciador puede causar la suprarregulación de genes en las inmediaciones, generalmente dentro de aproximadamente 10 kilobases (kb) de la inserción. Se expusieron plantas T1 al agente selectivo con objeto de recuperar específicamente las plantas transformadas que expresaban el marcador seleccionable y, por lo tanto, contenían inserciones de DNA-T. Se recogieron muestras de aproximadamente 15-20 semillas T2 de plantas transformadas T1 y se extrajeron los lípidos de las semillas completas. Se llevó a cabo un análisis por cromatografía en fase gaseosa (GC; del inglés, gas chromatography) para determinar el contenido y la composición de ácidos grasos de las muestras de semillas.

Se identificó una línea de *Arabidopsis* que mostraba un fenotipo con alto contenido de aceite, en donde los aceites (es decir, los ácidos grasos) constituían aproximadamente el 35% de la masa de semilla. Se descubrió la asociación del gen HIO30 con el fenotipo con alto contenido de aceite mediante el análisis de la secuencia de DNA genómico que flanquea la inserción de DNA-T en la línea identificada. En consecuencia, se pueden emplear genes y/o polipéptidos HIO30 en el desarrollo de plantas genéticamente modificadas que tienen un fenotipo con contenido modificado de aceite ("un fenotipo HIO30"). Se pueden usar genes HIO30 en la generación de cultivos de semillas oleaginosas que proporcionen una producción de aceite mejorada a través del procesamiento de semillas oleaginosas, y en la generación de cultivos de granos para pienso que proporcionen una energía aumentada para la alimentación de animales. Los genes HIO30 se pueden usar además para aumentar el contenido de aceite de cultivos de aceites especiales, con objeto de aumentar la producción de ácidos grasos deseados poco habituales. Las plantas transgénicas que han sido genéticamente modificadas para que expresen HIO30 se pueden usar en la producción de aceite, en donde se cultivan las plantas transgénicas y se obtiene el aceite de partes de la planta (por ejemplo, la semilla) usando métodos estándares.

#### Ácidos nucleicos y polipéptidos HIO30

La secuencia de ácido nucleico (DNA genómico) HIO30 de *Arabidopsis* se proporciona en la ID. SEC. n° 1 y en la entrada GI#22331747:54-938 de Genbank. La correspondiente secuencia proteica se proporciona en la ID. SEC. n° 2 y en GI#22331748. Más adelante, en el Ejemplo 3, se describen ácidos nucleicos y/o proteínas que son ortólogos o parálogos de HIO30 de *Arabidopsis*.

Como se utiliza en esta memoria, la expresión "polipéptido HIO30" se refiere a una proteína HIO30 de longitud completa o a un fragmento, derivado (variante) u ortólogo de la misma que es "funcionalmente activo", lo que significa que el fragmento, derivado u ortólogo de la proteína presenta una o más de las actividades funcionales asociadas con el polipéptido de ID. SEC. n° 2. En una realización preferida, un polipéptido HIO30 funcionalmente activo causa un fenotipo con contenido alterado de aceite cuando se expresa incorrectamente en una planta. En otra realización preferida, la expresión incorrecta del polipéptido HIO30 causa un fenotipo con alto contenido de aceite en una planta. En otra realización, un polipéptido HIO30 funcionalmente activo es capaz de rescatar la actividad HIO30 endógena incompleta (incluyendo deficiente) cuando se expresa en una planta o en células vegetales; el polipéptido de rescate puede ser de la misma especie o de una especie diferente como aquella con actividad incompleta. En otra realización, un fragmento funcionalmente activo de un polipéptido HIO30 de longitud completa (es decir, un polipéptido nativo que tiene la secuencia de ID. SEC. n° 2 o un ortólogo del mismo presente en la naturaleza) conserva una o más de las propiedades biológicas asociadas con el polipéptido HIO30 de longitud completa, tal como actividad de señalización, actividad de unión, actividad catalítica o actividad de localización celular o extracelular. Un fragmento de HIO30 comprende preferiblemente un dominio de HIO30, tal como un dominio C- o N-terminal o catalítico, entre otros, y comprende preferiblemente al menos 10, preferiblemente al menos 20, más preferiblemente al menos 25, y lo más preferiblemente al menos 50, aminoácidos contiguos de una proteína HIO30. Los dominios funcionales se pueden identificar utilizando el programa PFAM (A. Bateman et al., 1999, *Nucleic Acids Res.* 27: 260-262; sitio web en pfam.wustl.edu). Un fragmento de HIO30 preferido comprende un dominio de seudouridilato sintasa (PFam00849). Las variantes funcionalmente activas de polipéptidos HIO30 de longitud completa o de fragmentos de los mismos incluyen polipéptidos con inserciones, supresiones o sustituciones de aminoácidos que conservan una o más de las propiedades biológicas asociadas con el polipéptido HIO30 de longitud completa. En ciertos casos, se generan variantes que cambian el procesamiento postraduccional de un polipéptido HIO30. Por ejemplo, en comparación con el polipéptido nativo, las variantes pueden tener unas características de transporte proteico o localización proteica alteradas o una semivida proteica alterada.

Como se utiliza en esta memoria, la expresión "ácido nucleico HIO30" abarca ácidos nucleicos con la secuencia proporcionada, o con una secuencia complementaria de la secuencia proporcionada en, la ID. SEC. nº 1, así como fragmentos, derivados u ortólogos funcionalmente activos de los mismos. Un ácido nucleico HIO30 de esta invención puede ser DNA, derivado de DNA genómico o cDNA, o RNA.

5 En una realización, un ácido nucleico HIO30 funcionalmente activo codifica, o es complementario de un ácido nucleico que codifica, un polipéptido HIO30 funcionalmente activo. Dentro de esta definición se incluye el DNA genómico que sirve como molde para un transcrito de RNA primario (es decir, un precursor de mRNA) que requiere procesamiento, tal como corte y empalme, antes de que codifique el polipéptido HIO30 funcionalmente activo. Un ácido nucleico HIO30 puede incluir otras secuencias no codificadoras, las cuales pueden ser transcritas o no; dichas  
10 secuencias incluyen, entre otras, 5'- y 3'-UTRs, señales de poliadenilación y secuencias reguladoras que controlan la expresión génica, como es sabido en la técnica. Algunos polipéptidos requieren eventos de procesamiento, tales como escisión proteolítica, modificación covalente, etc., con objeto de que lleguen a ser totalmente activos. En consecuencia, los ácidos nucleicos funcionalmente activos pueden codificar el polipéptido HIO30 maduro o el preprocesado, o una forma intermedia. Un polinucleótido HIO30 puede incluir también secuencias de codificación heterólogas, por ejemplo, secuencias que codifiquen un marcador incluido para facilitar la purificación del polipéptido  
15 fusionado, o un marcador de transformación.

En otra realización, se puede utilizar un ácido nucleico HIO30 funcionalmente activo en la generación de fenotipos HIO30 con "pérdida de función" a través de, por ejemplo, supresión antisentido, cosupresión, etc.

20 En una realización preferida, un ácido nucleico HIO30 utilizado en los métodos de esta invención comprende una secuencia de ácido nucleico que codifica, o es complementaria de una secuencia que codifica, un polipéptido HIO30 que tiene una identidad secuencial de al menos 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95% o más con respecto a la secuencia polipeptídica presentada en la ID. SEC. nº 2.

25 En otra realización, un polipéptido HIO30 de la invención comprende una secuencia polipeptídica con una identidad de al menos 50% o 60% con respecto a la secuencia polipeptídica HIO30 de ID. SEC. nº 2, y puede tener una identidad secuencial de al menos 70%, 80%, 85%, 90%, 95% o más con respecto a la secuencia polipeptídica HIO30 de ID. SEC. nº 2. En otra realización, un polipéptido HIO30 comprende una secuencia polipeptídica con una identidad secuencial de al menos 50%, 60%, 70%, 80%, 85%, 90%, 95% o más con respecto a un fragmento funcionalmente activo del polipéptido presentado en la ID. SEC. nº 2, tal como un dominio de pseudouridilato sintasa. En aún otra realización, un polipéptido HIO30 comprende una secuencia polipeptídica con una identidad de al  
30 menos 50%, 60%, 70%, 80% o 90% con respecto a la secuencia polipeptídica de ID. SEC. nº 2, en toda su longitud, y comprende un dominio de pseudouridilato sintasa.

35 En otro aspecto, una secuencia polinucleotídica HIO30 tiene una identidad de al menos 50% a 60%, en toda su longitud, con respecto a la secuencia de ácido nucleico HIO30 presentada como ID. SEC. nº 1, o a secuencias de ácido nucleico que son complementarias de dicha secuencia HIO30, y puede comprender una identidad secuencial de al menos 70%, 80%, 85%, 90%, 95% o más con respecto a la secuencia HIO30 presentada como ID. SEC. nº 1 o a un fragmento funcionalmente activo de la misma, o a secuencias complementarias.

40 Como se emplea en esta memoria, "porcentaje (%) de identidad secuencial" con respecto a una secuencia objetivo especificada, o una porción especificada de la misma, se define como el porcentaje de nucleótidos o aminoácidos de la secuencia derivada candidata que son idénticos a los nucleótidos o aminoácidos de la secuencia objetivo (o la porción especificada de la misma) después de alinear las secuencias e introducir huecos, si fuera necesario, para alcanzar el máximo porcentaje de identidad secuencial, según se genera mediante el programa WU-BLAST-2.0a19 [Altschul et al., J. Mol. Biol. (1997) 215: 403-410; sitio web en blast.wustl.edu/blast/README.html) con los parámetros de búsqueda ajustados a los valores por omisión. Los parámetros HSP S y HSP S2 son valores  
45 dinámicos y son establecidos por el propio programa dependiendo de la composición de la secuencia particular y de la composición de la base de datos particular frente a la cual se busca la secuencia de interés. Se determina un "% de valor de identidad" mediante el número de correspondientes nucleótidos o aminoácidos idénticos, dividido por la longitud de la secuencia para la cual se presenta el porcentaje de identidad. Se determina el "porcentaje (%) de similitud de secuencias de aminoácidos" realizando el mismo cálculo que para la determinación del % de identidad de secuencias de aminoácidos pero incluyendo en el cálculo las sustituciones de aminoácido conservativas además de los aminoácidos idénticos. Una sustitución de aminoácido conservativa es aquella en que se sustituye un aminoácido por otro aminoácido que tiene unas propiedades similares para que la plegadura o actividad de la proteína no resulte significativamente afectada. Los aminoácidos aromáticos que pueden ser sustituidos entre sí son fenilalanina, triptófano y tirosina; los aminoácidos hidrófobos intercambiables son leucina, isoleucina, metionina y valina; los aminoácidos polares intercambiables son glutamina y asparagina; los aminoácidos básicos  
50 intercambiables son arginina, lisina e histidina; los aminoácidos ácidos intercambiables son ácido aspártico y ácido glutámico; y los aminoácidos pequeños intercambiables son alanina, serina, treonina, cisteína y glicocola.

60 Las moléculas de ácido nucleico derivadas de las moléculas de ácido nucleico objetivo incluyen secuencias que se hibridan selectivamente con la secuencia de ácido nucleico de ID. SEC. nº 1. El rigor de la hibridación puede ser controlado mediante la temperatura, la fuerza iónica, el pH y la presencia de agentes desnaturizantes, tal como formamida, durante la hibridación y el lavado. Las condiciones rutinariamente utilizadas son bien conocidas [véanse,

por ejemplo, "Current Protocols in Molecular Biology", volumen 1, capítulo 2.10, John Wiley & Sons, Publishers (1994); y Sambrook et al., "Molecular Cloning", Cold Spring Harbor (1989)]. En ciertas realizaciones, una molécula de ácido nucleico de la invención es capaz de hibridarse con una molécula de ácido nucleico que contiene la secuencia de nucleótidos de ID. SEC. nº 1 bajo unas condiciones de hibridación rigurosas que son: prehibridación de los filtros que contienen ácido nucleico durante un período de 8 horas a toda la noche, a 65 °C, en una disolución que comprende citrato de fuerza única (SSC; del inglés, single strength citrate) 6X (SSC 1X es NaCl 0,15 M y citrato sódico 0,015 M, pH de 7,0), disolución de Denhardt 5X, pirofosfato sódico al 0,05% y 100 µg/ml de DNA de esperma de arenque; hibridación durante 18-20 horas a 65 °C en una disolución que contiene SSC 6X, disolución de Denhardt 1X, 100 µg/ml de tRNA de levadura y pirofosfato sódico al 0,05%; y lavado de los filtros a 65 °C durante 1 hora en una disolución que contiene SSC 0,1X y dodecilsulfato sódico (SDS; del inglés, sodium dodecyl sulfate) al 0,1%. En otras realizaciones se utilizan unas condiciones de hibridación moderadamente rigurosas que son: pretratamiento de los filtros que contienen ácido nucleico durante 6 horas a 40 °C en una disolución que contiene formamida al 35%, SSC 5X, Tris-HCl 50 mM (pH de 7,5), EDTA 5 mM, PVP al 0,1%, Ficoll al 0,1%, BSA al 1% y 500 µg/ml de DNA de esperma de salmón desnaturalizado; hibridación durante 18-20 horas a 40 °C en una disolución que contiene formamida al 35%, SSC 5X, Tris-HCl 50 mM (pH de 7,5), EDTA 5 mM, PVP al 0,02%, Ficoll al 0,02%, BSA al 0,2%, 100 µg/ml de DNA de esperma de salmón, y sulfato de dextrano al 10% (peso/volumen); y luego lavado dos veces durante 1 hora a 55 °C en una disolución que contiene SSC 2X y SDS al 0,1%. Alternativamente, se pueden utilizar unas condiciones de bajo rigor que comprenden: incubación durante un período de 8 horas a toda la noche, a 37 °C, en una disolución que comprende formamida al 20%, SSC 5X, fosfato sódico 50 mM (pH de 7,6), disolución de Denhardt 5X, sulfato de dextrano al 10%, y 20 µg/ml de DNA de esperma de salmón fragmentado y desnaturalizado; hibridación en el mismo tampón durante un período de 18 a 20 horas; y lavado de los filtros en SSC 1X durante 1 hora a aproximadamente 37 °C.

Como resultado de la degeneración del código genético, se pueden producir diversas secuencias polinucleotídicas que codifiquen un polipéptido HIO30. Por ejemplo, se pueden seleccionar codones para aumentar la velocidad a que tiene lugar la expresión del polipéptido en una especie huésped particular, de acuerdo con la óptima utilización de codones dictada por el organismo huésped particular (véase, por ejemplo, Nakamura et al., 1999, Nucleic Acids Res. 27: 292). Dichas variantes secuenciales pueden ser utilizadas en los métodos de esta invención.

En los métodos de la invención se pueden usar ortólogos del HIO30 de *Arabidopsis*. Los métodos para identificar los ortólogos en otra especie vegetal son conocidos en la técnica. Normalmente, los ortólogos de diferentes especies conservan la misma función a causa de la presencia de uno o más motivos proteicos y/o estructuras tridimensionales. En la evolución, cuando un evento de duplicación génica sigue a la especiación, un único gen en una especie, tal como de *Arabidopsis*, puede corresponder a múltiples genes (parálogos) en otra. Como se usa en esta memoria, el término "ortólogos" abarca parálogos. Cuando se dispone de datos de secuencias para una especie vegetal particular, los ortólogos son generalmente identificados mediante un análisis de homologías de secuencias, tal como un análisis por BLAST, usando normalmente secuencias proteicas señuelo. Se asignan secuencias a un posible ortólogo si la secuencia más satisfactoria del resultado del BLAST directo recupera la secuencia original en cuestión en el BLAST inverso [M. A. Huynen y P. Bork, Proc. Natl. Acad. Sci. (1998), 95: 5849-5856; M. A. Huynen et al., Genome Research (2000), 10: 1204-1210]. Se pueden usar programas para el alineamiento múltiple de secuencias, tal como CLUSTAL (J. D. Thompson et al., 1994, Nucleic Acids Res. 22: 4673-4680), para resaltar regiones y/o restos conservados de proteínas ortólogas y generar árboles filogenéticos. En un árbol filogenético que representa múltiples secuencias homólogas de diversas especies (por ejemplo, recuperadas a través del análisis por BLAST), las secuencias ortólogas de dos especies aparecen generalmente más cerca en el árbol con respecto a todas las demás secuencias de estas dos especies. El reconocimiento de plegaduras ("threading") estructural u otro análisis de la plegadura de proteínas (por ejemplo, usando un software de ProCeryon Biosciences, Salzburgo, Austria) también permite identificar posibles ortólogos. Los métodos de hibridación de ácido nucleico pueden ser también utilizados para hallar genes ortólogos y son preferidos cuando no se dispone de datos sobre secuencias. La PCR degenerada y la exploración de bancos de cDNA o DNA genómico son métodos habituales para hallar secuencias génicas relacionadas y son bien conocidas en la técnica (véanse, por ejemplo, Sambrook, *supra*; C. Dieffenbach y G. Dveksler (redactores), "PCR Primer: A Laboratory Manual", Cold Spring Harbor Laboratory Press, New York, 1989). Por ejemplo, en Sambrook et al., *supra*, se describen métodos para generar un banco de cDNA a partir de la especie vegetal de interés y sondar el banco con sondas génicas parcialmente homólogas. Como sonda se puede usar una porción muy conservada de la secuencia de codificación de HIO30 de *Arabidopsis*. Los ácidos nucleicos HIO30 ortólogos se pueden hibridar con el ácido nucleico de ID. SEC. nº 1 bajo condiciones de alto, moderado o bajo rigor. Después de la multiplicación o el aislamiento de un segmento de un supuesto ortólogo, ese segmento puede ser clonado y secuenciado mediante técnicas estándares y ser utilizado como sonda para aislar un clon de cDNA o genómico completo. Alternativamente, es posible iniciar un proyecto EST para generar una base de datos sobre información de secuencias para la especie vegetal de interés. En otro planteamiento, para el aislamiento de ortólogos se usan anticuerpos que se unen específicamente a polipéptidos HIO30 conocidos (véase, por ejemplo, E. Harlow y D. Lane, "Using Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, 1999, New York). Un análisis por transferencia Western permite determinar que un ortólogo de HIO30 (es decir, una proteína ortóloga) está presente en un extracto crudo de una especie vegetal particular. Cuando se observa reactividad, la secuencia que codifica el ortólogo candidato puede ser aislada al explorar bancos de expresión que representan la especie vegetal particular. Se pueden construir bancos de expresión en una diversidad de vectores comercialmente asequibles, incluyendo lambda gt11, como se describe en

Sambrook et al., *supra*. Una vez que se ha(n) identificado el(los) ortólogo(s) candidato(s) por cualquiera de estos medios, se usan secuencias ortólogas candidatas como señuelo (la "cuestión") para el BLAST inverso frente a secuencias de *Arabidopsis* u otra especie en que se han identificado secuencias de ácido nucleico y/o polipéptido HIO30.

- 5 Se pueden obtener ácidos nucleicos y polipéptidos HIO30 usando cualquier método disponible. En este campo técnico son bien conocidas, por ejemplo, las técnicas para aislar secuencias de cDNA o DNA genómico de interés explorando bancos de DNA o usando la reacción en cadena de la polimerasa (PCR; del inglés, polymerase chain reaction), como se describió previamente. Alternativamente, se puede sintetizar una secuencia de ácido nucleico. Se puede usar cualquier método conocido, tal como la mutagénesis dirigida al sitio (T. A. Kunkel et al., *Methods Enzymol.* 204: 125-39, 1991), para introducir cambios deseados en un ácido nucleico clonado.

En general, los métodos de la invención implican incorporar la forma deseada del ácido nucleico HIO30 en un vector de expresión vegetal para la transformación de células vegetales, y que el polipéptido HIO30 se exprese en la planta huésped.

- 15 Una molécula de ácido nucleico HIO30 aislada no está en la forma ni el ambiente en que se encuentra en la naturaleza, y es identificada y separada de al menos una molécula de ácido nucleico contaminante con la que se asocia normalmente en la fuente natural del ácido nucleico HIO30. Sin embargo, una molécula de ácido nucleico HIO30 aislada incluye moléculas de ácido nucleico HIO30 contenidas en células que expresan normalmente HIO30, donde, por ejemplo, la molécula de ácido nucleico está en una posición cromosómica distinta a la de las células naturales.

- 20 Generación de plantas genéticamente modificadas con un fenotipo con contenido alterado de aceite

Se pueden usar ácidos nucleicos y polipéptidos HIO30 en la generación de plantas genéticamente modificadas que tengan un fenotipo con contenido modificado de aceite. Como se emplea en esta memoria, un " fenotipo con contenido modificado de aceite" se puede referir a un contenido modificado de aceite en cualquier parte de la planta; el contenido modificado de aceite se observa a menudo en las semillas. En una realización preferida, la expresión alterada del gen HIO30 en una planta se usa para generar plantas con un fenotipo con alto contenido de aceite.

- Los métodos descritos en la presente memoria son generalmente aplicables a todas las plantas. Aunque el etiquetado por activación y la identificación génica se llevan a cabo en *Arabidopsis*, el gen HIO30 (o un ortólogo, variante o fragmento del mismo) se puede expresar en cualquier tipo de planta. En una realización preferida, la invención se dirige a plantas productoras de aceite, que producen y almacenan triacilglicerol en órganos específicos, principalmente en las semillas. Dichas especies incluyen soja (*Glycine max*), semilla de colza y canola (incluyendo *Brassica napus* y *B. campestris*), girasol (*Helianthus annuus*), algodón (*Gossypium hirsutum*), maíz (*Zea mays*), cacao (*Theobroma cacao*), cártamo (*Carthamus tinctorius*), palma de aceite (*Elaeis guineensis*), palma de coco (*Cocos nucifera*), lino (*Linum usitatissimum*), ricino (*Ricinus communis*) y cacahuete (*Arachis hypogaea*). La invención también se puede dirigir a plantas portadoras de frutos y vegetales, plantas productoras de granos, plantas productoras de nueces, especies de *Brassica* de ciclo rápido, alfalfa (*Medicago sativa*), tabaco (*Nicotiana*), gramíneas (familia *Poaceae*), otros cultivos de plantas forrajeras, y especies silvestres que pueden ser una fuente de ácidos grasos únicos.

- El técnico experto reconocerá que en este campo técnico existe una gran variedad de técnicas de transformación y que continuamente se llega a disponer de nuevas técnicas. Dentro del alcance de la presente invención se puede emplear cualquier técnica que sea adecuada para la planta huésped diana. Por ejemplo, las construcciones se pueden introducir en una diversidad de formas, incluyendo, pero sin limitarse a, como una cadena de DNA, en un plásmido o en un cromosoma artificial. La introducción de las construcciones en las células vegetales diana puede ser llevada a cabo mediante una diversidad de técnicas, incluyendo, pero sin limitarse a, transformación mediada por *Agrobacterium*, electroporación, microinyección, bombardeo de microproyectiles, coprecipitación de fosfato cálcico-DNA y transformación mediada por liposomas con un ácido nucleico heterólogo. La transformación de la planta es preferiblemente permanente, es decir, por integración de las introducidas construcciones de expresión en el genoma de la planta huésped, para que las construcciones introducidas sean hechas pasar a las sucesivas generaciones de la planta. Dependiendo del uso previsto, una construcción de ácido nucleico heteróloga que comprende un polinucleótido HIO30 puede codificar la proteína entera o una porción biológicamente activa de la misma.

- 50 En una realización, se pueden usar sistemas vector binarios basados en Ti para transferir polinucleótidos. Los vectores binarios estándares de *Agrobacterium* son conocidos por quienes tienen experiencia en la técnica, y muchos son comercialmente asequibles (por ejemplo, pBI121, Clontech Laboratories, Palo Alto, California).

- El procedimiento óptimo para la transformación de plantas con vectores de *Agrobacterium* variará con el tipo de planta que se vaya a transformar. Los métodos ejemplares para transformación mediada por *Agrobacterium* incluyen la transformación de explantes de tejido de hipocótilo, punta de brote, tallo u hoja, procedente de plantones y/o plántulas estériles. Dichas plantas transformadas se pueden reproducir sexualmente o mediante cultivo celular o tisular. Se ha descrito previamente la transformación con *Agrobacterium* para un gran número de tipos diferentes de plantas, y en la bibliografía científica se pueden encontrar métodos para dicha transformación. Son particularmente



relevantes los métodos para transformar cultivos comercialmente importantes, tales como los de semilla de colza [De Block et al., *Plant Physiol.* (1989) 91: 694-701], girasol [Everett et al., *Bio/Technology* (1987) 5: 1201] y soja [Christou et al., *Proc. Natl. Acad. Sci. USA* (1989) 86: 7500-7504; Kline et al., *Nature* (1987) 327: 70].

5 Se puede regular la expresión (incluyendo transcripción y traducción) de HIO30 con respecto al nivel de expresión, el(los) tipo(s) de tejido donde tiene lugar la expresión, y/o la fase de desarrollo de la expresión. Se dispone de diversas secuencias reguladoras heterólogas (por ejemplo, promotores y potenciadores) para controlar la expresión de un ácido nucleico HIO30. Estas incluyen promotores constitutivos, inducibles y regulables, así como promotores y potenciadores que controlan la expresión de un modo tisular o temporalmente específico. Los promotores constitutivos ejemplares incluyen el promotor E4 de frambuesa (Patentes de EE.UU. números 5.783.393 y 10 5.783.394), el promotor 35S de CaMV (J. D. Jones et al., *Transgenic Res.* 1: 285-297, 1992), el promotor de CsVMV (B. Verdaguer et al., *Plant Mol. Biol.* 37: 1055-1067, 1998) y el promotor de actina del melón (solicitud PCT publicada WO0056863). Los promotores tisularmente específicos ejemplares incluyen los promotores E4 y E8 del tomate (Patente de EE.UU. nº 5.859.330) y el promotor del gen 2All del tomate (M. J. J. Van Haaren et al., *Plant Mol. Biol.* 21: 625-640, 1993).

15 En una realización preferida, la expresión de HIO30 está bajo el control de secuencias reguladoras de genes cuya expresión está asociada con el desarrollo precoz de semillas y/o embriones. Los genes de legumbres cuyos promotores están asociados con el desarrollo precoz de semillas y embriones incluyen legumina de *V. faba* (Baumlein et al., 1991, *Mol. Gen. Genet.* 225: 121-8; Baumlein et al., 1992, *Plant J.* 2: 233-9), *usp* de *V. faba* (Fiedler et al., 1993, *Plant Mol. Biol.* 22: 669-79), convicilina de guisante (Bown et al., 1988, *Biochem. J.* 251: 717-26), lectina de guisante (dePater et al., 1993, *Plant Cell* 5: 877-86), beta-faseolina de *P. vulgaris* (Bustos et al., 1991, *EMBO J.* 10: 1469-79), DLEC2 y PHS (beta) de *P. vulgaris* (Bobb et al., 1997, *Nucleic Acids Res.* 25: 641-7) y beta-conglicinina de soja, proteína 7S de almacenamiento (Chamberland et al., 1992, *Plant Mol. Biol.* 19: 937-49). Los genes de cereales cuyos promotores están asociados con el desarrollo precoz de semillas y embriones incluyen glutelina de arroz ("GluA-3," Yoshihara y Takaiwa, 1996, *Plant Cell. Physiol.* 37: 107-11; "GluB-1," Takaiwa et al., 1996, *Plant Mol. Biol.* 30: 1207-21; Washida et al., 1999, *Plant Mol. Biol.* 40: 1-12; "Gt3," Leisy et al., 1990, *Plant Mol. Biol.* 14: 41-50), prolamina de arroz (Zhou y Fan, 1993, *Transgenic Res.* 2: 141-6), prolamina de trigo (Hammond-Kosack et al., 1993, *EMBO J.* 12: 545-54), zeína de maíz (Z4, Matzke et al., 1990, *Plant Mol. Biol.* 14: 323-32) y hordeínas B de cebada (Entwistle et al., 1991, *Plant Mol. Biol.* 17: 1217-31). Otros genes cuyos promotores están asociados con el desarrollo precoz de semillas y embriones incluyen GLO7A de palma de aceite (globulina 7S, Morcillo et al., 2001, *Physiol. Plant* 112: 233-243), napina de *Brassica napus*, proteína 2S de almacenamiento, y el gen napA (Josefsson et al., 1987, *J. Biol. Chem.* 262: 12.196-201; Stalberg et al., 1993, *Plant Mol. Biol.* 1993, 23: 671-83; Ellerstrom et al., 1996, *Plant Mol. Biol.* 32: 1019-27), oleosina de *Brassica napus* (Keddie et al., 1994, *Plant Mol. Biol.* 24: 327-40), oleosina de *Arabidopsis* (Plant et al., 1994, *Plant Mol. Biol.* 25: 193-205), FAE1 de *Arabidopsis* (Rossak et al., 2001, *Plant Mol. Biol.* 46: 717-25), conA de *Canavalia gladiata* (Yamamoto et al., 1995, *Plant Mol. Biol.* 27: 729-41) y estrictosidina sintasa de *Catharanthus roseus* (Str; Ouwkerk y Memelink, 1999, *Mol. Gen. Genet.* 261: 635-43). En otra realización preferida, se usan secuencias reguladoras de genes expresados durante la biosíntesis de aceite (véase, por ejemplo, la Patente de EE.UU. nº 5.952.544). Son promotores alternativos los de genes de proteínas de almacenamiento de plantas (Bevan et al., 1993, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 342: 209-15).

40 Se pueden llevar a cabo pruebas moleculares y genéticas estándares para analizar adicionalmente la asociación entre un gen y un fenotipo observado. A continuación se describen técnicas ejemplares.

#### 1. Análisis de DNA/RNA

Se pueden determinar los patrones de expresión génica específicos de la fase y el tejido en líneas mutantes frente a líneas de tipo silvestre mediante, por ejemplo, hibridación in situ. Se puede llevar a cabo el análisis del estado de metilación del gen, especialmente de las regiones reguladoras flanqueadoras. Otras técnicas adecuadas incluyen sobreexpresión, expresión ectópica, expresión en otra especie vegetal e inactivación ("knock-out") génica [genética inversa, inactivación dirigida, silenciamiento génico víricamente inducido (VIGS; del inglés, *viral induced gene silencing*; véase D. Baulcombe, *Arch. Virol.*, Supl. 15: 189-201, 1999)].

50 En una aplicación preferida, se utiliza la perfiladura de la expresión, generalmente mediante el análisis con micromatrices, para medir simultáneamente diferencias o cambios inducidos en la expresión de muchos genes diferentes. Las técnicas para el análisis con micromatrices son bien conocidas en este campo técnico [M. Schena et al., *Science* (1995) 270: 467-470; D. Baldwin et al., *Curr. Opin. Plant Biol.* 2 (2): 96-103, 1999; F. Dangond, *Physiol. Genomics* (2000) 2: 53-58; N. L. van Hal et al., *J. Biotechnol.* (2000) 78: 271-280; T. Richmond y S. Somerville, *Curr. Opin. Plant Biol.* (2000) 3: 108-116]. Se puede llevar a cabo la perfiladura de la expresión de líneas etiquetadas individuales. Dicho análisis permite identificar otros genes que son coordinadamente regulados como consecuencia de la sobreexpresión del gen de interés, lo que puede ayudar a colocar un gen desconocido en una ruta particular.

#### 2. Análisis de productos génicos

El análisis de productos génicos puede incluir expresión de proteínas recombinantes, producción de antisueros, inmunolocalización, ensayos bioquímicos en cuanto a actividad catalítica u otra actividad, análisis del estado de

fosforilación, y análisis de la interacción con otras proteínas por medio de ensayos de dos híbridos en levadura.

### 3. Análisis de rutas

El análisis de rutas puede incluir situar un gen o producto génico dentro de una ruta bioquímica, metabólica o de señalización particular basándose en su fenotipo de expresión incorrecta o mediante su homología secuencial con genes relacionados. Alternativamente, el análisis puede comprender cruces genéticos con líneas de tipo silvestre y otras líneas mutantes (creándose mutantes dobles) para ordenar el gen en una ruta, o determinar el efecto de una mutación sobre la expresión de genes "informadores" aguas abajo en una ruta.

Generación de plantas mutadas con un fenotipo con contenido alterado de aceite

La invención proporciona además un método para identificar plantas que tienen mutaciones en el HIO30 endógeno que confieren un contenido alterado de aceite, y generar una progenie de estas plantas con contenido alterado de aceite que no esté genéticamente modificada. En un método, llamado "TILLING" (detección de lesiones locales inducidas, en genomas; del inglés, *targeting induced local lesions in genomes*), se inducen mutaciones en la semilla de una planta de interés, por ejemplo, usando un tratamiento con EMS. Las plantas resultantes son cultivadas y autofertilizadas, y se usa la progenie para preparar muestras de DNA. Se usa una PCR específica de HIO30 para identificar si una planta mutada presenta una mutación de HIO30. Las plantas que presentan mutaciones de HIO30 pueden ser luego examinadas en cuanto a un contenido alterado de aceite, o, alternativamente, se pueden examinar las plantas en cuanto a un contenido alterado de aceite y usar luego la PCR específica de HIO30 para determinar si una planta que tiene un contenido alterado de aceite tiene un gen HIO30 mutado. El TILLING permite identificar mutaciones que pueden alterar la expresión de genes específicos o la actividad de proteínas codificadas por estos genes [véanse Colbert et al., (2001), *Plant Physiol.* 126: 480-484; y McCallum et al. (2000), *Nature Biotechnology* 18: 455-457].

En otro método, se puede usar un planteamiento de gen candidato/locus de rasgo cuantitativo (QTL; del inglés, *quantitative trait locus*) en un programa de cría asistido por marcadores para identificar, en el gen HIO30 o en ortólogos de HIO30, alelos o mutaciones que puedan conferir un contenido alterado de aceite (véanse Bert et al., *Theor. Appl. Genet.*, junio de 2003, 107 (1): 181-9; y Lionneton et al., *Genome*, diciembre de 2002, 45 (6): 1203-15). De este modo, en un aspecto más de la invención, se usa un ácido nucleico HIO30 para determinar si una planta que tiene un contenido alterado de aceite presenta una mutación en el HIO30 endógeno o tiene un alelo particular que causa el contenido alterado de aceite.

### Ejemplos

#### Ejemplo 1

Generación de plantas con un fenotipo HIO30 por transformación con una construcción de etiquetado por activación

Se generaron mutantes usando el vector de etiquetado "ACTTAG" por activación, pSKI015 (GI#6537289; D. Weigel et al., *supra*). Se usaron métodos estándares para la generación de plantas transgénicas de *Arabidopsis*, que fueron esencialmente como los descritos en la solicitud PCT publicada WO0183697. En resumen, se transformaron plantas de *Arabidopsis* T0 (Col-0) con *Agrobacterium* que portaba el vector pSKI015, que comprende DNA-T procedente del plásmido Ti de *Agrobacterium*, un gen marcador seleccionable de resistencia a herbicidas, y el elemento potenciador 35S de CaMV 4X. Se seleccionaron plantas transgénicas de la generación T1 basándose en la resistencia a herbicidas. Se recogió la semilla T2 de plantas T1 y se guardó en una colección indexada, y se accedió a una porción de la semilla T2 para la exploración.

La determinación cuantitativa del contenido de ácidos grasos de la semilla se llevó a cabo usando los métodos siguientes. Una parte alícuota de 15 a 20 semillas T2 de cada línea examinada, que contenía generalmente inserción homocigótica, tipo silvestre homocigótico y genotipos heterocigóticos en una relación estándar 1:1:2, fue concentrada en una ultramicrobalanza UMT-2 (Mettler-Toledo Co., Ohio, EE.UU.) y fue luego transferida a un vial de vidrio para extracción. Se transesterificaron las semillas completas en 500 µl de H<sub>2</sub>SO<sub>4</sub> al 2,5% en MeOH durante 3 horas a 80 °C, siguiendo el método de Browse et al. (*Biochem. J.* 235: 25-31, 1986) con modificaciones. Se incluyó una cantidad conocida de ácido heptadecanoico en la mezcla de reacción como patrón interno. Se añadieron 750 µl de agua y 400 µl de hexano a cada vial, que fue luego sacudido enérgicamente, y se dejó que se separaran las fases. Los viales de reacción se cargaron directamente en un cromatógrafo de gases para el análisis y se tomaron muestras de la fase superior de hexano mediante el automuestreador. Se usó cromatografía en fase gaseosa con detección mediante ionización por llama para separar y cuantificar los ésteres metílicos de los ácidos grasos. Se usaron cromatógrafos de gases Agilent 6890 Plus para la separación, con columnas Agilent Innowax (30 m x 0,25 mm de diámetro interno; 250 µm de espesor de película). El gas portador fue hidrógeno a un caudal constante de 2,5 ml/minuto. Se inyectó 1 µl de muestra en modo "sin división" (temperatura de entrada de 220 °C, caudal de purga de 15 ml/min en 1 minuto). Se programó el horno con una temperatura inicial de 105 °C y un tiempo inicial de 0,5 minutos, seguidos de una rampa de 60 °C por minuto hasta 175 °C y una rampa de 40 °C/minuto hasta 260 °C, con un tiempo de espera final de 2 minutos. La detección fue mediante ionización por llama (temperatura de 275 °C, caudal de combustible de 30,0 ml/min, caudal de agente oxidante de 400,0 ml/min). El control del instrumento y la recogida y el análisis de datos fueron mediante el Millennium Chromatography Management System (versión 3.2,

Waters Corporation, Milford, Massachusetts, EE.UU.). La integración y la cuantificación se llevaron a cabo automáticamente, pero posteriormente, antes de la inclusión de los resultados derivados en el estudio, todos los análisis se examinaron manualmente para verificar una correcta identificación de picos y una aceptable relación de señal a ruido.

- 5 Se reconoció que la línea ACTTAG denominada W000086431 tenía un fenotipo con alto contenido de aceite. Específicamente, el aceite constituía el 34,8% de la masa de semilla (peso/peso), en comparación con un contenido medio de aceite de 28,7% de otras líneas ACTTAG cultivadas y analizadas en las mismas condiciones (es decir, las líneas de referencia). Se llevó a cabo un nuevo análisis de la misma semilla por triplicado. El aceite constituía el 32,1% de la masa de semilla, lo que confirmaba un aumento del contenido de aceite con respecto a la referencia.
- 10 La línea W000086431 tiene tres loci de DNA-T. Un locus no estaba ligado al fenotipo, y se identificaron dos loci íntimamente ligados que se cosegregaban con el fenotipo con alto contenido de aceite. Los individuos T2 homocigóticos para los loci de alto contenido de aceite producían una semilla con un contenido de aceite del 115,4% de la referencia, los individuos T2 hemicigóticos para los loci producían una semilla con un contenido de aceite del 118,4% de la referencia. Los individuos T2 que carecían del locus HIO30 tenían contenidos de aceite del 105% de la referencia. Puesto que los homocigotos y los hemicigotos para los loci con alto contenido de aceite presentan un aumento similar en el contenido de aceite, se determinó que el fenotipo de W000086431 con alto contenido de aceite es dominante.

#### Ejemplo 2

Caracterización de la inserción de DNA-T en plantas que presentan el fenotipo con contenido alterado de aceite

- 20 Se llevaron a cabo análisis moleculares estándares, esencialmente como se describe en la solicitud PCT de patente WO0183697, para determinar el sitio de la inserción de DNA-T asociada con el fenotipo con contenido alterado de aceite. En resumen, se extrajo DNA genómico de las plantas que presentaban el fenotipo con contenido alterado de aceite. Una PCR confirmó, usando cebadores específicos para el vector pSKI015, la presencia del potenciador 35S en plantas de la línea W000086431, y un análisis por transferencia Southern permitió verificar la integración genómica del DNA-T ACTTAG.

- 25 Se usaron rescate plasmídico y/o PCR inversa para recuperar el DNA genómico que flanqueaba la inserción de DNA-T, el cual fue luego sometido a un análisis de secuencia usando una búsqueda BLASTN básica y/o una búsqueda en la base de datos de The Arabidopsis Information Resource (TAIR) (disponible en el sitio web arabidopsis.org). Se recogieron semillas de 18 familias T3 descendientes del mutante. Se determinó el contenido de aceite de las semillas de estas familias del modo descrito en el Ejemplo 1. Se determinaron los genotipos de estas familias con respecto a una inserción de DNA-T mediante una PCR específica para DNA-T, usando cebadores que son específicos para una de las correspondientes regiones genómicas. El contenido medio de aceite de las familias T3 que contenían el inserto de DNA-T en los loci 2 y 3 era mayor que el de las familias que carecían del inserto en los correspondientes loci. Por lo tanto, se concluyó que el locus 2 y/o 3 está ligado al fenotipo con alto contenido de aceite. Por contraste, el contenido medio de aceite de las familias T3 que contenían el inserto de DNA-T en el locus 1 era menor/aproximadamente igual que el de las familias que carecían del inserto en el locus correspondiente, y se concluyó que el locus 1 no está ligado al fenotipo con alto contenido de aceite.

- 35 El análisis de secuencias reveló que el codón de inicio de la secuencia de nucleótidos presentada como ID. SEC. nº 1, que fue denominada HIO30, estaba aproximadamente a 5,6 kb 5' del borde cadena arriba del inserto de DNA-T en el locus 3.

#### Ejemplo 3

Análisis de la secuencia HIO30 de *Arabidopsis*

- 45 Se llevaron a cabo análisis de secuencias con BLAST (Altschul et al., 1997, J. Mol. Biol. 215: 403-410), PFAM (Bateman et al., 1999, Nucleic Acids Res. 27: 260-262), PSORT (K. Nakai y P. Horton, 1999, Trends Biochem. Sci. 24: 34-6) y/o CLUSTAL (J. D. Thompson et al., 1994, Nucleic Acids Res. 22: 4673-4680).

El BLASTN de la ID. SEC. nº 1 frente a secuencias de GenBank identificó 6 ESTs de *Arabidopsis* procedentes de clones de cDNA derivados de la misma región genómica que At3g52260. Parece que este gen se presenta como una sola copia en *Arabidopsis*.

- 50 El BLASTN también identificó ESTs adicionales de otras especies vegetales. Se identificaron las siguientes tres ESTs de patata (*Solanum tuberosum*) como ortólogos candidatos (las calificaciones de BLAST están entre paréntesis): GI#s 13614224 (1054 4.2e-41), 18257466 (1031 4.8e-40) y 14266906 (1140 7.2e-50). Se clasificaron otras ESTs de plantas por especie y luego se ensamblaron en el menor número de cóntigos, representados por las ID. SEC. números 3-6 y descritos adicionalmente más adelante. En la mayoría de los casos, los cóntigos de EST representan regiones de codificación parciales. No obstante, la secuencia de cDNA completa del ortólogo puede ser determinada por alguien experto en técnicas de biología molecular.

## ES 2 400 907 T3

La ID. SEC. nº 3 es de tomate (*Lycopersicon esculentum*), y es un cóntigo de las secuencias siguientes: GI#s 5894121, 9503437, 5894459 y 5889410. Presenta un 56% de identidad con la ID. SEC. nº 1.

La ID. SEC. nº 4 es de soja (*Glycine max*), y es un cóntigo de las GI#s 6913831 y 5760781. Presenta un 66% de identidad con la ID. SEC. nº 1.

- 5 La ID. SEC. nº 5 es de maíz (*Zea mays*), y es un cóntigo de las GI#s 21215230, 5268759 y 22521540. Presenta un 57% de identidad con la ID. SEC. nº 1.

La ID. SEC. nº 6 es de trigo (*Triticum aestivum*), y es un cóntigo de las GI#s 19955658 y 9364987. Presenta un 62% de identidad con la ID. SEC. nº 1.

- 10 El análisis con BLASTP devolvió entradas redundantes para el producto del gen At3g52260. Los únicos otros resultados devueltos fueron numerosos productos de genes bacterianos correspondientes a pseudouridilato sintasa de RNA 23S. No se devolvieron otras secuencias de plantas.

- 15 El análisis con PFam detectó un dominio de pseudouridilato sintasa de RNA (PF00849). De acuerdo con esta predicción funcional, PSORT2 predice que el producto del gen At3g52260 es citoplásmico (44%). El gen puede tener una función reguladora en la expresión de genes implicados en el metabolismo de ácidos grasos o en rutas relacionadas.

### Ejemplo 4

#### Confirmación de la asociación fenotipo/genotipo

- 20 Un análisis por RT-PCR mostró que el gen HIO30 estaba sobreexpresado en plantas de la línea que presenta el fenotipo HIO30. Específicamente, se extrajo RNA de hojas en roseta y/o silicuas de plantas que presentan el fenotipo HIO30, recogidas en una diversidad de fases de desarrollo y reunidas. La RT-PCR se llevó a cabo usando cebadores específicos para la secuencia presentada como ID. SEC. nº 1, para otros genes pronosticados en las inmediaciones de la inserción de DNA-T, y para un gen de actina constitutivamente expresado (testigo positivo). Los resultados mostraron que las plantas que presentaban el fenotipo HIO30 sobreexpresaban el mRNA del gen HIO30, lo que indica que la expresión potenciada del gen HIO30 se correlaciona con el fenotipo HIO30.

- 25 El patrón de herencia dominante del fenotipo HIO30 se confirma por medio de un análisis genético. En general, el análisis genético implica la producción y el análisis de híbridos F1. Típicamente, se llevan a cabo cruces F1 recogiendo polen de plantas T2, que se usa para polinizar plantas de tipo silvestre. Dichos cruces se llevan a cabo tomando aproximadamente 4 flores de cada planta individual seleccionada y usando la flor T2 como dador de polen masculino, y flores de las plantas de tipo silvestre como hembras. Se realizan 4-5 cruces para un individuo de  
30 interés. Las semillas obtenidas de cruces del mismo individuo son reunidas, plantadas y cultivadas hasta la madurez como híbridos F1.

**LISTADO DE SECUENCIAS**

<110> Agrinomics LLC  
 <120> GENERACIÓN DE PLANTAS CON CONTENIDO ALTERADO DE ACEITE.  
 <130> AG03-076C-PC  
 5 <150> 60/434.763  
 <151> 18-12-2002  
 <160> 6  
 <170> PatentIn versión 3.2  
 <210> 1  
 10 <211> 5034  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <400> 1

atgccgcagg atcacgcttc gtgggatcgg aaagagctct tgaggcaaag gaaacacgat 60  
 aggcctgaac aatcttttga atccccgcct ttctgatgga gggattcgcc ttcttctcac 120  
 catgttcctc gagagttttc ttctogtttg ggatctggag acttccgcag accttcttca 180  
 ttaacacagc tcttaagatt gataggaagt gaattgatga ggttactctt caaaggagca 240  
 atttttaata ctccaggttg acggcaccag tttgtggagg agactagtca tggatacaca 300  
 tcttctcggg ccagtgcccg aatgtttgat aattataggc catcagcatc gcgtggagac 360  
 tggagatata ccaggaattg cagggatgat agagtttctg taagccaaa ggaatggaaa 420  
 tgcaatacat gggagatgag caacggatct tctagaagt ttgagaggcc atttggatt 480  
 agaaatggtc ggaggtcagt tgatgaaagg ccgctacatg cttcagatac tcattctacc 540  
 gtggtgaact ctttggatcc agccaactcg gctcattatc tggacaatga gatcagtacc 600  
 ccagtacggg ctcttaaaat taaaaatgag cataaatttt cagatcaaag gttatcactt 660  
 cttcagatc ctcatcttga atgtattagc ttgtttgaac ggccttcttc tgagaacaat 720  
 tatggcaata aggtttgttc accagcaaag caatgcaatg atttgatgta tggtcgaagg 780  
 ttagttagtg ataattcatt agatgctcca atccccaatg cagagctgga ggggacttgg 840  
 gaacaacttc gcctgaaaga cccgcaagat aacaatagtt tacatggtat caatgatata 900  
 gacggtgata ggaatgtgc aaaggagagt tctctgggag caactgggaa acttccactg 960  
 tggaatagtt ctgggagttt tgcattctcag agttcagggt ttagtcatc aagtagcttg 1020  
 aaaagcttgg gggctgttga ttccagcgat cggaagattg aggttcttcc taaaattggt 1080  
 actgtgactc aatcttcttc aggagacgct actgcctgtg ccacaactac tcatcttctc 1140  
 gaggagatga gttctagaaa gaaacaacgt ctccgggtggg gtgagggact ggcgaaatat 1200  
 gagaaaaaga aagttgatgt taacccaat gaagatggaa caacattgat ggaaaacggg 1260

ES 2 400 907 T3

ttagaggAAC tacattcgtt aaacaaaaat attgctgata aaagtccac agcagccatt 1320  
 gttccagatt atggttcccc tacaacacca tcctctgtag cttgcagttc atcaccaggg 1380  
 tttgctgata aatcatctcc gaaggctgct atagctgcta gtgatgtcag taacatgtgc 1440  
 cgttcgccta gtcccgtgtc tagtattcac cttgaacgat tcccaatcaa tatcgaggag 1500  
 ctcgataaca tctcaatgga gcgttttggc tgtttactca atgagttact tggactgat 1560  
 gattctggta caggggattc cagttctgtc caattgacat caatgaacac attacttgc 1620  
 tggaaagggtg aaattttgaa agctgtggag atgactgaat cagaaattga tctccttgaa 1680  
 aacaaacata ggacactaaa gcttgaaggT agaagacact ctcgtgttgt tggaccagT 1740  
 tcatactgtt gtgatggaga tgcaaatgtg cccaaggagc aggcttcttg tagtttgat 1800  
 cctaaggcaa cagcttcttc tgtagctaaa aactggtga gagctcctgt gcatcaggct 1860  
 ggtttagcca aggttctctc tgatgttttt gaagatagtc ctggggaagt taaacctcta 1920  
 tcccaatctt ttgccactgt tgaaagagag gaagatatac tgccatacc atctatgaag 1980  
 gcagctgttt cttcgaaaga gattaacaca cctgcttttg ccaatcagga aactattgag 2040  
 gtttctctctg ctgatgacag catggcctcc aaagaagact tgttctgggc taagttatta 2100  
 tctgccaata agaaatatgc ttgtgaatca tctggagtat tcaatcaatt gctccaaga 2160  
 gattttaatt cgtctgacaa ctcaagattc cctggcatat gtcaaacgca gtttgattct 2220  
 catgtccaag aaaaaattgc agatagggtA ggcctattga gagctagggA gaaaatttta 2280  
 ctccttcagt ttaaagcgtt tcagctctca tggaaagaag atttgatca gctagcttta 2340  
 gcaaagtacc aatcaaagtc tagcaaaaaa acagaactat atccgaatgc aaaaaatgga 2400  
 gggatctga agcttccccA atctgtacgc ctgaggttct cttcttcagc tccaagaagg 2460  
 gatagtgtag tccccacaac agagctcgta agttatatgg aaaagctact tccgggtacc 2520  
 catctaaagc cttttagaga cttttgaaa atgcctgcta tgattttgga tgagaaagag 2580  
 agggtgatgt cgaggtttat ttctagcaat ggactgattg aagatccatg tgacgttgag 2640  
 aaggaaagaa caatgattaa tccttggacc tcagaggaga aagaaatctt tctgaatttg 2700  
 ctagcaatgc atgggaagga tttcaagaag attgcttcat ctcttaccA aaagacaact 2760  
 gcggactgta ttgattacta ctacaaaaac cacaagtctg attgttttgg gaaaataaag 2820  
 aagcagcgtg cttatggtaa ggaagggag cacacctaca tgttggtcc acgaaaaaag 2880  
 tggaaacgtg agatgggggc tgectctctt gatattttag gggatgtctc cattatagca 2940  
 gcaaacgctg gaaagggtgc atcaaccagg ccgatctctt ccaaaaagat caccctaga 3000  
 ggttgacgca gtgctaattc attgcagcac gatggaaata actctgaagg gtgctctac 3060  
 agttttgatt tcccacgtaa gagaactgct ggtgcagatg ttttagctgt tggctcttg 3120

ES 2 400 907 T3

tcaccagagc agataaatc ttgcttaagg acttctgtga gctctagaga gaggtgtatg 3180  
gatcatctga agtttaatca tgtcgtaaag aaacctcgga tatctcatac tctacataat 3240  
gagaacagca atactctaca caatgagaac agcaacgaag aagatgactc atgttcggaa 3300  
gagagctgtg gggaaacagg tcctattcac tggacagatg atgagagatc tgcctttata 3360  
cagggttttt cgctttttgg caagaatfff gcttcaatat caaggtacgt cgggacaaga 3420  
tctccagatc agtgtaagg tttcttcagc aaagttcgga aatgtcttgg gttggaatct 3480  
ataaagtttg gatctggaaa tgtaagcaca tccgtaagtg ttgataatgg caatgagggg 3540  
gggaggagcg acttgaaga tccttgcct atggagagta actctggcat agtgaataat 3600  
ggagtttgtg ccaagatggg tatgaattct cctacctcac cttttaatat gaatcaggat 3660  
gggtttaatc aatcaggctc tgcaaatgtg aaagccgacc ttagtagatc agaagaagag 3720  
aatgggcaga aatatttgtg tctgaaagat gataataatc tcgtgaacaa tgcatatgtc 3780  
aatggcggtt tcccagatct agtttcagaa tctttagag atttggtaga tattaatact 3840  
gttgagagcc agtctcaggc tgccggaaa agcaagagca atgatctcat gtcaatggaa 3900  
atcgatgaag gtgtcttaac atctgtcact atatcttccg agccattgta ttgtggccta 3960  
agtgttcttt ccaatgttat tgtggaaacc cctacagaaa tctcacgaaa gggctcagga 4020  
gatcaagggtg ctacaatgcc taaathtagt tcaaagaatc aagatggagt gatgcaagct 4080  
gcaaacagaa ccagaaatc tggccttgaa cctgaaagtg caccttcagg tttcaggtag 4140  
cctgagtgtc ttcaccatgt tccgattgag gtgtgtacgg aaaaccctat aggcgtcagt 4200  
gcaccacgag gaaatccaaa ttgccatgca gagtccgagt caggaaatc tcttgttggg 4260  
caagttgacg aaacacatga cttgggttgg cccaagaaca atctggaatt ggatgggagg 4320  
cttcaggttt taggccatgt aaaccctgag cagattgggtc tactaaaagc gaccaataca 4380  
gaatcttgtc aaaatcccca gagatcagtc acccaagatc tgagcaggat aagtagatca 4440  
aaatctgatt tgatcgtaaa aacccaacgt acaggtgaag gcttctcact caccaagtgt 4500  
actagttcag ctccaaagcc tctggcagta tcccataaag agggcagatc tggcatagc 4560  
aggagccatt cgtttagttt gtctgatact gagagactcc acaagaatgg agatgtgaaa 4620  
ctgtttggtg cagtacttac tactgatgag aatggaataa acaaaaaaca caatccatgt 4680  
ggaattgtca ggtcatcatc aaccttgagc agggaccatg atacaagaca tcattacatt 4740  
aatcagcaac accttcagaa cgttcccatt acgagctacg gtttttggga tggcaacaga 4800  
attcaaaccg ggctcacatc tttgccagag tggccaagt tgcctgcaag ttgccctgaa 4860  
gcattttcca cgcactaaa gcagcaagtt ggtaacagca aagagattct ggtggatggt 4920  
aatgggtggaa ttttgagctt tggtaaagcat aacgaagata gagctgagtc ctcaagcgt 4980  
aaggatgaag gtaacatagg aggggtaaat ggtgtagcag aggcagccac gtga 5034

<210> 2  
<211> 1677  
<212> PRT  
<213> Arabidopsis thaliana

5

ES 2 400 907 T3

<400> 2

Met Pro Gln Asp His Ala Ser Trp Asp Arg Lys Glu Leu Leu Arg Gln  
 1 5 10 15

Arg Lys His Asp Arg Pro Glu Gln Ser Phe Glu Ser Pro Pro Phe Arg  
 20 25 30

Trp Arg Asp Ser Pro Ser Ser His His Val Pro Arg Glu Phe Ser Ser  
 35 40 45

Arg Leu Gly Ser Gly Asp Phe Arg Arg Pro Ser Ser Leu Thr Gln Leu  
 50 55 60

Leu Arg Leu Ile Gly Ser Glu Leu Met Arg Leu Leu Phe Lys Gly Ala  
 65 70 75 80

Ile Phe Asn Thr Gln Gly Gly Arg His Gln Phe Val Glu Glu Thr Ser  
 85 90 95

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

Arg Pro Ser Ala Ser Arg Gly Asp Trp Arg Tyr Thr Arg Asn Cys Arg  
 115 120 125

Asp Asp Arg Val Ser Val Ser Gln Lys Glu Trp Lys Cys Asn Thr Trp  
 130 135 140

Glu Met Ser Asn Gly Ser Ser Arg Ser Phe Glu Arg Pro Phe Gly Ile  
 145 150 155 160

Arg Asn Gly Arg Arg Ser Val Asp Glu Arg Pro Leu His Ala Ser Asp  
 165 170 175

Thr His Ser Thr Val Val Asn Ser Leu Asp Pro Ala Asn Ser Ala His  
 180 185 190

Tyr Leu Asp Asn Glu Ile Ser Thr Pro Val Arg Ser Leu Lys Ile Lys  
 195 200 205



ES 2 400 907 T3

Asn Glu His Lys Phe Ser Asp Gln Arg Leu Ser Leu Pro Ser Asp Pro  
 210 215 220

His Ser Glu Cys Ile Ser Leu Phe Glu Arg Pro Ser Ser Glu Asn Asn  
 225 230 235 240

Tyr Gly Asn Lys Val Cys Ser Pro Ala Lys Gln Cys Asn Asp Leu Met  
 245 250 255

Tyr Gly Arg Arg Leu Val Ser Asp Asn Ser Leu Asp Ala Pro Ile Pro  
 260 265 270

Asn Ala Glu Leu Glu Gly Thr Trp Glu Gln Leu Arg Leu Lys Asp Pro  
 275 280 285

Gln Asp Asn Asn Ser Leu His Gly Ile Asn Asp Ile Asp Gly Asp Arg  
 290 295 300

Lys Cys Ala Lys Glu Ser Ser Leu Gly Ala Thr Gly Lys Leu Pro Leu  
 305 310 315 320

Trp Asn Ser Ser Gly Ser Phe Ala Ser Gln Ser Ser Gly Phe Ser His  
 325 330 335

Ser Ser Ser Leu Lys Ser Leu Gly Ala Val Asp Ser Ser Asp Arg Lys  
 340 345 350

Ile Glu Val Leu Pro Lys Ile Val Thr Val Thr Gln Ser Ser Ser Gly  
 355 360 365

Asp Ala Thr Ala Cys Ala Thr Thr Thr His Leu Ser Glu Glu Met Ser  
 370 375 380

Ser Arg Lys Lys Gln Arg Leu Gly Trp Gly Glu Gly Leu Ala Lys Tyr  
 385 390 395 400

Glu Lys Lys Lys Val Asp Val Asn Pro Asn Glu Asp Gly Thr Thr Leu  
 405 410 415

Met Glu Asn Gly Leu Glu Glu Leu His Ser Leu Asn Lys Asn Ile Ala  
 420 425 430

Asp Lys Ser Pro Thr Ala Ala Ile Val Pro Asp Tyr Gly Ser Pro Thr  
 435 440 445

Thr Pro Ser Ser Val Ala Cys Ser Ser Ser Pro Gly Phe Ala Asp Lys

ES 2 400 907 T3

450						455											460
Ser	Ser	Pro	Lys	Ala	Ala	Ile	Ala	Ala	Ser	Asp	Val	Ser	Asn	Met	Cys		
465					470					475					480		
Arg	Ser	Pro	Ser	Pro	Val	Ser	Ser	Ile	His	Leu	Glu	Arg	Phe	Pro	Ile		
				485					490					495			
Asn	Ile	Glu	Glu	Leu	Asp	Asn	Ile	Ser	Met	Glu	Arg	Phe	Gly	Cys	Leu		
			500					505					510				
Leu	Asn	Glu	Leu	Leu	Gly	Thr	Asp	Asp	Ser	Gly	Thr	Gly	Asp	Ser	Ser		
		515					520					525					
Ser	Val	Gln	Leu	Thr	Ser	Met	Asn	Thr	Leu	Leu	Ala	Trp	Lys	Gly	Glu		
	530					535					540						
Ile	Leu	Lys	Ala	Val	Glu	Met	Thr	Glu	Ser	Glu	Ile	Asp	Leu	Leu	Glu		
545					550					555					560		
Asn	Lys	His	Arg	Thr	Leu	Lys	Leu	Glu	Gly	Arg	Arg	His	Ser	Arg	Val		
				565					570					575			
Val	Gly	Pro	Ser	Ser	Tyr	Cys	Cys	Asp	Gly	Asp	Ala	Asn	Val	Pro	Lys		
			580					585					590				
Glu	Gln	Ala	Ser	Cys	Ser	Leu	Asp	Pro	Lys	Ala	Thr	Ala	Ser	Ser	Val		
		595					600					605					
Ala	Lys	Thr	Leu	Val	Arg	Ala	Pro	Val	His	Gln	Ala	Gly	Leu	Ala	Lys		
	610					615					620						
Val	Pro	Ala	Asp	Val	Phe	Glu	Asp	Ser	Pro	Gly	Glu	Val	Lys	Pro	Leu		
625					630					635					640		
Ser	Gln	Ser	Phe	Ala	Thr	Val	Glu	Arg	Glu	Glu	Asp	Ile	Leu	Pro	Ile		
				645					650					655			
Pro	Ser	Met	Lys	Ala	Ala	Val	Ser	Ser	Lys	Glu	Ile	Asn	Thr	Pro	Ala		
			660						665				670				
Phe	Ala	Asn	Gln	Glu	Thr	Ile	Glu	Val	Ser	Ser	Ala	Asp	Asp	Ser	Met		
		675					680					685					
Ala	Ser	Lys	Glu	Asp	Leu	Phe	Trp	Ala	Lys	Leu	Leu	Ser	Ala	Asn	Lys		
	690					695					700						

ES 2 400 907 T3

Lys Tyr Ala Cys Glu Ser Ser Gly Val Phe Asn Gln Leu Leu Pro Arg  
 705 710 715 720  
 Asp Phe Asn Ser Ser Asp Asn Ser Arg Phe Pro Gly Ile Cys Gln Thr  
 725 730 735  
 Gln Phe Asp Ser His Val Gln Glu Lys Ile Ala Asp Arg Val Gly Leu  
 740 745 750  
 Leu Arg Ala Arg Glu Lys Ile Leu Leu Leu Gln Phe Lys Ala Phe Gln  
 755 760 765  
 Leu Ser Trp Lys Lys Asp Leu Asp Gln Leu Ala Leu Ala Lys Tyr Gln  
 770 775 780  
 Ser Lys Ser Ser Lys Lys Thr Glu Leu Tyr Pro Asn Ala Lys Asn Gly  
 785 790 795 800  
 Gly Tyr Leu Lys Leu Pro Gln Ser Val Arg Leu Arg Phe Ser Ser Ser  
 805 810 815  
 Ala Pro Arg Arg Asp Ser Val Val Pro Thr Thr Glu Leu Val Ser Tyr  
 820 825 830  
 Met Glu Lys Leu Leu Pro Gly Thr His Leu Lys Pro Phe Arg Asp Ile  
 835 840 845  
 Leu Lys Met Pro Ala Met Ile Leu Asp Glu Lys Glu Arg Val Met Ser  
 850 855 860  
 Arg Phe Ile Ser Ser Asn Gly Leu Ile Glu Asp Pro Cys Asp Val Glu  
 865 870 875 880  
 Lys Glu Arg Thr Met Ile Asn Pro Trp Thr Ser Glu Glu Lys Glu Ile  
 885 890 895  
 Phe Leu Asn Leu Leu Ala Met His Gly Lys Asp Phe Lys Lys Ile Ala  
 900 905 910  
 Ser Ser Leu Thr Gln Lys Thr Thr Ala Asp Cys Ile Asp Tyr Tyr Tyr  
 915 920 925  
 Lys Asn His Lys Ser Asp Cys Phe Gly Lys Ile Lys Lys Gln Arg Ala  
 930 935 940  
 Tyr Gly Lys Glu Gly Lys His Thr Tyr Met Leu Ala Pro Arg Lys Lys



ES 2 400 907 T3

Cys Pro Met Glu Ser Asn Ser Gly Ile Val Asn Asn Gly Val Cys  
 1190 1195 1200  
 Ala Lys Met Gly Met Asn Ser Pro Thr Ser Pro Phe Asn Met Asn  
 1205 1210 1215  
 Gln Asp Gly Val Asn Gln Ser Gly Ser Ala Asn Val Lys Ala Asp  
 1220 1225 1230  
 Leu Ser Arg Ser Glu Glu Glu Asn Gly Gln Lys Tyr Leu Cys Leu  
 1235 1240 1245  
 Lys Asp Asp Asn Asn Leu Val Asn Asn Ala Tyr Val Asn Gly Gly  
 1250 1255 1260  
 Phe Pro Ser Leu Val Ser Glu Ser Cys Arg Asp Leu Val Asp Ile  
 1265 1270 1275  
 Asn Thr Val Glu Ser Gln Ser Gln Ala Ala Gly Lys Ser Lys Ser  
 1280 1285 1290  
 Asn Asp Leu Met Ser Met Glu Ile Asp Glu Gly Val Leu Thr Ser  
 1295 1300 1305  
 Val Thr Ile Ser Ser Glu Pro Leu Tyr Cys Gly Leu Ser Val Leu  
 1310 1315 1320  
 Ser Asn Val Ile Val Glu Thr Pro Thr Glu Ile Ser Arg Lys Gly  
 1325 1330 1335  
 Ser Gly Asp Gln Gly Ala Thr Met Pro Lys Phe Ser Ser Lys Asn  
 1340 1345 1350  
 Gln Asp Gly Val Met Gln Ala Ala Asn Arg Thr Arg Asn Ser Gly  
 1355 1360 1365  
 Leu Glu Pro Glu Ser Ala Pro Ser Gly Phe Arg Tyr Pro Glu Cys  
 1370 1375 1380  
 Leu His His Val Pro Ile Glu Val Cys Thr Glu Asn Pro Ile Gly  
 1385 1390 1395  
 Val Ser Ala Pro Arg Gly Asn Pro Asn Cys His Ala Glu Ser Glu  
 1400 1405 1410  
 Ser Gly Asn Ser Leu Val Gly Gln Val Asp Glu Thr His Asp Leu

ES 2 400 907 T3

1415						1420									1425
Gly	Trp	Pro	Lys	Asn	Asn	Leu	Glu	Leu	Asp	Gly	Arg	Leu	Gln	Val	
1430						1435					1440				
Leu	Gly	His	Val	Asn	Pro	Glu	Gln	Ile	Gly	Leu	Leu	Lys	Ala	Thr	
1445						1450					1455				
Asn	Thr	Glu	Ser	Cys	Gln	Asn	Pro	Gln	Arg	Ser	Val	Thr	Gln	Asp	
1460						1465					1470				
Leu	Ser	Arg	Ile	Ser	Arg	Ser	Lys	Ser	Asp	Leu	Ile	Val	Lys	Thr	
1475						1480					1485				
Gln	Arg	Thr	Gly	Glu	Gly	Phe	Ser	Leu	Thr	Lys	Cys	Thr	Ser	Ser	
1490						1495					1500				
Ala	Pro	Lys	Pro	Leu	Ala	Val	Ser	His	Lys	Glu	Gly	Arg	Ser	Gly	
1505						1510					1515				
His	Ser	Arg	Ser	His	Ser	Phe	Ser	Leu	Ser	Asp	Thr	Glu	Arg	Leu	
1520						1525					1530				
His	Lys	Asn	Gly	Asp	Val	Lys	Leu	Phe	Gly	Thr	Val	Leu	Thr	Thr	
1535						1540					1545				
Asp	Glu	Asn	Gly	Ile	Lys	Gln	Lys	His	Asn	Pro	Cys	Gly	Ile	Val	
1550						1555					1560				
Arg	Ser	Ser	Ser	Thr	Leu	Ser	Arg	Asp	His	Asp	Thr	Arg	His	His	
1565						1570					1575				
Tyr	Ile	Asn	Gln	Gln	His	Leu	Gln	Asn	Val	Pro	Ile	Thr	Ser	Tyr	
1580						1585					1590				
Gly	Phe	Trp	Asp	Gly	Asn	Arg	Ile	Gln	Thr	Gly	Leu	Thr	Ser	Leu	
1595						1600					1605				
Pro	Glu	Ser	Ala	Lys	Leu	Leu	Ala	Ser	Cys	Pro	Glu	Ala	Phe	Ser	
1610						1615					1620				
Thr	His	Leu	Lys	Gln	Gln	Val	Gly	Asn	Ser	Lys	Glu	Ile	Leu	Val	
1625						1630					1635				
Asp	Val	Asn	Gly	Gly	Ile	Leu	Ser	Phe	Gly	Lys	His	Asn	Glu	Asp	
1640						1645					1650				

ES 2 400 907 T3

Arg Ala Glu Ser Ser Ser Ala Lys Asp Glu Gly Asn Ile Gly Gly  
 1655 1660 1665

Val Asn Gly Val Ala Glu Ala Ala Thr  
 1670 1675

<210> 3  
 <211> 1012  
 <212> DNA  
 5 <213> Lycopersicon esculentum

<400> 3  
 caggaaattt gtagaatttg aatttgagtt taatattttg gccagaaatt tgttgatttc 60  
 ttcaagtttt ggattaatct gctgctgatt gtttcaggaa gttgcttctg gcgatattcc 120  
 ctactccgag ttacgaatgt cagaaaaagc agagcttgca gcggtgaggg cgtactttgg 180  
 agtgctgtgg ccgcaacgta atgagggatt atcctacat gacatcgtcc gaccaccga 240  
 tgctggctctt acattgatcg aattctactt taggaagtac aaaaattcag ctcctttaca 300  
 aggttggttg cagagaattc aaaataaaca gataacaatt gatggtaaag ttgttatctt 360  
 accagatact gaactcagag caggtgctga attagtatat catcgccttc cttggagaga 420  
 acctgatgca ccttacttgc tagaagtact atttgaagat gactacttga ttgttgtaaa 480  
 taaaccttct ggtttgcaag ttcttctctg agggttatat cagcagcggg ccgtcttgac 540  
 gcaactccag tggcatgcat gtaagctgac aaccacttgc tcaggttgtc aaaaaacaca 600  
 tccagtcca gttcatcgct taggaagggg tacatcagga atactgctct gtgcaaaaac 660  
 aaagctttgt aaatctcgcc ttgcagcata ttttgctgag gggacgtcag ttgttgaaga 720  
 aaaatgcacc aactcagagt gcaatacaat gaggaagatt tgcaagatat atcgggcgct 780  
 agtaagtggg gtgatggata tggatgaggg tgatcatcaag caaccaattg gtacaattaa 840  
 atatcctgga gttgctaaag gttgtatgt tgcttctcct tcaggaagc cagctttgag 900  
 cagtgttcgc gttcttgaaa gagattcaga gagtaactgc acattggttc aggttgaaat 960  
 tcaatctgga aggccacacc aaatccgcat ccacctctct ttcataggat at 1012

<210> 4  
 <211> 737  
 10 <212> DNA  
 <213> Glycine max

<400> 4  
 cttggcctga atgcaacgac ggtctgtcct acgacgacgt cgttcgagcc tctgatgccg 60  
 gggcgacact catagagttt tactccacca agtacaagag ttctgctccc ttacaaggtt 120  
 ggttgcagcg aataaaaagt gggcagataa cagttgatgg aggagtgtt actgattcta 180

ES 2 400 907 T3

acacagtcct cagagttgga tcaaagctaa tctatcatag acttccatgg aaggagccag 240  
atgcaccgca catgatcgac gtcttatatg aagatgatga catgattgct ctaaataaac 300  
cgtctggcct gcaagttttg cctggaggtc tctaccagca gaggacaatt ttaacacagc 360  
ttcaatggga agccaacaat cagggtagct gtgaaatgca caaaaggctg cattctggtc 420  
ccgtgcatcg cctagggagg gggacttcag gaattttatt atgtgcgaag acaaaactag 480  
ccagagctcg tcttgcatct cattttgctg acggaacttc tcacgttggg ggaaaaagag 540  
atacaaagca ggaacttggg aagattgcaa agatgtaccg agctcttgtg agtgggatag 600  
ttgagaatga caagtgact attaataaac caattggaat agtaaaatat cctggtggtg 660  
ctaaaggggt atacgttgct tctgaatcag gaaaaccagc actcagtgtg gtggacattc 720  
tagagacgaa catacaa 737

<210> 5  
<211> 1308  
<212> DNA  
5 <213> Zea mays

<220>  
<221> característica\_misclánea  
<222> (661)..(665)  
<223> n es a, c, g, o t

10 <220>  
<221> característica\_misclánea  
<222> (1186)..(1186)  
<223> n es a, c, g, o t

<400> 5

gcacgaggcg gccgcggcag aaacacagat gggcgacggg gctccgccgc caggggctct 60  
atactccttc ggaacgccgt ggccggagct caatcaaggc ctcacctaca ccgacacggt 120  
ccgttgcgct gatgcggacg ccgctaccac cttgattgag ttctactcta ctaaccacaa 180  
gagctcggcg ccattgccag ggtggatcaa gaggattcgt aatgggcaga taaccgttga 240  
tggtgaagtt gtcactgac cagatatgac tctggtggat gggctctaagt tggatatca 300  
tcgttttcct tggcaggagc catttgccc gtatttgctg gaagtgtctc acgaggatga 360  
tgacatgggt gcccttaata agccttctgg cttgcaagtt ctgcctaaag gactgtttca 420  
gcagcgaact gttttagcac agcttcaatt gaaagactgg aagatggcct catcttgccg 480  
gttcaagaga aaagatgtgc agtcacatcc agtacctgtt catcgtttag ggaggggaac 540  
atcaggcctc ctgctttgtg ccaagacaaa agttgccaaa gttcgacttg catcttattt 600  
tgctgaaggt gctataaatg ctgcaaagaa aagggataaa tcagagttca gtgaagagcg 660  
15 nnnnntttca aaattttacc gagccttagt gactggcata cttgatgatg atgaggttgt 720



ES 2 400 907 T3

tgttacgcaa cctatagggg tagttcatta tcctggagtt gcagagggac tttatgcagc 780  
atggttcctca ggaaagccag caatgagcaa agtatgtggt cttgagagac ttgcacacca 840  
aatcacaca ctgggccagg ttgaaattca ttcaggacga cctcaccaaa taaggataca 900  
ccttgacatac attggggcacc cacttgtaga tgacctctc tatggtattg gtggggcacc 960  
caatthttgtt gagccagaat ctactggcac agatagttct tttgcatctg atgggggtta 1020  
tgagagacct ttgcaacctg ttcttgaga ctgtgggtat cacctgcatg cacattggct 1080  
ggttctttgc cacccaacaa ccaataagat ggtaaaaatt accgctctc ttccacaaat 1140  
tctacagaca cgggaggaac gccgcgctgc agctgagcaa accggnggtt gaacatgtag 1200  
aatcttgaaa atgtatattt cttgaagtta gcaagcagca ggttctcaca gacgttagag 1260  
ttagacactc agacatctgc tcctctgtca actgtacaac ggcgagct 1308

<210> 6  
<211> 566  
<212> DNA  
5 <213> Triticum aestivum

<400> 6  
atactccgcc gctgcaggcg aactacgtct tcgggagggc atggccggat ctcaacgaag 60  
gactctccta caccgatacg ttccgcggcg ctgatgcgga aaccaccgcc accttgacca 120  
atthctactc tgagaactac aagagctcgg cgccattgcc aggggtggatt cataggattc 180  
gcaatggaca gataacgggt gatggccaag ttgtcactga tccagatatg attctcaggg 240  
agggttctaa gttggtatat catcgcctcg catggaagga gccatttgca ccacatttgc 300  
ttcaagtgct ttatgaagat gacgacatgg tagcccttaa taagccttcc ggthttgcaag 360  
ttctgcaaaa aggactcttc cagcagcgca ctgttctagc acaacttcag tggaaagagt 420  
ggaagatgcc cccatcaagc tgctctaaga gaaaaaatgt gcagttacat cctgtacctg 480  
ttcatcgatt aggaaggggc acgtcaggtc tactgctttg tgccaagaca aagcttgcca 540  
aagttcaact tgcattttat tttgca 566

**REIVINDICACIONES**

- 5 1. Una planta transgénica que comprende un vector de transformación de plantas que comprende una secuencia de nucleótidos que codifica un polipéptido que tiene una identidad secuencial de al menos 70% con respecto a la ID. SEC. nº 2, por lo que la planta transgénica tiene un porcentaje aumentado de masa de semilla que es aceite en relación con plantas testigo.
2. Una planta transgénica que comprende una secuencia polinucleotídica heteróloga que codifica un polipéptido que tiene una identidad secuencial de al menos 70% con respecto a la ID. SEC. nº 2, por lo que la planta transgénica tiene un porcentaje aumentado de masa de semilla que es aceite en relación con plantas testigo.
- 10 3. La planta transgénica de la Reivindicación 1 ó 2, en donde dicho polipéptido tiene una identidad secuencial de al menos 80% con respecto a la ID. SEC. nº 2.
4. La planta transgénica de la Reivindicación 3, en donde dicho polipéptido tiene una identidad secuencial de al menos 90% con respecto a la ID. SEC. nº 2.
5. La planta transgénica de cualquiera de las Reivindicaciones 1 a 4, en donde dicho polipéptido comprende la secuencia de aminoácidos de ID. SEC. nº 2.
- 15 6. La planta transgénica de cualquiera de las Reivindicaciones 1 a 5, que es seleccionada del grupo que consiste en semilla de colza, soja, maíz, girasol, algodón, cacao, cártamo, palma de aceite, palma de coco, lino, ricino y cacahuete.
7. Un tejido u órgano transgénico de la planta de acuerdo con cualquiera de las Reivindicaciones 1 a 5.
8. Un tejido u órgano transgénico de acuerdo con la Reivindicación 7, que es una semilla.
- 20 9. Un método para producir aceite, que comprende cultivar la planta transgénica de cualquiera de las Reivindicaciones 1 a 5 y recuperar aceite de dicha planta.
10. Un método para producir una planta que tiene un porcentaje aumentado de masa de semilla que es aceite, método que comprende:
  - 25 a) introducir en células progenitoras de la planta un vector de transformación de plantas que comprende una secuencia de nucleótidos que codifica un polipéptido que tiene una identidad secuencial de al menos 70% con respecto a la ID. SEC. nº 2, y
  - b) cultivar las células progenitoras transformadas para producir una planta transgénica, en donde dicha secuencia polinucleotídica se expresa y dicha planta transgénica presenta un porcentaje aumentado de masa de semilla que es aceite en relación con plantas testigo.
- 30 11. El método de la Reivindicación 10, en donde dicho polipéptido tiene una identidad secuencial de al menos 80% con respecto a la ID. SEC. nº 2.
12. El método de la Reivindicación 10 u 11, en donde dicho polipéptido tiene una identidad secuencial de al menos 90% con respecto a la ID. SEC. nº 2.
- 35 13. El método de cualquiera de las Reivindicaciones 10 a 12, en donde dicho polipéptido comprende la secuencia de aminoácidos de ID. SEC. nº 2.
14. El método de cualquiera de las Reivindicaciones 10 a 13, en donde dicha planta es seleccionada del grupo que consiste en semilla de colza, soja, maíz, girasol, algodón, cacao, cártamo, palma de aceite, palma de coco, lino, ricino y cacahuete.
- 40 15. Un método para generar una planta que tiene un porcentaje aumentado de masa de semilla que es aceite, en el que se emplea la metodología TILLING, que comprende:
  - inducir mutaciones en la semilla de la planta;
  - cultivar plantas a partir de la semilla mutada; e
  - identificar una planta que tiene una mutación en la secuencia génica ID. SEC. nº 1 que da lugar a un porcentaje aumentado de masa de semilla que es aceite en comparación con las plantas que carecen de la mutación.
- 45 16. El método de la Reivindicación 15, en donde las mutaciones se inducen mediante tratamiento con EMS.
17. El método de la Reivindicación 15, en donde la identificación de la planta que tiene una mutación en la secuencia génica ID. SEC. nº 1 comprende:

## ES 2 400 907 T3

autofertilizar las plantas cultivadas a partir de la semilla mutada, para obtener una progenie;

aislar DNA de las plantas de la progenie;

utilizar una PCR para identificar mutantes en la secuencia ID. SEC. nº 1; y

examinar el contenido de aceite de las plantas que tienen mutaciones en la secuencia ID. SEC. nº 1.

5