



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 331 228**

51 Int. Cl.:

C12N 15/54 (2006.01) **C12N 15/82** (2006.01)
C12N 9/10 (2006.01) **A01H 5/00** (2006.01)
C07K 16/40 (2006.01) **C12P 7/64** (2006.01)
C11B 1/00 (2006.01) **C11C 1/00** (2006.01)
G01N 33/573 (2006.01) **C12Q 1/48** (2006.01)
A61K 38/45 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **01913791 .8**

96 Fecha de presentación : **08.02.2001**

97 Número de publicación de la solicitud: **1254238**

97 Fecha de publicación de la solicitud: **06.11.2002**

54

Título: **Nuevo gen de elongasa y proceso para la preparación de ácidos grasos poliinsaturados.**

30

Prioridad: **09.02.2000 DE 100 05 973**
17.05.2000 DE 100 23 893
19.12.2000 DE 100 63 387

45

Fecha de publicación de la mención BOPI:
28.12.2009

45

Fecha de la publicación del folleto de la patente:
28.12.2009

73

Titular/es: **BASF SE**
67056 Ludwigshafen, DE

72

Inventor/es: **Heinz, Ernst;**
Zank, Thorsten;
Zähringer, Ulrich;
Lerchl, Jens y
Renz, Andreas

74

Agente: **Carvajal y Urquijo, Isabel**

ES 2 331 228 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

Nuevo gen de elongasa y proceso para la preparación de ácidos grasos poliinsaturados.

5 **Campo de la invención**

La invención se refiere a nuevos genes de elongasa con las secuencias mencionadas en las secuencias: SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 9 o sus homólogos, derivados o análogos, un constructo génico que contiene estos genes o sus homólogos, derivados o análogos, así como a su uso. La invención se refiere a los vectores u organismos transgénicos, que contienen un gen de elongasa con la secuencia: SEQ ID NO: 1, SEQ ID NO:3 y SEQ ID NO: 9 o sus homólogos, derivados o análogos. Adicionalmente la invención se refiere al uso de las secuencias de genes de elongasa por si solas o en combinación con otras elongasas y/u otros genes de biosíntesis de ácidos grasos. El presente invento se refiere a un nuevo gen de elongasa con la secuencia: SEQ ID NR: 1 o sus homólogos, derivados o análogos.

Además la invención se refiere a un procedimiento para la preparación de ácidos grasos poliinsaturados, así como un procedimiento para la incorporación de ADN en organismos no humanos, que producen grandes cantidades de aceites y especialmente, aceites con altos contenidos en ácidos grasos poliinsaturados.

20 **Antecedentes de la invención**

Determinados productos y productos secundarios de procesos metabólicos naturales en células son útiles para un amplio espectro de industrias, incluyendo la industria alimentaria y de nutrición animal, la industria cosmética y farmacéutica. A estas moléculas que se denominan generalmente "químicos finos" pertenecen también los lípidos y ácidos grasos, entre los que figura, por ejemplo, una clase que son los ácidos grasos poliinsaturados. Ácidos grasos poliinsaturados (PUFAs) se agregan, por ejemplo, a formulaciones de alimentos para niños para producir un mayor valor nutritivo de estas formulaciones. Los PUFAs tienen, por ejemplo, un efecto positivo en el nivel de colesterol en la sangre del hombre, por lo que son apropiados para la profilaxis contra enfermedades cardíacas. Los químicos finos y los ácidos grasos poliinsaturados (polyunsaturated fatty acids, PUFAs) pueden aislarse de fuentes animales, por ejemplo, pescado o de microorganismos. Mediante la cría de estos microorganismos pueden producirse y aislarse grandes cantidades de una o varias moléculas deseadas.

Microorganismos especialmente apropiados para la preparación de PUFAs son los microorganismos tales como las cepas de *Thraustochytrium* o *Schizochytrium*, las algas, tales como *Phaeodactylum tricorutum* o especies de *Cryptocodinium*, ciliatos, tales como *Stylonychia* o *Colpidium*, hongos, tales como *Mortierella*, *Entomophthora* o *Mucor*. Por selección de cepas se ha desarrollado un número de cepas de mutantes de los microorganismos correspondientes, que producen una serie de compuestos deseados, incluyendo los PUFAs. Sin embargo, la selección de cepas con una producción mejorada de una determinada molécula representa un proceso muy complicado que requiere mucho tiempo. Otra desventaja es que con un microorganismo definido sólo se puede producir determinados ácidos insaturados o bien únicamente un determinado espectro de ácidos grasos.

Alternativamente, puede realizarse la producción de químicos finos, adecuadamente, a gran escala por vía de la producción de plantas, que están desarrolladas de tal manera, que generen los PUFAs arriba mencionados. Plantas especialmente bien apropiadas para este fin son las plantas oleaginosas, que contienen grandes cantidades de compuestos lípidos, tales como colza, cáñola, linaza, soya, girasol, borraja y onagra. Pero también otras plantas útiles, que contienen aceites o lípidos y ácidos grasos, son bien apropiadas, tal y como se mencionan en la descripción sucesiva de esta invención. Por medio del cultivo convencional se ha desarrollado una serie de mutantes de estas plantas, que producen un espectro de lípidos y ácidos grasos, cofactores y enzimas deseados. Sin embargo, la selección de nuevas especies de plantas con una producción mejorada de una determinada molécula, es un proceso complicado que requiere mucho tiempo, o es aún imposible, cuando el compuesto no existe por naturaleza en la planta correspondiente, como sucede en el caso de ácidos grasos poliinsaturados con 20 átomos de carbono y ácidos grasos poliinsaturados con 22 átomos de carbono y aquellos con cadenas de carbono más largas.

55 **Resumen de la invención**

La invención provee nuevas moléculas de ácido nucleico, que son apropiadas para la identificación y el aislamiento de genes de elongasa de la biosíntesis de PUFA y que pueden usarse para la modificación de aceites, ácidos grasos, lípidos y compuestos procedentes de lípidos y muy preferentemente, para la obtención de ácidos grasos poliinsaturados, ya que sigue existiendo una gran demanda de nuevos genes, que codifican para enzimas, que participan en biosíntesis de ácidos grasos poliinsaturados, y que permiten producir estos a escala industrial. Especialmente, existe una demanda de enzimas de la biosíntesis de ácidos grasos, que permiten la elongación de ácidos grasos poliinsaturados, preferentemente, con dos o más dobles enlaces en la molécula. Los ácidos nucleicos de la invención codifican para enzimas, que poseen esta capacidad.

Microorganismos, tales como *Phaeodactylum*, *Colpidium*, *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodinium*, así como otras algas y hongos y plantas, especialmente, plantas oleaginosas, suelen usarse en la industria para la producción a gran escala de un gran número de productos químicos finos.

Siempre que estén disponibles vectores de clonado y técnicas para la manipulación genética de los microorganismos y ciliatos arriba mencionados, tal y como se divulgan en la WO 98/01572 y la WO 00/23604, o algas y organismos similares, como por ejemplo *Phaeodactylum tricorutum*, que se describen en Falciatore *et al.* [1999, Marine Biotechnology 1 (3): 239-251] i así como Dunahay *et al.* [1995, Genetic transformation of diatoms, J. Phycol. 31:10004-1012] y la literatura allí citada, las moléculas de ácido nucleico de la invención pueden usarse para la modificación genética tecnológica de estos organismos, de modo que estos se conviertan en productores mejores o más eficientes de uno o varios productos químicos finos, especialmente, ácidos grasos insaturados. Esta producción o eficiencia mejorada de la producción de un producto químico fino puede provocarse mediante una acción directa de la manipulación de un gen de la invención o mediante una acción indirecta de esta manipulación.

Los mohos y algas son los únicos sistemas vegetales conocidos, que producen cantidades considerables de ácidos grasos poliinsaturados, como por ejemplo ácido araquidónico (= ARA) y/o ácido eicosapentaenoico (= EPA) y/o ácido docosahexaénico (= DHA). Los mohos contienen los PUFAs en lípidos de membranas, mientras que las algas, los organismos análogos a algas, especialmente, ácidos grasos insaturados. Esta producción o eficiencia mejorada de la producción de un producto químico fino puede provocarse mediante una acción directa de la manipulación de un gen de la invención o mediante una acción indirecta de esta manipulación.

Los mohos y algas son los únicos sistemas vegetales conocidos, que producen cantidades considerables de ácidos grasos poliinsaturados, como por ejemplo ácido araquidónico (= ARA) y/o ácido eicosapentaenoico (= EPA) y/o ácido docosahexaénico (= DHA). Los mohos contienen los PUFAs en lípidos de membranas, mientras que las algas, los organismos análogos a algas, especialmente, ácidos grasos insaturados. Esta producción o eficiencia mejorada de la producción de un producto químico fino puede provocarse mediante una acción directa de la manipulación de un gen de la invención o mediante una acción indirecta de esta manipulación.

Especialmente los microorganismos, tales como *Cryptocodinium cohnii* y especies de *Thraustochytrium* son microorganismos, que acumulan a PUFAs, tales como ARA, EPA o DRA en triacilglicerinas. *Thraustochytrium* también son estrechamente parecidas en sentido filogenético con las cepas de *Schizochytrium*. Aunque estos organismos no son estrechamente análogos a mohos, como por ejemplo *Physcomitrella*, a nivel de la secuencia de ADN y, especialmente, a nivel de polipéptido se encuentran similitudes en las secuencias en tal grado, que en experimentos de hibridización heterológica, comparaciones de secuencias y experimentos en los que se vale de reacciones en cadena de polimerasa pueden identificarse y aislarse y caracterizarse funcionalmente moléculas de ADN también a partir de organismos muy lejanamente parecidas en sentido evolutivo. Especialmente, pueden derivarse secuencias de consenso, que son apropiadas para la clasificación heterológica o complementación funcional y previsión de funciones genéticas en terceras especies. Por tanto, esta capacidad de identificar estas funciones, por ejemplo la previsión de la especificidad de sustrato de enzimas, puede ser de importancia significativa. Además, estas moléculas de ácido nucleico pueden servir de secuencias de referencia para la cartografía (mapping) de genomas análogos o para la derivación de primers de PCR.

Las nuevas moléculas de ácido nucleico codifican para proteínas, que se denominan en la presente: elongasas específicas de PUFA (= PSEs o en singular: PSE). Estas PSEs pueden desempeñar, por ejemplo, una función que participa en el metabolismo (por ejemplo en la biosíntesis o la degradación) de compuestos que son necesarios para la síntesis de lípidos o ácidos grasos, tales como los PUFAs, o que participan en el transporte a través de las membranas de uno o más compuestos de lípidos y/o ácidos grasos hacia adentro o hacia afuera de la célula.

En esta nueva solicitud se representa el aislamiento de tales genes de elongasa novedosos en detalle. Hemos aislado por primera vez genes de elongasa, que son apropiados para la producción de ácidos poliinsaturados de cadena larga, preferentemente, con más de dieciocho o veinte átomos de carbono en el esqueleto fundamental de carbono del ácido graso y/o con por lo menos dos dobles enlaces en la cadena de carbono, y que proceden de organismos típicos, que contienen altas cantidades de PUFAs en la fracción de triacilglicerol. Por tanto, se habla aquí en singular de un gen o una proteína de PSE o bien en plural de genes o proteínas de PSE. Otras solicitudes de patente y publicaciones conocidas no divulgan ni presentan un gen de PSE funcionalmente activo, aunque hay diferentes otras solicitudes de patente conocidas, que muestran una elongación de ácidos grasos insaturados con cadena corta o mediana (WO 98/46776 y US 5,475,099) o la elongación o la producción de ácidos grasos de cadena larga, pero que no contienen más de un doble enlace, o que proporcionan ésteres de ceras de ácido graso de cadena larga (ver: WO 98/54954, WO 96/13582, WO 95/15387). La presente invención describe el aislamiento de elongasas novedosas con propiedades nuevas. Partiendo de la secuencia mencionada en SEQ ID NO:1, pudieron encontrarse otros ácidos nucleicos que codifican para elongasas, que alongan ácidos grasos insaturados.

WO 99/64616, WO 98/46763, WO 98/46764, WO 98/46765 describen la preparación de PUFAs en plantas transgénicas y muestran la clonación y la expresión funcional de las correspondientes actividades de desaturasa, especialmente, a partir de hongos, pero no presentan ningún gen codificante de PSE y ninguna actividad funcional de PSE. La expresión de las actividades de desaturasa resulta en un desplazamiento de la composición de ácido graso en plantas transgénicas, no se observó ningún aumento del contenido en ácidos grasos insaturados. Se muestra la preparación de ácido triénico con una cadena de carbono de 18 átomos y mediante un ácido gama-linolénico se ha reivindicado esta preparación, pero hasta la fecha no se ha presentado la obtención de ácidos grasos poliinsaturados con cadena muy larga (con cadena de carbono de 20 átomos de C y más larga, así como de ácidos triénicos y tipos más altamente insaturados).

Para la obtención de PUFAs de cadena larga se deben alargar los ácidos poliinsaturados con 16 o 18 átomos de carbono mediante actividad enzimática de una elongasa en por lo menos dos átomos de carbono. La secuencia de ácido nucleico de la invención, SEQ ID NO: 1, codifica para la primera elongasa vegetal, que es capaz de alargar los ácidos grasos con 16 o bien 18 átomos de carbono con por lo menos dos dobles enlaces en el ácido graso, en por lo menos dos átomos de carbono. Después de una serie de elongación, esta actividad enzimática proporciona ácidos grasos con 20 átomos de carbono, y después de dos, tres y cuatro series de elongación produce ácidos grasos con 22, 24 o bien 26 átomos de carbono. Con la ayuda de otras de las elongasas divulgadas (SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 y SEQ ID NO: 11) también pueden sintetizarse PUFAs de cadena larga. Estos pueden aplicarse por si solos a varios o, por ejemplo, en forma aditiva a la elongasa de PUFA a partir del moho *Physcomitrella patens* (SEQ ID NO: 1) para aumentar el contenido del PUFA en un nuevo procedimiento para la producción de PUFAs. La actividad de las elongasas de la invención proporciona, preferentemente, ácidos grasos con 20 átomos de carbono con por lo menos dos dobles enlaces en la molécula del ácido graso, preferentemente, con tres o cuatro dobles enlaces, muy preferentemente, con tres dobles enlaces en la molécula de ácido graso, y/o ácidos grasos con 22 átomos de carbono con por lo menos dos dobles enlaces en la molécula de ácido graso, preferentemente, con cuatro, cinco o seis dobles enlaces, muy preferentemente, con cinco o seis dobles enlaces en la molécula. Después de que haya tenido lugar la elongación con la enzima de la invención, pueden realizarse más etapas de de-saturación, para obtener los ácidos grasos altamente de-saturados. Por tanto, los productos de las actividades de elongasa y la posible desaturación siguiente proporcionan los PUFAs preferidos con un grado de desaturación más alto, como por ejemplo ácido docosadiénico, ácido araquidónico, ácido ω 6-eicosatriendihomo- γ -linolénico, ácido eicosapentaénico, ácido ω 3-eicosatriénico, ácido ω 3-eicosatetraénico, ácido docosapentaénico ó ácido docosaheptaénico. Sustratos de la actividad enzimática de la invención son, por ejemplo, ácido taxólico; ácido 7,10,13-hexadecatriénico, ácido 6,9-octadecadiénico, ácido linólico, ácido pinolénico, ácido α - o γ -linolénico, así como ácido araquidónico o ácido estearidónico, eicosatetraénico, ácido docosapentaénico, ácido eicosapentaénico. Sustratos preferidos son el ácido linólico, ácido γ -linolénico y/o ácido α -linolénico, así como ácido araquidónico, ácido eicosatetraénico, ácido docosapentaénico, ácido eicosapentaénico. Se prefieren, sobre todo, el ácido araquidónico, ácido docosapentaénico, ácido eicosapentaénico. Los ácidos grasos con 16 o 18 átomos de carbono con por lo menos dos dobles enlaces en el ácido graso pueden alongarse mediante la actividad enzimática de la invención en forma del ácido graso libre o en forma de los ésteres, tales como fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicerina, diacilglicerina o triacilglicerina. Para la alimentación humana es especialmente importante el ácido linólico conjugado o "CLA". Por CLA se entienden, especialmente, ácidos grasos, tales como C18: 2 9 cis, 11trans o el isómero C18: 2 10trans, 12 cis que debido a los sistemas enzimáticos humanos pueden ser desaturados o bien alongados después de su absorción en el cuerpo y que pueden contribuir a efectos benéficos para la salud. Con las elongasas de la invención pueden alongarse también aquellos ácidos grasos conjugados con al menos dos dobles enlaces en la molécula, introduciendo con ello estos ácidos grasos saludables en la alimentación humana. Otros ejemplos de ácidos grasos conjugados son el ácido alfa-parinarico, ácido eleosteárico y ácido calendulárico.

Siempre que hayan vectores de clonación apropiados para ser usados en plantas y en la transformación de plantas, como aquellos publicados y citados en: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), capítulo 6/7, páginas 71-119 (1993); F. F. White, Vectors for Gene Transfer in Higher Plants; en: Transgenic Plants, vol. 1, Engineering and Utilization, Eds.: Kung y R. Wu, Academic Press, 1993, 15-38; B. Jenes *et al.*, Techniques for Gene Transfer, en: Transgenic Plants, vol. 1, Engineering and Utilization, Eds.: Kung y R. Wu, Academic Press (1993), 128 - 143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225), pueden usarse los ácidos nucleicos de la invención para la modificación tecnológica genética de un amplio espectro de plantas, de modo que estas se conviertan en un productor mejor y más eficiente o uno modificado de uno o varios productos derivados de lípidos, como por ejemplo PUFAs. Esta mejor producción o eficiencia de la producción de un producto derivado de lípidos, como PUFAs, puede generarse por acción directa de la manipulación o por acción indirecta de esta manipulación.

Hay una serie de mecanismos mediante los cuales la modificación de una proteína de PSE de la invención puede influir, directamente en el rendimiento, la producción y/o eficiencia de la producción de un producto químico fino a partir de una planta oleaginosa o un microorganismo debido a una proteína modificada. La cantidad o la actividad de la proteína o del gen de PSE pueden elevarse, de manera que se producen mayores cantidades de este compuesto de novo, puesto que los organismos carecían de esta actividad y de la capacidad para la biosíntesis antes de la incorporación del gen correspondiente. También puede ser ventajoso usar diferentes secuencias divergentes, es decir, secuencias diferentes a nivel de la secuencia de ADN.

La incorporación del gen de PSE o de varios genes de PSE en el organismo o una célula no sólo puede aumentar el flujo de biosíntesis hacia el producto final, sino que también aumentar la composición correspondiente de triacilglicerina o crearla de nuevo. Igualmente, pueden aumentarse la cantidad o la actividad de otros genes, que participan en la importación de sustancias nutritivas necesarias para la biosíntesis de uno o más productos químicos finos (por ejemplo, ácidos grasos, lípidos polares y neutros), de manera que aumente la concentración de estos precursores, co-factores o compuestos precursores dentro de las células o dentro del compartimento de almacenamiento, por lo que se incrementa más la capacidad de las células de producir PUFAs, tal como se describe en lo sucesivo. Los ácidos grasos y lípidos mismos son productos químicos finos deseados; mediante optimización de la actividad o aumento de la cantidad de una o varias PSEs, que participan en la biosíntesis de estos compuestos, o mediante destrucción de la actividad de una o varias PSEs que participan en la degradación de estos compuestos, puede ser posible incrementar el rendimiento, la producción y/o eficiencia de la producción de moléculas de ácidos grasos y lípidos a partir de plantas o microorganismos.

La mutagénesis del gen de PSE de la invención puede resultar también en una proteína de PSE con actividades modificadas, que influye directamente o indirectamente en la producción de uno o varios productos químicos finos deseados. Por ejemplo, puede incrementarse la cantidad en número o la actividad de los genes de PSE de la invención, de manera, que los desechos o productos secundarios metabólicos normales de la célula (cuya cantidad podría aumentarse debido a una sobreproducción del producto químico fino deseado) sean exportados eficientemente, antes de que destruyan a otras moléculas o procesos dentro de la célula (que podrían reducir la viabilidad de la célula) o que podrían estorbar las vías de la biosíntesis de los productos químicos finos (por lo cual se reduce el rendimiento, la producción o eficiencia de la producción del producto químico fino deseado). Además, las cantidades intracelulares relativamente altas del producto químico fino mismo pueden ser tóxicas para la célula o estorbar los mecanismos de retroacción enzimática, como por ejemplo la regulación alostérica; por ejemplo, mediante aumento de la actividad o el número de enzimas que siguen corriente abajo o de enzimas de des-envenenamiento de la vía del PUFA podría aumentarse la localización de los PUFA en la fracción de triacilglicerina; podría aumentarse la viabilidad de células de semillas, lo que a su vez resultaría en un mejor desarrollo de células en el cultivo o en semillas que producen el producto químico fino deseado. El gen de PSE de la invención también puede manipularse de manera que se produzcan las cantidades correspondientes de las diferentes moléculas de lípido y ácido graso. Esto puede tener un marcado efecto en la composición del lípido de la membrana de la célula y genera aceites nuevos, adicionalmente a la presencia de PUFAs recién sintetizados. Ya que cada tipo de lípido tiene diferentes propiedades físicas, una modificación de la composición en lípido de una membrana cambia considerablemente la fluidez de la membrana. Los cambios en la fluidez de la membrana pueden afectar el transporte de moléculas por vía de la membrana, así como afectar la integridad de la célula, factores ambos que poseen un efecto decisivo en la producción del producto químico fino. En plantas, estos cambios pueden influir también en otras características, tales como la tolerancia frente a situaciones de estrés abiótico o biótico.

La tolerancia frente a estrés biótico y abiótico es una característica general que se desea transmitir por herencia a un amplio espectro de plantas, tales como maíz, trigo, cebada, avena, triticale, arroz, centeno, soya, cacahuete, algodón, colza y canola, yuca, pimienta, girasol y tagetes plantas solanáceas, tales como papas, tabaco, berenjenas y tomates, especies de vicia, arvejas, alfalfa, plantas de arbustos (café, cacao té), especies de salix, árboles (palma aceitera, coco) y gramíneas perennes y frutas de campo de nutrición animal. Estas frutas de campo representan en otra modalidad de la invención también plantas diana preferidas para la tecnología genética. Plantas según la invención especialmente preferidas son las plantas oleaginosas, tales como soya, cacahuete, colza, canola, girasol, cártamo, árboles (palma aceitera, coco) o frutas de capote, tales como maíz, trigo, cebada, avena, triticale, arroz, centeno, alfalfa o arbustos (café, cacao, té).

Por tanto, un aspecto de la invención se refiere a moléculas de ácido nucleico aisladas (por ejemplo cADNs), que comprenden secuencias de nucleótidos, que codifican a una PSE o varias PSEs o partes biológicas activas de las mismas, o fragmentos de ácidos nucleicos, que son apropiadas como primers o sondas de hibridación para la detección o la amplificación de ácidos nucleicos codificantes de PSE (por ejemplo, ADN o mRNA). En unas modalidades especialmente preferidas, la molécula de ácido nucleico comprende una de las secuencias de nucleótido representadas o la región codificante o un complemento de una de estas secuencias de nucleótidos. En otra modalidad particularmente preferida, la molécula de ácido nucleico aislada de acuerdo con la invención comprende una secuencia de nucleótidos, que hibrida a una secuencia de nucleótidos, tal y como está representada en la secuencia SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, o una parte de la misma, o es homóloga en por lo menos aproximadamente 50%, preferentemente, por lo menos aproximadamente 60%, más preferentemente, por lo menos aproximadamente 70%, 80% o 90% y aún más preferentemente, por lo menos aproximadamente 95%, 96%, 97%, 98%, 99% o más con la misma. En otras modalidades preferidas, la molécula de ácido nucleico aislado codifica una de las secuencias de aminoácido representadas en la secuencia. El gen de PSE preferido de acuerdo con la invención también posee, preferentemente, por lo menos una de las actividades de PSE aquí descritas.

En otra modalidad preferida, la molécula de ácido nucleico aislada codifica para una proteína o una parte de la misma, y la proteína o parte de la misma contiene una secuencia de aminoácido, que es suficientemente homóloga con una secuencia de aminoácido de las secuencias: SEQ ID NO: 2, SEQ ID NO:4, y SEQ ID NO: 10, para que la proteína o la parte de proteína mantenga una actividad de PSE. La proteína o la parte de proteína, que es codificada por la molécula de ácido nucleico, mantiene, preferentemente, la capacidad de participar en el metabolismo de compuestos necesarios para la composición de membranas celulares de plantas o en el transporte de moléculas por vía de estas membranas. En una modalidad, la proteína codificada por la molécula de ácido nucleico es homóloga en por lo menos aproximadamente un 50%, preferentemente, por lo menos aproximadamente un 60% y más preferentemente, por lo menos aproximadamente un 70%, 80% o 90% y todavía más preferentemente en por lo menos aproximadamente un 95%, 96%, 97%, 98%, 99% o más con una secuencia de aminoácido de la secuencia SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO: 10. En otra modalidad preferida, la proteína es una proteína de longitud completa, que en parte es substancialmente homóloga con una secuencia total de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 (que proviene del marco de lectura abierto mostrado en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9) y puede aislarse en su longitud completa mediante métodos y experimentos conocidos por la persona técnica en la materia.

En otra modalidad preferida, la molécula de ácido nucleico aislada proviene de *Phytophthora infestans*, *Physcomitrella patens*, o *Thraustochytrium* y codifica a una proteína (por ejemplo una proteína de fusión de PSE), que contiene un dominio biológicamente activo, que es homólogo en por lo menos aproximadamente un 50% o más con una secuencia de aminoácido de la secuencia: SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y que mantiene la capacidad de participar en el metabolismo de compuestos necesarios para la composición de membranas celulares de

ES 2 331 228 T3

plantas o de participar en el transporte de moléculas por vías de estas membranas, o que posee por lo menos una de las actividades de elongación que resultan en PUPAs, tales como ARA, EPA o DHA o sus moléculas precursoras, o que posee una de las actividades indicadas en la Tabla 1, y comprende también secuencias de ácido nucleico heterológicas, que codifican un polipéptido heterológico o proteínas reguladoras.

TABLA 1

Muestra de ácido graso de cinco cepas transgénicas de levadura en % molar. Las fracciones del ácido g-linolénico se resaltan mediante cifras en negrilla, las de productos alargados están subrayadas y las del ácido g-linolénico mediante cifras en negrilla (último renglón)

Acidos grasos [% molar]	pYES2	pY2PSE1a	pY2PSE1b	pY2PSE1c	pY2PSE1d
16: 0	17,0	17,6	16,4	16,3	17,6
16: 1Δ9	28,0	26,8	28,0	27,9	25,1
18: 0	6,5	6,0	6,4	5,6	6,1
18: 1Δ ⁹	25,9	23,5	27,0	25,2	21,4
18: 3Δ ^{6,9,12}	22,6	15,7	13,2	16,4	22,8
20: 3Δ ^{8,11,14}	-	10,3	9,0	8,6	7,1
18: 3Δ ^{6,9,12}	-	39,6	40,5	34,4	23,7
Elongación					

En otra modalidad, la molécula de ácido nucleico aislada tiene una longitud de por lo menos 15 nucleótidos e híbrida en condiciones estrictas a una molécula de ácido nucleico, que contiene una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9. La molécula de ácido nucleico aislada corresponde, preferentemente, a la molécula de ácido nucleico que existe en la naturaleza. Más preferentemente, la molécula de ácido nucleico aislada codifica PSE de *Phytophthora* o PSE de *Thraustochytrium*, que existen en la naturaleza, o una parte biológicamente activa de las mismas.

Otro aspecto de la invención se refiere a vectores, por ejemplo vectores de expresión recombinantes, que contienen por lo menos una molécula de ácido nucleico de la invención, y a células huésped en las que se encuentran incorporadas estos vectores, especialmente, microorganismos, células vegetales, tejidos u órganos vegetales o plantas enteras. En una modalidad, una célula huésped de este tipo puede almacenar compuestos de químicos finos, especialmente PUFAs; para aislar el compuesto deseado, se cosechan las células. El compuesto (aceites, lípidos, triacilglicéridos, ácidos grasos) o las PSE pueden aislarse luego a partir del medio o de las células huésped, que en caso de plantas son células, que contienen o almacenan productos químicos finos, lo más preferiblemente células de tejidos de almacenamiento, tales como cáscaras de semillas, bulbos, células epidérmicas y células de semillas.

Aún otro aspecto de la invención se refiere a una planta genéticamente modificada, preferentemente, una planta oleaginosa, como las arriba mencionadas, muy preferentemente, colza, linaza o *Physcomitrella patens*, en las que se ha incorporado un gen de PSE. En una modalidad, el genoma de colza, lino o *Physcomitrella patens* se ha modificado como transgen mediante incorporación de una molécula de ácido nucleico de la invención, que codifica a una secuencia de PSE de tipo silvestre o mutada. En otra modalidad, un gen de PSE endógeno en el genoma del organismo donante, *Physcomitrella patens*, *Phytophthora infestans* o *Thraustochytrium*, ha sido modificado, es decir, destruido funcionalmente, por ejemplo, por recombinación homóloga con un gen de PSE modificado o mutagenesis y detección mediante secuencias de ADN. En una modalidad preferida, el organismo vegetal pertenece a los géneros *Physcomitrella*, *Ceratodon*, *Funaria*, colza o lino, y se prefiere *Physcomitrella*, colza o lino. En una modalidad preferida, también se usa *Physcomitrella*, colza o lino para la producción de un compuesto deseado, como lípidos y ácidos grasos, y especialmente se prefieren los PUFAs.

En aún otra modalidad, el moho *Physcomitrella patens* puede ser usado para demostrar la función de un gen de elongasa, aplicando la recombinación homóloga a base de los ácidos nucleicos descritos en la presente invención.

Otro aspecto de la presente invención se refiere a un gen de PSE aislado o una parte del mismo, por ejemplo una parte biológicamente activa del mismo. En una modalidad preferida, la PSE aislada o una parte de la misma puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en un microorganismo o una célula vegetal, o en el transporte de moléculas por vía de sus membranas. En otra modalidad preferida, la PSE aislada o la parte de la misma es suficientemente homóloga con una secuencia de aminoácido de la SEQ ID NO:

ES 2 331 228 T3

2, SEQ ID NO:4, y SEQ ID NO:10 para que la proteína o parte de la misma mantenga la capacidad de participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en microorganismos o células de plantas, o en el transporte de moléculas por vía de estas membranas.

5 La invención también proporciona una preparación aislada de una PSE. En modalidades preferidas, el gen de PSE comprende una secuencia de aminoácido de la SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10. En otra modalidad preferida, la invención se refiere a una proteína de longitud completa aislada, que es substancialmente homóloga con una secuencia de aminoácido de la SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 (que son codificadas por los marcos de lectura abiertos mostrados en SEQ ID NO: 1, SEQ ID NO:3, y SEQ ID NO:9). En otra modalidad, la proteína es por lo menos en aproximadamente un 50%, preferentemente, por lo menos aproximadamente un 60% y más preferentemente, por lo menos aproximadamente un 70%, 80% o 90% y especialmente, por lo menos aproximadamente un 95%, 96%, 97%, 98%, 99% o más homóloga con una secuencia de aminoácido de la secuencia SEQ ID NO: 2, SEQ ID NO: 4, y SEQ ID NO:10. En otras modalidades, la PSE aislada comprende una secuencia de aminoácido, que es homóloga en por lo menos aproximadamente un 50% con las secuencias de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y participa en el metabolismo de compuestos necesarios para la composición de ácidos grasos en microorganismos o una célula vegetal, o en el transporte de moléculas por estas membranas o tiene una o varias de las actividades elongantes de PUFA, y ventajosamente se efectúa la elongación de cadenas de carbono de 16 o 20 átomos de C desaturados con dobles enlaces en por lo menos dos lugares.

20 Alternativamente, la PSE aislada puede comprender una secuencia de aminoácido, que es codificada por una secuencia de nucleótido, que hibrida con una secuencia de nucleótido de la SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 por ejemplo en condiciones severas, o que es homóloga en por lo menos aproximadamente un 50%, preferentemente, por lo menos aproximadamente un 60%, más preferentemente, por lo menos aproximadamente un 70%, 80% o 90% y aún más preferentemente, por lo menos aproximadamente un 95%, 96%, 97%, 98%, 99% o más con la misma. También se prefiere que las formas de PSE preferidas también tengan una de las actividades de PSE aquí descritas.

El polipéptido de PSE o una parte biológicamente activa del mismo, puede estar ligado en forma funcionalmente apta con un polipéptido que no sea de PSE, formando así una proteína de fusión. En modalidades preferidas, esta proteína de fusión tiene una actividad, que se diferencia de la actividad de la PSE por sí sola. En otras modalidades preferidas, esta proteína de fusión participa en el metabolismo de compuestos, que son necesarios para la síntesis de lípidos y ácidos grasos, co-factores y enzimas en microorganismos o plantas, o en el transporte de moléculas por vía de estas membranas. En modalidades especialmente preferidas, la incorporación de esta proteína de fusión en una célula huésped modula la producción por la célula de un compuesto deseado. En una modalidad preferida, estas proteínas de fusión contienen también actividades de $\Delta 4$ -, $\Delta 5$ - o $\Delta 6$ -, $\Delta 8$ -, $\Delta 15$ -, $\Delta 17$ - o $\Delta 19$ -desaturasa solas o en combinación.

35 Otro aspecto de invención se refiere a un procedimiento para la obtención de un producto químico fino. Este procedimiento comprende o bien la cría de un microorganismo adecuado o el cultivo de células vegetales, tejidos vegetales, órganos vegetales o plantas enteras, que abarcan las secuencias de nucleótido de la invención de la SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o sus homólogos, derivados o análogos o un constructo génico, que comprende la SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o sus homólogos, derivados o análogos, o un vector, que comprende estas secuencias o el constructo génico que provoca la expresión de una molécula de ácido nucleico de PSE según la invención, de modo que se produzca un producto químico fino. En una modalidad preferida, el procedimiento comprende además el paso de la generación de una célula, que contiene una tal secuencia de ácido nucleico de elongasa según la invención, y una célula se transforma con una secuencia de ácido nucleico de elongasa, un constructo génico o un vector, los cuales provocan la expresión de un ácido nucleico de PSE según la invención. En otra modalidad preferida, este procedimiento abarca adicionalmente la etapa de la generación del producto químico fino a partir del cultivo. En una modalidad especialmente preferida, la célula pertenece al orden de los ciliatos, a microorganismos, tales como hongos o al reino vegetal, especialmente a las plantas oleaginosas, muy preferentemente a los microorganismos o las plantas oleaginosas.

50 Otro aspecto de la invención se refiere a un procedimiento para la modulación de la producción de una molécula por medio de un microorganismo. Estos procedimientos abarcan la reunión de una célula con una sustancia que modula la actividad de PSE o la expresión de ácido nucleico de PSE, de tal manera que se modifique una actividad asociada con la célula en relación con la misma actividad en ausencia de la sustancia. En una modalidad preferida se modulan una o dos vías metabólicas de la célula para lípidos y ácidos grasos, co-factores y enzimas o el transporte de compuestos por vía de estas membranas, de manera que se mejore el rendimiento o el grado de producción por este microorganismo de un producto químico fino deseado. La sustancia que modula la actividad de PSE puede ser una sustancia que estimula la actividad de PSE o la expresión de ácido nucleico de PSE, o que puede usarse como producto intermedio en la biosíntesis de ácido graso. Ejemplos de sustancia, que estimulan a la actividad de PSE o la expresión de ácido nucleico de PSE son, entre otras, pequeñas moléculas, PSEs activas, así como ácidos nucleicos codificantes de PSEs, que han sido incorporados en la célula. Ejemplos de sustancias que inhiben la actividad de PSE o la expresión de la actividad de PSE son, entre otras, pequeñas moléculas y/o moléculas de ácido nucleico de PSE antisentido.

Otro aspecto de la invención se refiere a un procedimiento para la modulación de los rendimientos de un compuesto deseado a partir de una célula que comprende la incorporación de un gen de PSE de tipo silvestre o mutante que se mantiene sobre un plásmido separado o se incorpora al genoma de la célula huésped, en una célula. En la integración al genoma, la integración puede ser casual o efectuarse por una recombinación tal, que el gen nativo sea sustituido por la copia integrada, por lo cual se modula por la célula la producción del compuesto deseado, o usando un gen in trans,

de modo que el gen esté ligado en forma funcional con una unidad de expresión funcional, que contiene por lo menos una secuencia que garantice la expresión de un gen y por lo menos una secuencia que garantice la poliadenilación de un gen transcrito funcionalmente.

5 En una modalidad preferida, los rendimientos se modifican. En otra modalidad preferida, se incrementa el producto químico deseado es incrementado y los compuestos indeseados molestos pueden reducirse. En una modalidad especialmente preferida, el producto químico fino deseado es un lípido o un ácido graso, un co-factor o una enzima. En una modalidad especialmente preferida, este producto químico es un ácido poliinsaturado. Más preferentemente, está seleccionado de ácido araquidónico (= ARA), ácido eicosapentaénico (= EPA) o ácido docosahexaénico (= DHA).

10

Descripción detallada de la invención

15 La presente invención proporciona ácidos nucleicos de PSE y moléculas de proteínas de que participan en el metabolismo de lípidos y ácidos grasos, co-factores de PUFA y enzimas en el moho, *Physcomitrella patens*, *Phytophthora infestans* o *Traustochytrium*, o participan en el transporte de compuestos lipófilos por membranas. Los compuestos de la invención pueden usarse para la modulación de la producción de químicos finos a partir de organismos, por ejemplo, microorganismos, tales como ciliatos, hongos, levaduras, bacterias, algas y/o plantas, tales como maíz, trigo, centeno, avena, triticale, arroz, cebada, soya, cacahuete, algodón, especies de Brassica, tales como colza, canola y nabo silvestre, pimienta, girasol, borraja, onagra y tagetes, plantas solanáceas, tales como papas, tabaco, berenjena y tomates, especies de Vicia, arvejas, mandioca, alfalfa, arbustos (café, cacao, té), especies de Salix, árboles (palma aceitera, coco) y gramíneas perennes y frutos de campo de nutrición animal, o bien directamente (por ejemplo cuando la sobreexpresión u optimización de una proteína de biosíntesis de ácido graso influye directamente en el rendimiento, la producción y/o eficiencia de la producción del ácido graso a partir de organismos modificados), o pueden tener un efecto indirecto, que resulta no obstante en un incremento del rendimiento, de la producción y/o eficiencia de la producción de un compuesto deseado o una reducción de compuestos indeseados (por ejemplo cuando la modulación del metabolismo de lípidos y ácidos grasos, cofactores y enzimas produce una alteración del rendimiento, producción y/o eficiencia de la producción o la composición de los compuestos deseados dentro de las células, lo que a su vez puede influir en la producción de uno o varios productos químicos finos).

30

Aspectos de la invención se ilustran en lo sucesivo en más detalle.

I. Productos químicos finos y PUFAs

35 El término "químico fino" es conocido en el campo técnico y comprende moléculas producidas por un organismo, que se usan en diferentes industrias, tales como, sin que estén limitados a, la industria farmacéutica, industria agrícola, industria alimentaria y la industria cosmética. Estos compuestos abarcan lípidos, ácidos grasos, cofactores y enzimas etc. (tal y como se describen, por ejemplo, en Kuninaka, A. (1996) Nucleotides and related compounds, páginas 561-612, en *Biotechnology* vol. 6, Rehm *et al.*, Eds., VCH: Weinheim y la literatura allí citada, lípidos, ácidos grasos saturados y no saturados (por ejemplo ácido araquidónico), vitaminas y cofactores (tal y como se describen en Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, Vitamins, páginas 443-613 (1996) VCH: Weinheim y la literatura allí citada; y Ong, A.S., Niki, E., & Packer, L. (1995) Nutrition, Lipids, Health and Disease Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, celebrado en 1-3 de septiembre de 1994 en Penang, Malasia, AOCS Press (1995)), enzimas y todos los otros productos químicos descritos por Gutcho (1983) en *Chemicals by Fermentation*, Noyes Data Corporation, ISBN: 0818805086, y la literatura allí citada. El metabolismo y los usos de determinados químicos finos se ilustran en lo sucesivo en más detalle.

50 Por la combinación de diferentes moléculas precursoras y enzimas de biosíntesis se obtienen diferentes moléculas de ácido graso, lo que afecta decisivamente la composición de la membrana. Puede suponerse que los PUFAs no sólo se incorporan en triacilglicerina, sino también en lípidos de membranas.

55 La síntesis de membranas es un proceso bien caracterizado, en el que participan un número de componentes, incluyendo los lípidos, como parte de la membrana biestratificada. De allí que la producción de nuevos ácidos grasos, como por ejemplo PUFAs, puede generar nuevas propiedades de funciones de membrana dentro de una célula o un organismo.

60 Las membranas de células sirven para una cantidad de funciones en una célula. Primero, y en primera línea, una membrana limita el contenido de una célula del entorno, con lo que confiere integridad a la célula. Membranas también pueden servir como barreras frente a la entrada de compuestos peligrosos o indeseados y también frente a la salida de compuestos deseados.

65 Para una descripción detallada y la participación de membranas y mecanismos implicados ver: Bamberg, E., *et al.* (1993) Charge transport of ion pumps on lipid bilayer membranes, *Q. Rev. Biophys.* 26:1-25; Gennis, R.B. (1989) Pores, Channels and Transporters, in: *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, páginas 270-322; y Nikaido, H., y Saier, H. (1992) Transport proteins in bacteria: common themes in their design, *Science* 258:936-942, y la literatura allí citada.

La síntesis de lípidos puede subdividirse en dos secciones: la síntesis de ácidos grasos y su enlace con sn-glicerina-3-fosfato, así como la adición o modificación de un grupo polar de cabeza. Lípidos usuales, que se usan en membranas, comprenden fosfolípidos, glicolípidos, esfingolípidos y fosfoglicéridos. La síntesis de ácido graso comienza con la transformación de acetil-CoA o bien en malonil-CoA por medio de la acetil-CoA-carboxilasa, o en acetil-ACP por medio de la acetiltransacilasa. Después de una reacción de condensación, se forman estas dos moléculas de producto junto con acetoacetil-ACP, que es transformado por una serie de reacciones de condensación, reducción y deshidratación, de manera que se obtiene una molécula de ácido graso saturado con la longitud de cadena deseada. La producción de los ácidos grasos no saturados a partir de estas moléculas es catalizada por desaturasas específicas, en forma aeróbica mediante oxígeno molecular, o en forma anaeróbica (para la síntesis de ácido graso en microorganismos ver F.C. Neidhardt *et al.* (1996) *E. coli* y *Salmonella*. ASM Press: Washington, D.C., páginas 612-636 y la literatura allí citada; Lengeler *et al.* (editores) (1999) *Biology of Prokaryotes*. Thieme: Stuttgart, New York, y la literatura allí citada, así como Magnuson, K., *et al.* (1993) *Microbiological Reviews* 57:522-542 y la literatura allí citada).

Los precursores para la biosíntesis de PUFA son, por ejemplo, ácido linólico y ácido linoléico. Estos ácidos grasos con 18 átomos de carbono tienen que ser alargados a 20 o bien 22 átomos de carbono, con el fin de obtener ácidos grasos de tipo de cadena eícosa y docosa. Con la ayuda de diferentes desaturasas como, por ejemplo, enzimas que presentan una actividad de $\Delta 6$ -desaturasas, $\Delta 5$ - y $\Delta 4$ -desaturasa, pueden obtenerse ácido araquidónico, ácido eicosapentaénico y ácido docosahexaénico, así como diferentes PUFAs de cadena larga, que se pueden extraer y usar para diferentes fines en aplicaciones alimentarias, de nutrición animal, cosméticas o farmacéuticas.

Para la obtención de PUFAs de cadena larga se deben elongar, como se ha indicado arriba, los ácidos grasos poliinsaturados de 16 ó 18 o bien 20 átomos de carbono en por lo menos dos átomos de carbono mediante la actividad enzimática de una elongasa. Las secuencias de ácido nucleico de la invención codifican a primeras elongasas microbianas a partir de PUFA típico en productores que contienen fracción de triacilglicerol, que son capaces de elongar los ácidos grasos de 16, 18 o bien 20 átomos de carbono con por lo menos dos dobles enlaces en el ácido graso, en por lo menos dos átomos de carbono, o que realizan esta elongación, por ejemplo, en forma secuencial sucesiva, transformando un ácido graso con 16 o bien 18 átomos de carbono en un ácido graso con 20 átomos de carbono y transformando, luego, un ácido graso de 20 átomos de carbono en uno con 22 átomos de carbono o un ácido graso con un número más alto de unidades pares de átomos de 2C. Después de una serie de elongación, esta actividad enzimática proporciona ácidos grasos con 20 átomos de carbono, y después de dos, tres y cuatro series de elongación, se obtienen ácidos grasos con 22, 24 o bien 26 átomos de carbono.

Con la elongasa según la invención se pueden sintetizar también PUFAs más largos. La actividad de las elongasas de la invención resulta, preferentemente, en ácidos grasos con 20 y/o 22 átomos de carbono con por lo menos dos dobles enlaces en la molécula del ácido graso, ácidos grasos con 20 átomos de carbono, preferentemente, con tres, cuatro o cinco dobles enlaces, muy preferentemente con tres dobles enlaces en la molécula de ácido graso, ácidos grasos con 22 átomos de carbono, preferentemente, con tres, cuatro, cinco o seis dobles enlaces, muy preferentemente, cinco o seis dobles enlaces en la molécula del ácido graso. Después de la elongación con la enzima de la invención se pueden efectuar etapas de desaturación adicionales. Por tanto, los productos de las actividades de elongasa y de la posible desaturación adicional proporcionan los PUFAs preferidos con un grado de desaturación más alto, como por ejemplo ácido docosadiénico, ácido araquidónico, ácido $\omega 6$ -eicosatriendihomo- γ -linoléico, ácido eicosapentaénico, ácido $\omega 3$ -eicosatriénico, ácido $\omega 3$ -eicosatetraénico, ácido docosapentaénico o ácido docosahexaénico. Sustratos de esta actividad enzimática de la invención son, por ejemplo, el ácido taxólico, ácido 7,10,13-hexadecatriénico, ácido 6,9-octadecadiénico, ácido linólico, ácido γ -linoléico, ácido pinoléico, ácido α -linoléico, ácido araquidónico, ácido eicosapentaénico o ácido estearidónico. Sustratos preferidos son el ácido linólico, ácido γ -linoléico y/o ácido α -linoléico o bien el ácido araquidónico, ácido eicosatetraénico o el ácido eicosapentaénico. Los ácidos grasos con 16 ó 18 ó 20 átomos de carbono con por lo menos dos doble enlaces en el ácido graso se pueden elongar mediante la actividad enzimática de la invención en forma de los ácidos grasos libres o en forma de los ésteres, como por ejemplo fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicerina, diacilglicerina o triacilglicerina.

Además, los ácidos grasos deben ser transportados a continuación a diferentes modificaciones y luego ser incorporados en el lípido de almacenamiento de triacilglicerina. Otra etapa importante en la síntesis del lípido es la transferencia de ácidos grasos a los grupos de cabeza polares, por ejemplo, mediante aciltransferasa de glicerina-ácido graso (ver Frentzen, 1998, *Lipid*, 100(4-5):161-166).

Véanse publicaciones sobre la biosíntesis de ácido graso vegetal, la desaturación, el metabolismo lipídico y el transporte por membrana de compuestos que contienen grasa, la betaoxidación, modificación de ácido graso y cofactores, el almacenamiento y el ensamblaje en triacilglicerina, inclusive la literatura allí citada, en los siguientes artículos: Kinney, 1997, *Genetic Engineering*, editor: JK Setlow, 19:149-166; Ohlrogge y Browne, 1995, *Plant Cell* 7:957-970; Shanklin y Cahoon, 1998, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611-641; Voelker, 1996, *Genetic Engineering*, editor: JK Setlow, 18:111-113; Gerhardt, 1992, *Prog. Lipid R.* 31:397-417; Gühnemann-Schäfer & Kindl, 1995, *Biochim. Biophys. Acta* 1256:181-186; Kunau *et al.*, 1995, *Prog. Lipid Res.* 34:267-342; Stymne *et al.*, 1993, in: *Biochemistry and Molecular Biology of Membrane and Storage Lipids of Plants*, editor: Murata y Somerville, Rockville, American Society of Plant Physiologists, 150-158, Murphy & Ross 1998, *Plant Journal*. 13(1):1-16.

Vitaminas, cofactores y "nutracéuticos", como por ejemplo los PUFAs, abarcan un grupo de moléculas, que ya no pueden sintetizar los animales superiores y, por tanto, deben ser ingeridos, o que no pueden sintetizar los animales superiores en grado suficiente por sí mismos y que, por tanto, deben ser ingeridos adicionalmente, no obstante

el hecho que estos pueden ser sintetizados fácilmente por otros organismos, tales como bacterias. La biosíntesis de estas moléculas en organismos que son capaces de producirlos, como por ejemplo bacterias, ya se ha caracterizado en términos generales y completos (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, páginas 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E., & Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research Asia, celebrado en 1.-3 de septiembre de 1994 in Penang, Malasia, AOCS Press, Champaign, IL X, 374 S).

Las moléculas arriba mencionadas mismas son moléculas biológicamente activas o representan productos precursores de sustancias biológicamente activas, que sirven como transferidores de electrones o productos precursores en un gran número de vías metabólicas. Estos compuestos también tienen, además de su valor nutritivo, un valor industrial significativo en calidad de colorantes, antioxidantes y catalizadores y otros auxiliares de elaboración. (Para una sinopsis de la estructura, actividad y las aplicaciones industriales de estos compuestos, ver, por ejemplo, Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, págs. 443-613, VCH: Weinheim, 1996).

Ácidos grasos poliinsaturados tienen diferentes funciones y efectos benéficos para la salud, por ejemplo, en enfermedades coronarias, mecanismos antiflogísticos, la alimentación de niños, etc. Véanse publicaciones y literatura, inclusive la literatura allí citada, en: Simopoulos, 1999, Am. J. Clin. Nutr. 70 (3. Suppl.):560-569, Takahata *et al.*, Biosc. Biotechnol. Biochem. 1998, 62(11):2079-2085, Willich y Winther, 1995, Deutsche Medizinische Wochenschrift 120(7):229 y siguientes.

II. Elementos y procedimientos de la invención

La presente invención se basa, por lo menos en parte, en el descubrimiento de nuevas moléculas, que se denominan en la presente: moléculas de ácido nucleico de PSE y moléculas de proteína de PSE, que influyen en la producción de membranas celulares en *Physcomitrella patens*, *Cryptocodinium cohnii*, *Phytophthora infestans*, *Thraustochytrium* y/o *Ceratodon purpureus* y afectan, por ejemplo, el movimiento de moléculas por vía de estas membranas. En una modalidad, las moléculas de PSE participan en el metabolismo de compuestos necesarios para la composición de membranas celulares en organismos, tales como microorganismos y plantas, o influyen indirectamente en el transporte de moléculas por vía de estas membranas. En una modalidad preferida, la actividad de las moléculas de PSE según la invención para la regulación de la producción de componentes de membrana y del transporte de membrana influye en la producción por medio de este organismo del producto químico fino deseado. En una modalidad especialmente preferida, la actividad de las moléculas de PSE de la invención está modulada, de manera que estén modulados el rendimiento, la producción y/o la eficiencia de la producción de las vías metabólicas de microorganismos o plantas, que regulan a las PSEs de la invención, y esté modificada la eficiencia del transporte de compuestos por las membranas, lo cual modula directa o indirectamente el rendimiento, la producción y/o la eficiencia de la producción de un determinado producto químico fino por microorganismos y plantas.

Los términos PSE o polipéptido de PSE abarcan proteínas, que participan en el metabolismo de compuestos necesarios para la composición de membranas celulares en organismos tales como microorganismos y plantas, o en el transporte de moléculas por estas membranas. Ejemplos de PSEs se divulgan en las SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9 y SEQ ID NO: 11 o sus homólogos, derivados o análogos. Los términos PSE o secuencia(s) de ácido nucleico de PSE comprenden secuencias de ácido nucleico que codifican a una PSE y en las que una parte puede ser una región codificante y también correspondientes regiones de secuencia 5' y 3' no transladas. Ejemplos de genes de PSE se representan en las secuencias representadas en SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 y SEQ ID NO:11. Los términos producción o productividad son conocidos en el campo técnico y se refieren a la concentración del producto de fermentación (por ejemplo, del producto químico fino deseado) que se forma en un determinado período de tiempo y en un determinado volumen de fermentación (por ejemplo, kg de producto por hora por litro). El término eficiencia de la producción comprende el tiempo necesario para alcanzar una determinada cantidad de producto (por ejemplo, el tiempo que necesita una célula para establecer un determinado volumen de caudal de un producto químico fino). El término rendimiento o rendimiento de producto/carbono es conocido en el campo técnico y abarca la eficiencia de la transformación de la fuente de carbono en el producto (a saber, el producto químico fino). Esto se suele expresar, por ejemplo, como kg de producto por kg de fuente de carbono. Aumentando el rendimiento o la producción del compuesto, se incrementa la cantidad de las moléculas generadas o de las moléculas apropiadas generadas de este compuesto en una determinada cantidad de cultivo durante un período de tiempo determinado. Los términos biosíntesis o vía de biosíntesis son conocidos en el campo técnico y abarcan la síntesis de un compuesto, preferentemente de un compuesto orgánico, por parte de una célula a partir de compuestos precursores, por ejemplo, en un proceso de varias etapas fuertemente regulado. Los términos degradación o vía de degradación son conocidos en el campo técnico y abarcan la escisión de un compuesto, preferentemente, de un compuesto orgánico, por parte de una célula en productos de degradación (dicho en términos más generales: moléculas más pequeñas o menos complejas), por ejemplo en un proceso de varias etapas y fuertemente regulado. El término metabolismo es conocido en el campo técnico y comprende la totalidad de las reacciones bioquímicas que tienen lugar en un organismo. El metabolismo de un determinado compuesto (por ejemplo el metabolismo de un ácido graso) comprende, entonces, la totalidad de las vías de biosíntesis, modificación y degradación de este compuesto en la célula, con respecto a este compuesto.

En otra modalidad, las moléculas de PSE según la invención pueden modular la producción de una deseada molécula, por ejemplo un producto químico fino, en un microorganismo o en plantas. Hay una serie de mecanismos, mediante los cuales la modificación de un determinado PSE según la invención puede afectar directamente el rendimiento, producción y/o la eficiencia de la producción de un producto químico fino a partir de una cepa de microorganismos o plantas, que contienen dicha proteína modificada. El número o la actividad de PSEs, que participan en el transporte de moléculas del producto químico fino dentro o afuera de la célula, pueden ser aumentados, de modo que se transporten mayores cantidades de este compuesto por las membranas, a partir de las cuales se pueden generar y transformar más fácilmente. Además, los ácidos grasos, triacilglicerinas y/o lípidos mismos son químicos finos deseados; optimizando la actividad o aumentando el número de uno o varios PSEs de la invención que participan en la biosíntesis de estos compuestos, o estorbando la actividad de uno o varios PSEs que participan en la degradación de estos compuestos, se puede aumentar el rendimiento, la producción y/o la eficiencia de la producción de moléculas de ácido graso y lípidos a partir de organismos, tales como microorganismos o plantas.

La mutagénesis del gen de PSE de la invención puede producir también PSEs con actividades modificadas que afectan indirectamente la producción de uno o varios químicos finos deseados a partir de microorganismos o plantas. Por ejemplo, las PSEs de la invención que participan en la exportación de productos de degradación, pueden presentar un mayor número o una mayor actividad, de modo que los desperdicios metabólicos normales de la célula (cuya cantidad puede ser aumentada debido a la sobreproducción del producto químico fino deseado) sean exportados eficientemente, antes de que pueden dañar a célula (los que disminuiría la viabilidad de la célula) o estorbar las vías de biosíntesis de los químicos finos (lo que reduciría el rendimiento, la producción o la eficiencia de la producción de un producto químico fino deseado). Las cantidades intracelulares relativamente grandes del producto químico fino deseado mismo pueden ser, además, tóxicos para la célula, de manera que al aumentar la actividad o el número de transportadores que pueden exportar estos compuestos de la célula, se puede incrementar la viabilidad de la célula en el cultivo, lo que a su vez resulta en un número más alto de células en el cultivo, que pueden producir el producto químico fino deseado. Las PSEs de la invención también se pueden manipular de tal manera, que se produzcan las cantidades correspondientes de diferentes moléculas de lípidos y ácidos grasos. Esto puede tener un marcado efecto sobre la composición de lípido en la membrana celular. Ya que cada tipo de lípido presenta diferentes propiedades físicas, una modificación de la composición del lípido de una membrana puede cambiar significativamente la fluidez de la membrana. Cambios de la fluidez de la membrana pueden afectar el transporte de moléculas por las membranas, así como la integridad de la célula, lo que respectivamente influye considerablemente en la producción del producto químico fino a partir de microorganismos y plantas en cultivos de fermentación a gran escala. Las membranas de las plantas proporcionan propiedades específicas, tales como la tolerancia frente a calor, frío, sal, sequedad, así como la tolerancia frente a patógenos, tales como bacterias y hongos. Por tanto, la modulación de los componentes de membrana puede tener un efecto significativo sobre la viabilidad de las plantas bajo los parámetros de estrés arriba indicados. Esto puede realizarse mediante modificaciones en cascadas de señales o directamente por medio de una diferente composición de la membrana (véase, por ejemplo: Chapman, 1998, Trends in Plant Science, 3 (11): 419-426) y Signalkaskaden (véase Wang 1999, Plant Physiology, 120:645-651) o la tolerancia al frío tal como se divulga en WO 95/18222.

Las secuencias de ácido nucleico aisladas de la invención están contenidas, por ejemplo, en el genoma de una cepa de *Thraustochytrium*, que está disponible de American Type Culture Collection (colección de ATCC) con el No. de cepa: ATCC26185 (*Thraustochytrium*). En el caso de *Cryptocodinium*, las secuencias de ácidos nucleico aisladas pueden conseguirse, por ejemplo, de Provasoli-Guillard National Center for Culture of Marine Phytoplankton ((CCMP) West Boothbay Harbour, ME, USA) con el No. de cepa: CCMP316. En el caso de *Phytophthora infestans*, las moléculas de ácido nucleico mencionadas están aisladas de la cepa ATCC 48886.

La secuencia de nucleótido del cADN aislado de *Physcomitrella*, *Cryptocodinium*, *Phytophthora infestans* o *Thraustochytrium* y las secuencias de aminoácido derivadas de las PSEs de *Physcomitrella patens* están representadas en las SEQ ID NO: 1 hasta SEQ ID NO: 12. Se realizaron análisis computarizados, que clasificaron o bien identificaron a estas secuencias de nucleótido como secuencias que codifican a proteínas que participan en el metabolismo de componentes de membranas celulares o proteínas que participan en el transporte de compuestos por membranas celulares o bien de la biosíntesis de PUFA. EST's con los Nos. de entrada en el banco de datos: PP001019019F, CC001042041R, PI001002014R y TC002034029R, TC002034029R-11 y TC002014093R en el banco de datos del inventor, representan las secuencias de la invención en SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 y SEQ ID NO:11. De manera análoga se designaron los polipéptidos parciales como PP001019019F, CC001042041R, PI001002014R y TC002034029R, TC002034029R-11 y TC002014093R y estos representan las secuencias de la invención en SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 y SEQ ID NO:12 según la tabla 2. La inserción completa de ESTs TC002034029R fue secuenciada y dio la SEQ ID NO: 3, que es la secuencia completa de TC002034029R. TC002034029R-11 describe una secuencia de longitud completa de una elongasa a partir de *Thraustochytrium*. La denominación de los demás clones es análoga. Además, se asignaron en forma correspondiente nombres genéticos a los diferentes clones. Las abreviaturas significan: Tc = *Thraustochytrium*, Cc = *Cryptocodinium*, Pp = *Physcomitrella patens*, Pi = *Phytophthora infestans*.

65

TABLA 2

Denominación Nombre-est	Nombre de gen	Polipéptido SEQ ID NO	Ácido nucleico SEQ ID NO
PP001019019F	Pp_PSE1	2	1
TC002034029R	Tc_PSE1	4	3
TC002014093R	Tc_PSE2	6	5
CC001042041R	Cc_PSE1	8	7
TC002034029R-11	Tc_PSE1_1	10	9
PI001002014R	Pi_PSE1	12	11

La presente invención se refiere también a proteínas con una secuencia de aminoácido, que es substancialmente homóloga con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10. Tal como se usa aquí, una proteína con una secuencia de aminoácido, que es substancialmente homóloga con una secuencia de aminoácido seleccionada, es homóloga por lo menos en aproximadamente un 50% con la secuencia de aminoácido seleccionada, por ejemplo, de la secuencia completa de aminoácido seleccionada. Una proteína con una secuencia de aminoácido, que es substancialmente homóloga con una secuencia de aminoácido seleccionada, también puede ser homóloga en por lo menos aproximadamente un 50 a 60%, preferentemente, por lo menos aproximadamente un 60 a 70% y más preferentemente, por lo menos aproximadamente un 70 a 80%, 80 a 90% ó 90 a 95% y, especialmente preferible, en por lo menos aproximadamente un 96%, 97%, 98%, 99% o más con la secuencia de aminoácido seleccionada. La PSE de la invención o bien la parte biológicamente activa o el fragmento de la misma puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en microorganismos o plantas o en el transporte de moléculas por estas membranas, o puede poseer una o varias de las actividades necesarias para la elongación de PUFAs con 16 ó 18 o bien 20 átomos de carbono, de modo que se obtengan PUFAs con 20, 22 o bien 24 átomos de carbono, así como PUFAs relacionadas.

Diferentes aspectos de la invención se describen en más detalle en las sub-secciones siguientes.

A. Moléculas de ácido nucleico aisladas

Una modalidad de la invención son ácidos nucleicos aislados, que provienen de microorganismos productores de PUFA y que codifican para polipéptidos, que alongan ácidos carboxílicos con 16 ó 18 átomos de carbono con por lo menos dos enlaces dobles en el ácido graso en por lo menos dos átomos de carbono, o bien ácidos grasos con 20 átomos de carbono con por lo menos dos dobles enlaces en el ácido graso en por lo menos dos átomos de carbono.

Otra modalidad según la invención son ácidos nucleicos aislados, que comprenden secuencias de nucleótido, que codifican para polipéptidos que alongan ácidos grasos con 16, 18 o bien 20 átomos de carbono con por lo menos dos dobles enlaces en el ácido graso, y se seleccionan del grupo que se compone de:

- a) las secuencias de ácido nucleico representadas en SEQ ID NO: 1, SEQ ID NO: 3, o SEQ ID NO:9,
- b) una secuencia de ácido nucleico, que según el código genético degenerado proviene de una secuencia representada en SEQ ID NO:1, SEQ ID NO:3, o SEQ ID NO:9, o
- c) derivados de la secuencia representada en SEQ ID NO:1, SEQ ID NO:3, o SEQ ID NO:9 que codifica para polipéptidos de la secuencia de aminoácido representada en SEQ ID NO:2, SEQ ID NO:4, o SEQ ID NO:10 y que presenta por lo menos una homología de un 50% a nivel de aminoácido, sin que esto reduzca substancialmente la acción enzimática de los polipéptidos.

Las secuencias de ácido nucleico de la invención arriba mencionadas, que actúan como elongasas con 16, 18 o bien 20 átomos de carbono, provienen de organismos, tales como ciliatos, hongos, algas, plantas o dinoflagelatos, que son capaces de sintetizar los PUFAs, preferentemente, de plantas o algas, muy preferentemente, del género *Phytophthora*, *Physcomitrella*, *Thraustochytrium* o *Schizochytrium* y lo más preferible de *Phytophthora infestans*, *Physcomitrella patens*, o *Thraustochytrium sp.*, *Schizochytrium sp.* u organismos cercanamente relacionados.

Un aspecto de la invención se refiere a moléculas de ácido nucleico aisladas, que codifican a polipéptidos de PSE o partes biológicamente activas de los mismos, así como fragmentos de ácido nucleico, que son suficientes para ser usados como sondas de hibridación o primers para la identificación o ampliación de un ácido nucleico codificante de PSE (por ejemplo ADN de PSE). El término "molécula de ácido nucleico", tal como aquí se usa, debe comprender moléculas de ADN (por ejemplo, cADN o ADN genómico) y moléculas de ARN (por ejemplo mARN), así como análogos de ADN o ARN, que son generados mediante análogos de nucleótido. Este término comprende, además, la secuencia no translada, situada en el extremo 3' y en el extremo 5' de la región codificante del gen: por lo menos aproximadamente 100 nucleótidos de la secuencia corriente arriba del extremo 5' de la región codificante y por lo

menos aproximadamente 20 nucleótidos de la secuencia corriente abajo del extremo 3' de la región codificante del gen. La molécula de ácido nucleico puede ser mono o bicatenaria, pero preferentemente, es un ADN bicatenario. Una molécula de ácido nucleico "aislada" es separada de las demás moléculas de ácido nucleico, que están presentes en la fuente natural del ácido nucleico. Un ácido nucleico "aislado" no tiene preferentemente secuencias que flanquean de manera natural al ácido nucleico en el ADN genómico del organismo del cual proviene el ácido nucleico (por ejemplo secuencias, que se encuentran en los extremos 5' y 3' del ácido nucleico). En diferentes modalidades, la molécula aislada de ácido nucleico de PSE puede contener, por ejemplo, menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb o 0,1 kb de secuencias de nucleótido, que flanquean de manera natural a la molécula de ácido nucleico en el ADN genómico de la célula de la cual proviene el ácido nucleico (por ejemplo, una célula de *Physcomitrella patens*). Una molécula de ácido nucleico "aislada", como por ejemplo una molécula de cADN, puede estar, además, esencialmente libre de otro material celular o medio de cultivo, si se prepara por técnicas recombinantes, o estar libre de precursores químicos u otros químicos, cuando se sintetiza químicamente.

Una molécula de ácido nucleico de la invención, por ejemplo una molécula de ácido nucleico con una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o una parte de la misma, puede ser aislada aplicando las técnicas biológicas moleculares estándar y la información de secuencias aquí proporcionada. También se puede identificar con la ayuda de algoritmos comparativos, por ejemplo, una secuencia homóloga o regiones de secuencia homólogas conservadas a nivel de ADN o aminoácido. Por ejemplo, se puede aislar un cADN de *Phytophthora*, *Physcomitrella*, o *Thraustochytrium* a partir de un banco de *Phytophthora*, *Physcomitrella* o *Thraustochytrium*, usando la SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 completa o parte de la misma como sonda de hibridación, así como técnicas de hibridación estándar (tal y como se describen, por ejemplo, en Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*. 2a edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Además, se puede aislar una molécula de ácido nucleico, que comprende una secuencia completa de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o una parte de la misma, por reacción en cadena de polimerasa, usando primers de oligonucleótido obtenidos a base de esta secuencia o partes de la misma, especialmente, regiones alrededor de motivos de His-Box de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 o modificaciones de la misma en aminoácidos individuales definidos (por ejemplo, se puede aislar una molécula de ácido nucleico, que comprende la secuencia completa de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o una parte la misma, por reacción en cadena de polimerasa, - usando primers de oligonucleótido, preparados a base de la misma secuencia de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9). Adicionalmente, son especialmente apropiados para este fin aquellas secuencias parciales, que están representadas en la Figura 10. Por ejemplo, se puede aislar mRNA a partir de células (por ejemplo, por el procedimiento de extracción de tiocianato de Chirgwin *et al.* (1979) *Biochemistry* 18: 5294-5299) y preparar cADN mediante transcriptasa reversora (por ejemplo Moloney-MLV-Reversetranscriptasa, disponible de Gibco/BRL, Bethesda, MD, o AMV-Reverse-transcriptasa, disponible de Seikagaku America, Inc., St. Petersburg, FL). Primers de oligonucleótido sintéticos para la amplificación mediante la reacción en cadena de polimerasa se pueden obtener a base de una de las secuencias de nucleótido mostrada en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o con la ayuda de las secuencias de aminoácido representadas en la Figura 10. Un ácido nucleico de la invención se puede amplificar usando cADN o alternativamente ADN genómico como matriz y primers de oligonucleótido apropiados según las técnicas de amplificación de PCR estándar. El ácido nucleico amplificado de esta manera puede clonarse en un vector apropiado y caracterizarse mediante análisis secuencial de ADN. Oligonucleótidos, que corresponden a una secuencia de nucleótido de PSE se pueden preparar, por ejemplo, mediante procedimientos de síntesis estándar, por ejemplo con un aparato de síntesis de ADN automático.

El cADN presentado en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 comprende secuencias codificantes de PSEs, (a saber, "la región codificante") así como secuencias de 5' no transladas y secuencias de 3' no transladas. Alternativamente, la molécula de ácido nucleico puede comprender únicamente la región codificante de una de las secuencias de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o puede contener fragmentos genómicos enteros, aislados a partir de un ADN genómico.

Las SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 se identifican mediante el mismo número de entrada EST, como SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, de manera que son fáciles de correlacionar.

En otra modalidad preferida, una molécula de ácido nucleico aislada, de acuerdo con la invención, comprende una molécula de ácido nucleico, que es un complemento de una de las secuencias de nucleótido mostradas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o de una parte de la misma. Una molécula de ácido nucleico, que es complementaria a una de las secuencias de nucleótido representadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 es entonces suficientemente complementaria cuando es capaz de hibridar con una de las secuencias indicadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, por lo cual aparece un dúplex estable.

Homólogos de las nuevas secuencias de ácido nucleico de elongasa con la secuencia SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 son, por ejemplo, las variantes alélicas con una homología de por lo menos aproximadamente un 50 a 60%, preferentemente, por lo menos aproximadamente un 60 a 70%, mas preferentemente, por lo menos aproximadamente un 70 a 80%, 80 a 90% o 90 a 95% y aún más preferentemente, por lo menos aproximadamente un 95%, 96%, 97%, 98%, 99% o más con una secuencia de nucleótido representada en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o sus homólogos, derivados o análogos o partes de la misma. En otra modalidad preferida, una molécula de ácido nucleico aislada, de la invención, comprende una secuencia de nucleótido que hibrida a una secuencia de nucleótido representada en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o una parte de la misma, por ejemplo en condiciones severas. Variantes alélicas comprenden, especialmente, las variantes funcionales, que

ES 2 331 228 T3

se pueden obtener por delección, inserción o sustitución de nucleótidos de/en la secuencia representada en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 pero la intención es que se preserve, ventajosamente, la actividad enzimática de las proteínas sintetizadas procedentes de las mismas con respecto a la inserción de uno o varios genes. Proteínas que aún poseen la actividad enzimática de las elongasas son proteínas con por lo menos 10%, preferentemente 20%,
5 muy preferentemente 30%, especialmente 40% de la actividad enzimática original, en comparación con la proteína codificada por SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10. Elongasas, que presentan aún las actividades antes mencionadas, son elongasas cuya actividad enzimática no está reducida substancialmente.

10 Homólogos de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 son, por ejemplo, también los homólogos bacterianos, de hongos y vegetales, secuencias acortadas, ADN o ARN monocatenarios de la secuencia de ADN codificante y no codificante.

15 Homólogos de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 son también derivados, como por ejemplo, variantes de promotor. Los promotores corriente arriba de las secuencias indicadas de nucleótido pueden modificarse por uno o varios intercambios de nucleótidos mediante inserción/inserciones y/o delección/delecciones, pero sin que esto perjudique la funcionalidad o actividad de los promotores. Además, es posible, que la actividad de los promotores esté incrementada por modificación de su secuencia, o que esté completamente sustituida por promotores más activos, aún de organismos heterólogos.

20 Adicionalmente, la molécula de ácido nucleico de la invención puede comprender únicamente una parte de la región codificante de una de las secuencias en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, por ejemplo un fragmento, que puede usarse como sonda o primer, o un fragmento, que codifica una sección biológicamente activa de una PSE. Las secuencias de nucleótido determinadas del clonado del gen de PSE de *Physcomitrella patens*, *Phytophthora infestans* y *Thraustochytrium* permiten generar sondas y primers que están conformados para identificar y/o clonar
25 homólogos de PSE en otros tipos de células y otros organismos, así como homólogos de PSE a partir de otros mohos o especies relacionadas. Las sondas/los primers comprenden, generalmente, oligonucleótido sustancialmente purificado. El oligonucleótido comprende, generalmente, una región de secuencia de nucleótido, que hibrida bajo condiciones severas a por lo menos aproximadamente 12, preferentemente, aproximadamente 16, más preferentemente, aproximadamente 25, 40, 50 ó 75 nucleótidos sucesivos de una cadena de sentido de una de las secuencias indicadas en SEQ
30 ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, de una cadena antisentido de una de las secuencias indicadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o sus homólogos, derivados o análogos o mutantes de procedencia natural. Primers a base de una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 pueden ser usados en reacciones de PCR para el clonado de homólogos de PSE. Sondas a base de las secuencias de nucleótido de PSE se pueden usar para la detección de transcritos o secuencias genómicas, que codifican a proteínas idénticas u homólogos.
35 En las modalidades preferidas, la sonda comprende, además, un grupo de marcación ligado a la misma, por ejemplo un radioisótopo, un compuesto fluorescente, una enzima o un cofactor de enzima. Estas sondas se pueden usar como parte de un kit de ensayo para marcadores genómicos con el fin de identificar células, que expresan mal a una PSE, por ejemplo, midiendo una cantidad de un ácido nucleico que codifica PSE en una muestra de célula, por ejemplo midiendo el nivel de mRNA de PSE, o para determinar si un gen de PSE genómico está mutado o deletado.

40 En una modalidad, la molécula de ácido nucleico de la invención codifica a una proteína o parte de la misma, que comprende una secuencia de aminoácido, la cual es suficientemente homóloga con una secuencia de aminoácido de la SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 para que la proteína o la parte de la misma mantengan la capacidad de participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en
45 microorganismos o plantas o en el transporte de moléculas por estas membranas. Tal como aquí se usa, el término “suficientemente homólogo” se refiere a proteínas o partes de ellas, cuyas secuencias de aminoácido presentan un número mínimo de residuos de aminoácido idénticos o equivalentes (por ejemplo, un residuo de aminoácido con una cadena lateral similar, como un residuo de aminoácido en una de las secuencias de SEQ ID NO:2 a 12) con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, de manera que la proteína o la
50 parte de la misma puedan participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en microorganismos o plantas, o en el transporte de moléculas por estas membranas. Componentes proteicos de estas vías metabólicas para componentes de membrana o sistemas de transporte de membranas pueden, tal como está descrito aquí, desempeñar un papel en la producción y secreción de uno o varios productos químicos finos. Ejemplos de estas actividades también están descritos en la presente. Por tanto, la “función de una PSE” contribuye
55 directa o indirectamente al rendimiento, la producción y/o la eficiencia de la producción de uno o varios químicos finos. Ejemplos de especificidades de sustratos PSE de la actividad catalítica se indican en la Tabla 1.

En otra modalidad, los derivados de molécula de ácido nucleico de la invención codifican a proteínas con una homología de por lo menos aproximadamente un 50 a 60%, preferentemente, por lo menos aproximadamente un 60 a
60 70% y más preferentemente, por lo menos aproximadamente un 70 a 80%, 80 a 90%, 90 a 95% y, especialmente, por lo menos aproximadamente un 96%, 97%, 98%, 99% o más con una secuencia de aminoácido completa de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10. La homología de la secuencia de aminoácido se determinó por toda la región de secuencia con el programa: PileUp (J. Mol. Evolution., 25, 351-360, 1987, Higgins *et al.*, CABIOS, 5, 1989:151-153) o BESTFIT o GAP (Henikoff, S. y Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.)

Partes de proteínas que son codificadas por las moléculas de ácido nucleico de PSE de la invención, son preferentemente partes biológicamente activas de una de las PSEs. Tal como se usa aquí, el término “parte biológicamente

ES 2 331 228 T3

activa de una PSE” significa una sección, por ejemplo un dominio/un motivo de una PSE, que puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en microorganismos o plantas, o en el transporte de moléculas por estas membranas, o que presenta una de las actividades indicadas en la Tabla 1. Para determinar si una PSE o una parte biológicamente activa de la misma puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en microorganismos o plantas, o en el transporte de moléculas por estas membranas, se puede realizar un ensayo de la actividad enzimática. Estos procedimientos de ensayo, que se describen detalladamente en el ejemplo 8 de la sección de ejemplos, son conocidos por la persona técnica en la materia.

Fragmentos de ácido nucleico adicionales que codifican a secciones biológicamente activas de una PSE, se pueden obtener por aislamiento de una parte de las secuencias en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, expresión de la sección codificada de la PSE o del péptido (por ejemplo por expresión recombinante *in vitro*) y determinación de la actividad de la parte codificada de la PSE o del péptido.

La invención comprende, adicionalmente, moléculas de ácido nucleico, que se diferencian de una de las secuencias de nucleótido representadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 (o partes de ella) debido al código genético degenerado, y codifican por tanto la misma PSE que aquella que es codificada por las secuencias de nucleótido representadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9. En otra modalidad, la molécula de ácido nucleico aislada de la invención tiene una secuencia de nucleótido que codifica a una proteína con una de las secuencias representadas en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10. En otra modalidad, la molécula de ácido nucleico de la invención codifica a una proteína de PSE de longitud completa, que es sustancialmente homóloga con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 (que es codificada por un marco de lectura abierto representado en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 y que se puede identificar y aislar por métodos convencionales.

Adicionalmente, a las secuencias de nucleótido de PSE representadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 la persona técnica en la materia reconoce, que pueden existir polimorfismos de secuencias de ADN, que resultan en modificaciones en las secuencias de aminoácido de las PSEs, dentro de una población (por ejemplo la población de *Physcomitrella*, *Phytophthora* o *Thraustochytrium*). Estos polimorfismos genéticos en el gen de PSE pueden existir entre individuos dentro de una población debido a variaciones naturales. Tal como se usa aquí, los términos “gen” y “gen recombinante” significan moléculas de ácido nucleico con un marco de lectura abierto, que codifica a una PSE, preferentemente, una PSE de *Phytophthora*, *Physcomitrella* o *Thraustochytrium*. Estas variantes naturales producen, generalmente, una variación de 1 a 5% en la secuencia de nucleótido del gen de PSE. Todas y cualquiera de las variaciones de estos nucleótidos y los polimorfismos de aminoácido en la PSE resultantes, que se deben a variaciones naturales y que no modifican la actividad funcional de las PSEs, están comprendidas en su totalidad en la invención.

Moléculas de ácido nucleico, que corresponden a las variantes naturales, y que no son homólogos, derivados y análogos de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* de cADN de *Phytophthora*, *Physcomitrella*, o *Thraustochytrium*, se pueden aislar a base de su homología con el ácido nucleico de PSE de *Phytophthora*, *Physcomitrella* o *Thraustochytrium* divulgadas en la presente, usando cADN de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* o una parte del mismo, como sonda de hibridación según las técnicas de hibridación estándar bajo condiciones severas de hibridación. En otra modalidad, una molécula de ácido nucleico aislada de la invención tiene una longitud de por lo menos 15 nucleótidos e hibrida bajo condiciones severas con la molécula de ácido nucleico que comprende una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9. En otras modalidades, el ácido nucleico tiene una longitud de por lo menos 25, 50, 100, 250 o más nucleótidos. El término “hibrida bajo condiciones severas”, tal como aquí se usa, describe condiciones de hibridación y lavado, en las cuales las secuencias de nucleótido que tienen una homología de por lo menos un 60% entre sí, suelen quedar hibridadas la una a la otra. Las condiciones son preferentemente tales, que las secuencias que son homólogas entre sí en por lo menos aproximadamente un 65%, más preferentemente por lo menos aproximadamente un 70% y aún más preferentemente por lo menos aproximadamente un 75% o más, suelen quedar hibridadas. Estas condiciones severas son conocidas a la persona técnica en la materia y se encuentran descritas en: *Current Protocols in Molecular Biology*, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6. Un ejemplo preferido, no limitativo de condiciones de hibridación severas son hibridaciones en 6 x cloruro sódico/citrato sódico (sodium chloride/sodiocitrate = SSC) a aproximadamente 45°C, seguidas de una o varias etapas de lavado en 0,2 x SSC, 0,1% SDS a 50 hasta 65°C. La persona técnica en la materia sabe que estas condiciones de hibridación difieren según el tipo del ácido nucleico y, cuando están presentes solventes orgánicos, por ejemplo, con respecto a la temperatura y la concentración del búfer. La temperatura varía, por ejemplo, en “condiciones de hibridación estándar” según el tipo de ácido nucleico entre 42°C y 58°C en búfer acuoso con una concentración de 0,1 a 5 x SSC (pH 7,2). Cuando está presente un solvente orgánico en el búfer arriba mencionado, por ejemplo un 50% de formamida, entonces asciende la temperatura en condiciones estándar a aproximadamente 42°C. Se prefieren condiciones de hibridación para híbridos de ADN:ADN, por ejemplo, de 0,1 x SSC y 20°C a 45°C, preferentemente, entre 30°C y 45°C. Son preferidas las condiciones de hibridación para híbridos de ADN:ARN, por ejemplo, de 0,1 x SSC y 30°C a 55°C, preferentemente, entre 45°C y 55°C. Las temperaturas de hibridación arriba mencionadas son determinadas, por ejemplo, para un ácido nucleico con una longitud de aproximadamente 100 bp (= pares de bases) y un contenido de G + C del 50%, en ausencia de formamida. La persona técnica en la materia sabe cómo determinar las condiciones de hibridación necesarias por medio de manuales, como los mencionados previamente o de los siguientes libros de texto: Sambrook *et al.*, “Molecular Cloning”, Cold Spring Harbor Laboratory, 1989; Hames y Higgins (Eds.) 1985, “Nucleic Acids Hybridization: A Practical Approach”, IRL Press at Oxford University Press, Oxford Brown (editor) 1991, “Essential Molecular Biology: A Practical Approach”, IRL Press at Oxford University Press, Oxford.

ES 2 331 228 T3

Una molécula de ácido nucleico según la invención aislada que bajo condiciones severas hibrida a una secuencia de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 corresponde, preferentemente, a una molécula de ácido nucleico que existe en la naturaleza. Tal como aquí se usa, una molécula de ácido nucleico “que existe en la naturaleza” se refiere a una molécula de ARN o de ADN con una secuencia de nucleótido, que existe en la naturaleza (por ejemplo que codifica a una proteína natural). En una modalidad, el ácido nucleico codifica a una PSE de *Phycomitrella patens*, *Phytophthora infestans* o *Thraustochytrium* que existe en la naturaleza.

Adicionalmente a las variantes de la secuencia de PSE que existen en la naturaleza, que pueden estar presentes en la población, la persona técnica en la materia reconoce, además, que también se pueden insertar modificaciones mediante mutación en una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 lo que resulta en modificaciones de la secuencia de aminoácido de la PSE codificada, sin que esto perjudique la funcionalidad de la proteína de PSE. Por ejemplo, se pueden preparar sustituciones de nucleótido que en radicales de aminoácido “no esenciales” resultan en sustituciones de aminoácido, en una secuencia de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9. Un residuo de aminoácido “no esencial” es un residuo que puede ser modificado en una secuencia de tipo silvestre de una de las PSEs (SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10), sin que esto modifique la actividad de la PSE, mientras que un residuo de aminoácido “esencial” es indispensable para la actividad de la PSE. Sin embargo, otros residuos de aminoácido (por ejemplo aquellos, que no están conservados o tan sólo semiconservados en el dominio con actividad de PSE) pueden ser no esenciales para la actividad, por lo que probablemente pueden ser modificados, sin que esto modifique la actividad de PSE.

Por tanto, otro aspecto de la invención se refiere a moléculas de ácido nucleico, que codifican a PSEs, que contienen residuos de aminoácido modificados, que no son esenciales para la actividad de PSE. Estas PSEs se distinguen en la secuencia de aminoácido de una secuencia en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y mantienen, no obstante, al menos una de las actividades de PSE aquí descritas. La molécula de ácido nucleico aislada comprende en una modalidad una secuencia de nucleótido, que codifica a una proteína, y la proteína comprende una secuencia de aminoácido con una homología de por lo menos aproximadamente un 50% con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en *Phytophthora*, *Phycomitrella* o *Thraustochytrium*, o en el transporte de moléculas por estas membranas. La proteína codificada por la molécula de ácido nucleico tiene una homología de preferentemente por lo menos aproximadamente un 50 a 60% con una de las secuencias en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, más preferentemente de por lo menos aproximadamente un 60 a 70 % con una de las secuencias en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, aún más preferentemente, de por lo menos aproximadamente un 70 a 80%, 80 a 90%, 90 a 95% con una de las secuencias en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y especialmente, de por lo menos aproximadamente un 96%, 97%, 98% o 99% con una de las secuencias en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10.

Para la determinación de una homología porcentual de dos secuencias de aminoácido (por ejemplo una de las secuencias de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y una forma mutada de la misma) o de dos ácidos nucleicos, se escriben las secuencias una debajo de otra para efectos de una comparación óptima (por ejemplo se pueden introducir espacios blancos en la secuencia de una proteína o de un ácido nucleico, para alcanzar un alineamiento óptimo con la otra proteína o el otro ácido nucleico). Entonces, los residuos de aminoácido o nucleótidos en las correspondientes posiciones de aminoácido o nucleótido se comparan. Cuando una posición en una secuencia (por ejemplo una de las secuencias de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10) es ocupada por el mismo residuo de aminoácido o el mismo nucleótido que se encuentra en la posición correspondiente en otra secuencia (por ejemplo una forma mutada de la secuencia seleccionada de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10) entonces las moléculas son homólogas en esta posición (es decir, la “homología” de aminoácido o ácido nucleico, tal como aquí se usa, corresponde a la “identidad” de aminoácido y ácido nucleico). La homología porcentual entre ambas secuencias es una función del número de posiciones idénticas, que ambas secuencias tienen en común (a saber, % de homología = número de posiciones idénticas/número total de posiciones x 100).

Una molécula de ácido nucleico aislada, que codifica a una PSE, que es homóloga con una secuencia de proteína de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, puede ser generada por incorporación de una o varias sustituciones, adiciones o deleciones de nucleótido en una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 de modo que se incorporen una o varias sustituciones, adiciones o deleciones de aminoácido en la proteína codificada. Mutaciones se pueden incorporar en una de las secuencias de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 mediante técnicas estándar, tales como mutagénesis específica de posición y mutagénesis generada por PCR. Preferentemente, se realizan sustituciones de aminoácido conservadora en uno o varios residuos de aminoácido no esenciales previstos. En la “sustitución de aminoácido conservadora” se intercambia un residuo de aminoácido por otro residuo de aminoácido con una cadena lateral similar. En el campo técnico están definidas familias de residuos de aminoácidos con cadenas laterales similares. Estas familias comprenden aminoácidos con cadenas laterales básicas (por ejemplo lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (por ejemplo glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares, (por ejemplo alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales beta-ramificadas (por ejemplo treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo tirosina, fenilalanina, triptófano, histidina).

De esta manera, un residuo de aminoácido no esencial previsto en una PSE es sustituido, preferentemente, por otro residuo de aminoácido de la misma familia de cadenas laterales. Alternativamente, en otra modalidad se puede

introducir la mutación en forma aleatoria por toda la secuencia o una parte de la secuencia codificante de PSE, por ejemplo por mutagénesis de saturación, y los mutantes resultantes se pueden revisar con respecto a la actividad de PSE aquí descrita, con el fin de identificar mutantes que mantienen la actividad de PSE. Después de la mutagénesis de una de las secuencias de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 se puede expresar la proteína codificada en forma recombinante, y se puede determinar la actividad de la proteína, por ejemplo, aplicando los ensayos aquí descritos (véase la sección de ejemplos).

Adicionalmente a las moléculas de ácido nucleico, que codifican a las PSEs arriba descritas, otro aspecto de la invención se refiere a moléculas de ácido nucleico aisladas, que son moléculas “antisentido” a las mismas. Un ácido nucleico “antisentido” comprende una secuencia de nucleótido, que es complementaria a un ácido nucleico “sentido”, que codifica a una proteína, por ejemplo que es complementaria a la cadena codificante de una molécula de cADN bicatenaria o complementaria a una secuencia de mRNA. Por tanto, un ácido nucleico antisentido puede ligarse por vía de enlaces de puentes de hidrógeno con un ácido nucleico sentido. El ácido nucleico antisentido puede ser complementario a una cadena completa de PSE codificante o solamente a una parte de la misma. En una modalidad, la molécula de ácido nucleico antisentido es “antisentido” a una “región - codificante” de la cadena codificante de una secuencia de nucleótido, que codifica a una PSE. El término “región codificante” se refiere a la región de la secuencia de nucleótido, que comprende codones, que son traducidos en residuos de aminoácido (por ejemplo toda la región codificante, que comienza y termina con el codon de stop, es decir, el último codon antes del codon de stop). En otra modalidad, la molécula de ácido nucleico antisentido es “antisentido” a la “región no codificante” de la cadena codificante de una secuencia de nucleótido que codifica a PSE. El término “región no codificante” se refiere a secuencias de 5' y 3' que flanquean a las regiones codificantes y que no son traducidas en los aminoácidos (es decir, que también se denominan regiones de 5' y 3' no traducidas).

Bajo la condición de las secuencias codificantes de PSE aquí divulgadas de la cadena codificante (por ejemplo las secuencias representadas en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9) se pueden configurar los ácidos nucleicos antisentido de la invención, aplicando las reglas del apareamiento de Watson-Crick. La molécula de ácido nucleico antisentido puede ser complementaria a toda la región codificante del mRNA de PSE, pero más preferentemente, es un oligonucleótido, que es “antisentido” a solamente una parte de la región codificante o no codificante del mRNA de PSE. El oligonucleótido antisentido puede ser complementario, por ejemplo, a la región que rodea lugar de iniciación de la translación de mRNA de PSE. Un oligonucleótido antisentido puede tener una longitud de, por ejemplo, aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45 ó 50 y más nucleótidos. Ventajosamente, un oligonucleótido antisentido tiene una longitud de 15 a 25 nucleótidos. Un ácido nucleico antisentido según la invención puede ser construido, usando síntesis químicas y reacciones de ligación enzimáticas mediante procedimientos conocidos en el campo técnico. Un ácido nucleico antisentido (por ejemplo un oligonucleótido antisentido) puede ser sintetizado, por ejemplo químicamente, usando nucleótidos que existen en la naturaleza o nucleótidos modificados en diferente forma, que están configurados de tal forma, que incrementen la estabilidad biológica de las moléculas o que aumenten la estabilidad física del dúplex formado entre el ácido nucleico antisentido y el ácido nucleico sentido; por ejemplo, se pueden usar derivados de tioato de fósforo y nucleótidos sustituidos por acridina. Ejemplos de nucleótidos modificados que se usan para la generación del ácido nucleico antisentido son, entre otros: 5-fluoruracilo, 5-bromuracilo, 5-cloruracilo, 5-yoduracilo, hipoxantina, xantina, 5-(carboxihidroximetil)uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, metiladenina, metilcitosina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopentiladenina, ácido uracilo-5-oxiacético (v), wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metilo de ácido uracil-5-oxiacético, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w y 2,6-diaminopurina. Alternativamente, se puede preparar el ácido nucleico antisentido en forma biológica, usando un vector de expresión en el que se ha sub-clonado un ácido nucleico en dirección antisentido (es decir, ARN transcrito por un ácido nucleico incorporado, está orientado en un ácido nucleico destinatario de interés en dirección antisentido, lo que se describe en el próximo sub-capítulo en más detalle).

Las moléculas de ácido nucleico antisentido según la invención suelen ser administradas a una célula o son generadas *in situ*, de manera que hibriden con el mRNA celular y/o el ADN genómico, el cual codifica a una PSE, o que se ligan con la misma, para inhibir de este modo la expresión de la proteína, por ejemplo inhibiendo la transcripción y/o translación. La hibridación se puede realizar por complementariedad nucleotídica convencional, con formación de un dúplex estable o, por ejemplo, en el caso de moléculas de ácido nucleico antisentido, que liga a dúplex de ADN, por interacciones específicas en el gran surco de la doble hélice. La molécula antisentido puede estar modificada de tal manera, que se liga específicamente con un receptor o con un antígeno expresado sobre una superficie celular seleccionada, por ejemplo mediante enlace de la molécula de ácido nucleico antisentido con un péptido o un anticuerpo, que se liga con un receptor en la superficie celular o un antígeno. La molécula de ácido nucleico antisentido también se puede suministrar a las células usando los vectores aquí descritos. Para alcanzar una concentración intracelular suficiente de las moléculas antisentido se prefieren estructuras (constructos) de vectores, en los que se encuentra la molécula de ácido nucleico antisentido bajo el control de un promotor procarionta, viral o eucariota fuerte, inclusive un promotor vegetal.

En otra modalidad, la molécula de ácido nucleico antisentido según la invención es una molécula de ácido nucleico α -anómera. Una molécula de ácido nucleico α -anómera forma híbridos bicatenarios específicos con el ARN complementario, y las cadenas corren paralelamente entre sí, contrario a las unidades β usuales. [Gaultier *et al.* (1987)

Nucleic Acids Res. 15:6625-6641]. Además, la molécula de ácido nucleico antisentido puede comprender un 2'-o-metilribonucleótido [Inoue *et al.* (1987) Nucleic Acids Res, 15: 6131-6148] o un análogo de ARN-ADN quimérico [Inoue *et al.* (1987) FEBS Lett. 215:327-330J.

5 En otra modalidad, el ácido nucleico antisentido según la invención es un ribozima. Las ribozimas son moléculas de ARN catalíticas con actividad de ribonucleasa, que son capaces de escindir un ácido nucleico monocatenario, como por ejemplo un mRNA, con la que tienen una región complementaria. Por lo que las ribozimas, por ejemplo la ribozima de Hammerhead [descrita en Haselhoff y Gerlach (1988) Nature 334: 585 591] para la escisión catalítica de transcritos de mRNA de PSE, para así inhibir la translación de mRNA de PSE. Una ribozima con especificidad para un ácido nucleico codificante de PSE puede ser estructurada con base en la secuencia de nucleótido de uno de los cADN de PSE divulgados en la presente (es decir, 38ck21g07fwd en SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9) o con base en una secuencia heterológica a aislar según uno de los procedimientos enseñados en la presente. Por ejemplo, se puede construir un derivado de Tetrahymena-L-19-IVS-ARN, y la secuencia del nucleótido del lugar activo es complementaria a la secuencia de nucleótido que se desea escindir en un mRNA codificante de PSE. Véase, 10 por ejemplo, Cech *et al.*, US 4,987,071 y Cech *et al.*, US 5, 116, 742. Alternativamente, se puede usar mRNA de PSE para la selección de un ARN catalítico con una actividad de ribonucleasa específica a partir de un pool de moléculas de ARN [véase, por ejemplo, Bartel, D., y Szostak, J.W. (1993) Science 261:1411-1418].

Alternativamente, se puede inhibir la expresión genética de PSE, dirigiendo secuencias de nucleótido que son complementarias a regiones regulatorias de una secuencia de nucleótido de PSE (por ejemplo un promotor y/o enhancer de PSE) de tal manera, que se formen estructuras de triple hélice, que inhiben la transcripción de un gen de PSE en células destino [ver, en general, Helene, C. (1991) Anticancer Drug Res. 6(6) 569-84; Helene, C., *et al.* (1992) Ann. N.Y. Acad. Sci. 660:27-36; y Maher. L.J. Bioassays 14(12):807-815].

25 B. *Constructo génico*

Otra modalidad de la invención es un nuevo constructo de gen, que contiene un ácido nucleico aislado que proviene de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* y codifica a un polipéptido, que alonga a ácidos grasos de C₁₆, C₁₈ o C₂₀ con por lo menos dos dobles enlaces en el ácido graso, en por lo menos dos átomos de carbono, o la secuencia genética de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, sus homólogos, derivados o análogos, que están ligados de manera funcionalmente apta con una o varias señales de regulación, para aumentar ventajosamente la expresión del gen. Ejemplos de estas señales de regulación son secuencias, con las que se ligan inductores o represores, y así regulan la expresión del ácido nucleico. Adicionalmente a estas nuevas secuencias de regulación, todavía puede estar presente la regulación natural de estas secuencias antes de las estructuras de gen propiamente dichas y, si es adecuado, éstas pueden estar genéticamente modificadas, de manera que la regulación natural ha sido interrumpida y que la expresión de los genes ha sido incrementada. Sin embargo, el constructo de gen puede tener una estructura más sencilla. Es decir, no se han insertado señales de regulación adicionales antes de la secuencia SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 o sus homólogos, y se ha deletado el promotor natural con su regulación. En lugar de ello, se ha mutado la secuencia de regulación natural de tal forma que no tiene lugar ninguna regulación, y que la expresión del gen se incrementa. Además, el constructo de gen puede comprender ventajosamente una o varias secuencias llamadas secuencias enhancer, que están ligadas en forma funcionalmente apta con el promotor y hacen posible la expresión incrementada de la secuencia de ácido nucleico. Pero también es posible insertar en el extremo 3' de las secuencias de ADN adicionalmente secuencias ventajosas, por ejemplo, otros elementos de regulación o terminadores. Los genes de elongasa pueden estar presentes en el constructo de gen en una o varias copias. Para la inserción de otros genes en organismos es ventajoso que estén presentes otros genes en el constructo de gen.

Secuencias de regulación ventajosas para el nuevo procedimiento están presentes, por ejemplo en promotores, tales como promotor *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIq*, T7, T5, T3, *gal*, *trc*, *ara*, SP6, λ -PR o λ -PL y se usan ventajosamente en bacteria gram negativas. Otras secuencias de regulación ventajosas están presentes, por ejemplo en los promotores gram positivos: *amy* y SPO2, en los promotores de levadura o hongo ADC1, MF α , AC, P-60, CIC1, GAPDH, TEF, *rp28*, ADH o en los promotores vegetales CaMV/35S [Franck *et al.*, Cell 21 (1980) 285-294], PRP1 [Ward *et al.*, Plant. Mol. Biol. 22 (1993)], SSU, OCS, *lib4*, *usp*, STLS1, B33, *nos* o en el promotor de ubiquitina o faseolina. En este contexto también son ventajosos los promotores inducibles, como los promotores descritos en EP-A-0 388 186 (inducibles por bencilsulfonamid), Plant J., 1992:397-404 (Gatz *et al.*, inducibles por tetraciclina), EP-A-0 335 528 (inducibles por ácido abscisínico) o WO 93/21334 (inducibles por etanol o ciclohexenol). Otros promotores vegetales apropiados son el promotor de la FBPasa citosólica o el promotor de ST-LSI de la patata (Stockhaus *et al.*, EMBO J. 8, 1989, 2445), el promotor de fosforibosilpifosfatamidotransferasa a partir de Glycine 20 max (No. de acceso de banco genético; U87999) o el promotor específico de nodios descrito en la EP-A-0 249 676. Promotores especialmente ventajosos son aquellos, que hacen posible la expresión en tejidos que participan en la biosíntesis de ácido graso. Son especialmente ventajosos los promotores específicos de semillas, por ejemplo el promotor de *usp*, LEB4, faseolina o napina. Otros promotores especialmente ventajosos son los promotores específicos de semillas, que se pueden usar para plantas monocotiledóneas o dicotiledóneas, y que están descritas en la US 5,608,152 (promotor de napina a partir de colza), la WO 98/45461 (promotor de faseolina a partir de *Arobidopsis*), la US 5,504,200 (promotor de faseolina a partir de *Phaseolus vulgaris*), la WO 91/13980 (promotor de Bce4 a partir de Brassica), en Baumlein *et al.*, Plant J., 2, 2, 1992:233-239 (promotor de LEB4 a partir de una leguminosa), y estos promotores son adecuados para dicotiledóneas. Los siguientes promotores son apropiados, por ejemplo, para monocotiledóneas: promotor de *lpt-2* o *lpt-1* a partir de cebada (WO 95/15389 y WO 95/23230), promotor de hordeína a partir de cebada y otros promotores apropiados descritos en la WO 99/16890.

ES 2 331 228 T3

En principio, es posible usar todos los promotores naturales con sus secuencias de regulación, como las arriba mencionadas para el nuevo procedimiento. También es posible y ventajoso usar adicionalmente promotores sintéticos.

5 El constructo de gen también puede comprender otros genes, como se describe arriba, que deben ser insertados en los organismos. Es posible y ventajoso incorporar en los organismos huésped genes de regulación, tales como genes para inductores, represores o enzimas, que debido a su actividad enzimática intervienen en la regulación de uno o más genes de una vía de biosíntesis, y expresarlos en estos organismos. Estos genes pueden ser de origen heterológico u homólogo. Los genes insertados pueden tener su propio promotor o también estar bajo el control del promotor de la secuencia SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 y SEQ ID NO:11 o sus homólogos, derivados o análogos.

15 El constructo de gen comprende ventajosamente, para la expresión de los otros genes presentes, otras secuencias de regulación 3' y/o 5' terminales adicionales para incrementar la expresión, las cuales se seleccionan en función del organismo huésped elegido y del gen o de los genes para la expresión óptima.

Estas secuencias de regulación deben, tal como se ha indicado arriba, hacer posible la expresión específica de los genes y la expresión proteica. Esto puede significar, dependiendo del organismo huésped por ejemplo, que el gen es expresado o sobreexpresado sólo tras inducción o que es expresado y/o sobreexpresado inmediatamente.

20 Además, las secuencias o factores de regulación pueden tener ventajosamente un efecto ventajoso sobre la expresión de los genes insertados y, por tanto, aumentarla. De esta manera, se pueden intensificar los elementos de regulación usando fuertes señales de transcripción, tales como promotores y/o enhancers, ventajosamente, a nivel de transcripción. Sin embargo, también es posible intensificar la translación, por ejemplo, mejorando la estabilidad del mRNA.

25 Ventajosamente, las secuencias de ácido nucleico de la invención son clonadas conjuntamente con por lo menos un gen repórtter en un constructo génico (= cassette de expresión, estructura de ácido nucleico), que son insertados en el organismo por vía de un vector o directamente en el genoma.

30 Este gen repórtter debe ser fácil de detectar mediante un ensayo de crecimiento, fluorescencia, quimio- o bioluminiscencia o resistencia o mediante medidas fotométricas. Por ejemplo, son mencionados como genes repórtter los genes de resistencia contra antibióticos o herbicidas, genes de hidrolasa, genes de proteína de fluorescencia, genes bioluminiscentes, genes de metabolismo de azúcar o nucleótido o genes de biosíntesis, como por ejemplo del gen de Ura3, gen de Ilv2, el gen de luciferasa, gen de b-galactosidasa, el gen de gfp, el gen de 2-desoxiglucosa-6-fosfatofosfatasa, el gen de b-glucuronidasa, gen de b-lactamasa, el gen de neomicinofosfotransferasa, el gen de higromicinfosfotransferasa o el gen de BASTA (= resistencia a glufosinato). Estos genes permiten medir y cuantificar fácilmente la actividad de transcripción y con ello la expresión de los genes. De este modo se pueden identificar los lugares de genoma, que presentan una diferente productividad.

40 Las secuencias de ácido nucleico de la invención con SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9, que pueden codificar para elongasas, pueden contener en uno o varias copias en el casete de expresión (= constructo de gen).

45 En el casete de expresión (= estructura de gen, estructura de ácido nucleico) puede estar contenido por lo menos otro ácido nucleico, que codifica para un gen, preferentemente, de la biosíntesis de ácido graso, que deben ser insertados en los organismos huésped. Estos genes pueden estar bajo regulación separada o bajo la misma región de regulación que los genes de las elongasas de la invención. En estos genes se trata, por ejemplo, de otros genes de biosíntesis, ventajosamente de la biosíntesis de ácido graso, que permiten una síntesis incrementada. Por ejemplo, son mencionados los genes para la $\Delta 19$ -, $\Delta 17$ -, $\Delta 15$ -, $\Delta 12$ -, $\Delta 9$ -, $\Delta 8$ -, $\Delta 6$ -, $\Delta 5$ -, $\Delta 4$ -desaturasa, las diferentes hidroxilasas, la $\Delta 12$ -acetilasa, la acil-ACP-tioesterasas, β -cetoacil-ACP-sintasas o β -cetoacil-ACP-reductasas. Ventajosamente, se usan los genes de desaturasa en la estructura de ácido nucleico. También estos genes pueden estar contenidos en una o varias copias en el constructo de gen.

C. Vectores de expresión recombinantes y células huésped

55 Otro aspecto de la invención se refiere a vectores, preferentemente vectores de expresión, que contienen un ácido nucleico de la invención o un constructo génico según la invención, que codifican a una PSE (o una parte de la misma). Tal como se usa aquí, el término "vector" es una molécula de ácido nucleico, que es capaz de transportar a otro ácido nucleico, al que está ligada. Un tipo de vector es un "plásmido", que representa un bucle de ADN circular bicatenario, en el que pueden ligarse segmentos de ADN adicionales. Otro tipo de vector es un vector viral, y pueden ligarse segmentos de ADN adicionales en el genoma viral. Determinados vectores se pueden replicar autónomamente en una célula huésped, en la que se han insertado (por ejemplo, vectores de bacterias de origen de replicación bacteriana y vectores mamíferos episomales). Otros vectores (por ejemplo vectores mamíferos no episomales) son integrados en el genoma de la célula huésped en su inserción a la célula huésped y, por tanto, replicados conjuntamente con el genoma huésped. Además, determinados vectores pueden dirigir la expresión de genes, con los que están ligados en forma funcionalmente apta. Estos vectores se denominan en la presente: "vectores de expresión". Generalmente, los vectores de expresión, que son apropiados para técnicas de recombinación de ADN, tienen la forma de plásmidos. En la presente descripción, se pueden usar los términos "plásmido" y "vector" en forma intercambiable, ya que el plásmido es la forma de vector que más se usa. Sin embargo, la presente invención abarcará estas otras formas de vectores de expresión, tales como los vectores virales (por ejemplo retrovirus definidos por replicación, adenovirus y

virus similares a adenovirus), que ejercen funciones parecidos. Además, el término vector comprende también otros vectores conocidos a la persona técnica en la materia, tales como fagos, virus, tales como SV40, CMV, baculovirus, adenovirus, transposons, elementos de IS, fasmidos, fagémidos, cósmidos, o ADN circulares, así como ARN.

5 Los vectores de expresión recombinantes de la invención comprenden un ácido nucleico de la invención o un constructo génico según la invención en una forma, que es apropiada para la expresión del ácido nucleico en una célula huésped, lo que significa, que los vectores de expresión recombinantes comprenden una o varias secuencias de regulación, seleccionadas a base de las células huésped usadas para la expresión, que están ligadas en forma funcionalmente apta con la secuencia de ácido nucleico a expresar. En un vector de expresión recombinante, el término “ligado en forma funcionalmente apta” significa, que la secuencia de nucleótido de interés está ligada con la/las secuencia(s) de regulación de tal manera que sea posible la expresión de la secuencia de nucleótido y que estén unidas entre sí para que ambas secuencias cumplan la función prevista, asociada con la secuencia (por ejemplo, en un sistema de transcripción/translated *in vitro*, o en una célula huésped, cuando el vector está insertado en la célula huésped). El término “secuencia de regulación” comprende promotores, enhancers y otros elementos de control de la expresión (por ejemplo señales de poliadenilación). Estas secuencias de regulación están descritas, por ejemplo, en Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), o ver: Gruber y Crosby, en: Methods in Plant Molecular Biology and Biotechnology, CRe Press, Boca Raton, Florida, Eds.: Glick y Thompson, capítulo 7, 89-108, e inclusive la literatura allí citada.

20 Secuencias de regulación abarcan aquellas, que regulan la expresión constitutiva de una secuencia de nucleótido en muchos tipos de células huésped, y aquellas, que regulan la expresión directa de la secuencia de nucleótido únicamente en determinadas células huésped bajo determinadas condiciones. La persona técnica en la materia sabe que la configuración del vector de expresión puede depender de factores, tales como selección de la célula huésped a transformar, el grado de la expresión de la proteína deseada, etc. Los vectores de expresión de la invención pueden ser insertadas en células huésped y de este modo producir proteínas o péptidos, incluyendo proteínas de fusión o péptidos de fusión, los cuales se codifican por los ácidos nucleicos aquí descritos (por ejemplo, PSEs, formas mutantes de PSEs, proteínas de fusión, etc.).

Los vectores de expresión recombinantes de la invención pueden ser configuradas para la expresión de PSEs en células procariontas o eucariontas. Por ejemplo, se puede expresar genes de PSE en células bacterianas, por ejemplo *C. glutamicum*, células de insectos (usando vectores de expresión de baculovirus), células de levadura y otros hongos [véase Romanos, M.A., *et al.* (1992) “Foreign gene expression in yeast: a review”, *Yeast*8:423-488; vandenHondel, C.A.M.J.J., *et al.* (1991) “Heterologous gene expression in filamentous fungi”, en: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, Hrsgb., pág. 396-428: Academic Press: San Diego; y van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) “Gene transfer systems and vector development for filamentous fungi”, en: *Applied Molecular Genetics of Fungi*, Peberdy, J.F., *et al.*, Hrsgb., S. 1-28, Cambridge University Press: Cambridge], Algen [Falcitore *et al.*, 1999, *Marine Biotechnology*, 1, 3:239-251], ciliates de los tipos: Holotrichia, Peritrichia, Spirotrichia, Suctorina, Tetrahimena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Pseudocohnilembus, Euplotes, Engelmanniella y Stilonychia, en particular del género *Stilonychia lemnae*, con vectores según el proceso de transformación tal como se describe en WO 98/01572, así como células de vegetales multicelulares [véase Schmidt, R. y Willmitzer, L. (1988) “High efficiency *Agrobacterium tumefaciens*-mediated transformation of Arabidopsis thaliana leaf and cotyledon explants” *Plant Cell Rep.*:583-586; *Plant Molecular Biology and Biotechnology*, C Press, Boca Raton, Florida, Kapitel 6/7, S.71-119 (1993); F.F. White, B. Jenes *et al.*, *Techniques for Gene Transfer*, en: *Transgenic Plants*, Bd. 1, Engineering and Utilization, Hrsgb.: Kung y R. Wu, Academic Press (1993), 128-43; Potrykus, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991), 205-225 (y la literatura allí citada)] o células de mamífero. Células huésped apropiadas se discuten, además, en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990).

Alternativamente, se puede transcribir y translatar el vector de expresión recombinante *in vitro*, por ejemplo, usando secuencias de regulación de promotor T7 y T7-polimerasa.

La expresión de proteínas en procariontas tiene lugar, generalmente, con vectores, que contienen promotores constitutivos o inducibles, que regulan la expresión de proteínas de fusión o proteínas no fusionadas. Vectores de fusión agregan una serie de aminoácidos a una proteína allí codificada, generalmente en el aminotérmino de la proteína recombinante, pero también en el término C, o fusionan dentro de regiones apropiadas en las proteínas. Estos vectores de fusión suelen tener tres tareas: 1) intensificar la expresión de proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante y 3) apoyar la purificación de la proteína recombinante mediante acción como ligando en la purificación por afinidad, por ejemplo, por vía de los llamados His-Tags. En vectores de expresión de fusión, se inserta frecuentemente un lugar de rotura proteolítico en el lugar de unión de la unidad de fusión y la proteína recombinante, para que sea posible de este modo la separación de la proteína recombinante de la unidad de fusión después de la purificación de la proteína de fusión. Estas enzimas y sus correspondientes secuencias de reconocimiento abarcan el factor Xa, trombina y enteroquinasa.

Vectores de expresión de fusión típicos son, entre otros, pGEX [Pharmacia Biotech Inc; Smith, D. B., y Johnson, K. S. (1988) *Gene* 67:31-40], pMAL [New England Biolabs, Beverly, MA] y pRIT5 [Pharmacia, Piscataway, NJ], en los que se fusiona glutatona-S-transferasa (GST), proteína ligante de maltosa E o bien proteína A, con la proteína destino recombinante. En una modalidad, la secuencia codificante de PSE está clonada en un vector de expresión de pGEX, generando así un vector, que codifica a una proteína de fusión, que del término N al término C comprende proteína

ES 2 331 228 T3

X - lugar de rotura-trombina-GST. La proteína de fusión puede ser purificada por cromatografía de afinidad, usando resina de glutationa-agarosa. Puede ser generada PSE recombinante, que no está fusionada con GST, por escisión de la proteína de fusión con trombina.

5 Ejemplos de vectores de expresión no fusionados de *E. coli* inducibles, apropiados son, por ejemplo, pTrc (Amann *et al.* (1988) *Gene* 69:301-315) y pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). La expresión del gen destino del vector pTrc se basa en la transcripción por polimerasa de ARN huésped de un promotor de fusión de híbrido trp-lac. La expresión del gen destino a partir del vector pET 11d se basa en la transcripción de un promotor de fusión de T7-gn10-lac, que es inducida por polimerasa de ARN viral co-expresada (T7 gn1). Esta polimerasa viral es suministrada por las cepas huésped: BL21 (DE3) o HMS174 (DE3), de un λ -profago residente, que comprende un gen de T7 gnl bajo el control de transcripción del promotor lacUV 5.

15 Otros vectores apropiados en organismos procariotas son conocidos a la persona técnica en la materia. Estos vectores - se encuentran, por ejemplo, en *E. coli* pLG338, pACIC184, die pBR-Reihe, wie pBR322, die pUC-Reihe, wie pUC18 o pUC19, die M113mp-Reihe, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1, pgt11 o pBdCI, en *Streptomyces* pIJ101, pIJ364, pIJ702 o pIJ361, en *Bacillus* pUB110, pC194 o pBD214, en *Corynebacterium* pSA77 o pAJ667.

20 Una estrategia para la maximización de la expresión de proteína recombinante es la expresión de la proteína en una bacteria huésped, cuya capacidad de escisión procariota de la proteína recombinante es perturbada [Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128]. Otra estrategia es la modificación de la secuencia de ácido nucleico del ácido nucleico a insertar en el vector de expresión, de manera que los codones individuales para cada aminoácido son aquellos, que se usan preferentemente en una bacteria seleccionada para la expresión, como por ejemplo *C. glutamicum* [Wada *et al.* (1992) *Nucleic Acids Res.* 20: 2111-2118]. Esta modificación de las secuencias de ácido nucleico de la invención se realiza mediante técnicas estándar de síntesis de ADN.

30 En otra modalidad, el vector de expresión de PSE es un vector de expresión de levadura. Ejemplos de vectores para la expresión en la levadura, *S. cerevisiae*, comprenden: pYepSec1 [Baldari *et al.* (1987) *Embo J.* 6:229-234], pMFa [Kurjan y Herskowitz (1982) *Cell* 30:933-943], pJRY88 [Schultz *et al.* (1987) *Gene* 54:113-123] así como pYES2 [Invitrogen Corporation, San Diego, CA]. Vectores y procedimientos para la construcción de vectores apropiados para ser usados en otros hongos, tales como hongos filamentosos, comprenden aquellos, que se describen en detalle en: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi", en: Applied Molecular Genetics of fungi, J.F. Peberdy *et al.*, editores, pág. 1-28, Cambridge University Press: Cambridge, o en: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, editores, pág. 396-428: Academic Press: San Diego]. Otros vectores de levadura apropiados son, por ejemplo: pAG-1, YEp6, YEp13 o pEMBLYe23.

40 Alternativamente, se pueden expresar los PSEs de la invención en células de insectos, usando vectores de expresión de Baculovirus. Vectores de Baculovirus disponibles para la expresión de proteína en células de insectos cultivadas (por ejemplo células de Sf9), comprenden la serie pAc [Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165] y la serie pVL [Lucklow y Summers (1989) *Virology* 170:31-39].

45 Los vectores arriba mencionados sólo dan una pequeña idea de vectores adecuados posibles. Otros plásmidos son conocidos a la persona técnica en la materia y se describen, por ejemplo, en: Cloning Vectors (Eds. Pouwels, P.H., *et al.*, Elsevier, Amsterdam-New York-Oxford, 1985, ISBN O 444 904018).

50 En otra modalidad, se expresa un ácido nucleico de la invención en células de mamíferos, usando un vector de expresión mamífero. Ejemplos de vectores de expresión comprenden: pCDM8 [Seed, B. (1987) *Nature* 329:840] y pMT2PC [Kaufman, *et al.* (1987) *EMBO J.* 6:187-195]. Cuando se usan células de mamíferos, las funciones de control del vector de expresión frecuentemente son suministradas por elementos de regulación virales. Promotores generalmente usados provienen, por ejemplo, de Polyoma, Adenovirus2, Cytomegalievirus y Simian Virus 40. Para otros sistemas adecuados de expresión para células procariotas y eucariotas, véase en los capítulos 16 y 17 de Sambrook, J., Fritsch, E. F., y Maniatis, T., *Molecular Cloning: A Laboratory Manual*, 2. edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

60 En otra modalidad, el vector de expresión mamífero puede regular la expresión del ácido nucleico preferentemente en un determinado tipo de célula (por ejemplo, se usan elementos de regulación específicos de tejidos para la expresión del ácido nucleico). Elementos de regulación específicos de tejido son conocidos en el campo técnico.

Ejemplos no limitativos de promotores específicos de tejidos apropiados son, entre otros: el promotor de albúmina [específico del hígado; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277], promotores específicos limfoides [Calame y Eaton (1988) *Adv. Immunol.* 43:235-275], especialmente, promotores con receptores de células T [Winoto y Baltimore (1989) *EMBO J.* 8:729-733] e inmunoglobulinas [Banerji *et al.* (1983) *Cell* 33:729-740; Queen y Baltimore (1983) *Cell* 33:741-748], promotores específicos neuronales [por ejemplo promotor de neurofilamento; Byrne y Ruddle (1989) *PNAS* 86:5473-5477], promotores específicos de páncreas [Edlund *et al.*, (1985) *Science* 230: 912-916] y promotores específicos de glándulas mamarias [por ejemplo promotor de suero de leche; US 4,873,316 y EP-A-0 264 166]. También están comprendidos los promotores regulados por desarrollo, por ejemplo los promotores hox del ratón

ES 2 331 228 T3

[Kessel y Gruss (1990) *Science* 249:374-379] y el promotor de fetoproteína [Campes y Tilghman (1989) *Genes Dev.* 3:537-546].

En otra modalidad, las PSEs de la invención se pueden expresar en células de plantas unicelulares (por ejemplo algas), véase Falciatore *et al.*, 1999, *Marine Biotechnology* 1 (3):239-251 y la literatura allí citada, y células vegetales de plantas superiores (por ejemplo espermatofitos, tales como frutas de campo). Ejemplos de vectores de expresión vegetales comprenden aquellos que están descritos detalladamente en: Becker, D., Kemper, E., Schell, J., y Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20:1195-1197; y Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acids Res.* 12:8711-8721; *Vectors for Gene Transfer in Higher Plants*; en: *Transgenic Plants*, Bd. 1, Engineering and Utilization, editors: Kung y R. Wu, Academic Press, 1993, pág. 15-38. Otros vectores vegetales apropiados se describen, por ejemplo, en: "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), cap. 6/7, págs. 71 119.

Vectores ventajosos son los llamados vectores shuttle o vectores binarios, que se replican en *E. coli* y *Agrobacterium*.

Un casete de expresión vegetal contiene preferentemente secuencias de regulación que pueden regular la expresión genética en células vegetales y están ligados en forma funcionalmente apta, de modo que cada secuencia puede cumplir su función como terminación de la transcripción, por ejemplo, señales de poliadenilación. Señales de poliadenilación preferidas son aquellas, que provienen de *Agrobacterium tumefaciens*-t-ADN, como el gen 3 del Ti-Plásmido, pTiACH5, conocido como octopinsintasa [Gielen *et al.*, *EMBO J.* 3 (1984) 835 siguientes] o equivalentes funcionales del mismo, pero también todos los otros terminadores funcionalmente activos en plantas.

Ya que la expresión de genes vegetales frecuentemente no está limitada a niveles de transcripción, el casete de expresión vegetal contiene preferentemente otras secuencias ligadas en forma funcionalmente apta, tales como enhancers de translación, por ejemplo, la secuencia overdrive, que contiene la secuencia líder no translada en 5' a partir del virus de mosaico del tabaco, que aumenta la relación de proteína/ARN [Gallie *et al.*, 1987, *Nucl. Acids Research* 15:8693-8711].

La expresión de genes vegetales debe ser ligada en forma funcionalmente apta con un promotor apropiado, que realiza la expresión genética oportunamente de manera específica para la célula o el tejido. Promotores preferidos, que provocan la expresión constitutiva [Benfey *et al.*, *EMBO J.* 8 (1989) 2195-2202], como aquellos que proceden de virus vegetales, tal como 35S CAMV [Franck *et al.*, *Cell* 21 (1980) 285-294], 19S CaMV (ver también US 5,352,605 y WO 84/02913) o promotores vegetales, tal como la subunidad de rubisco descrita en la US 4,962,028.

Otras secuencias preferidas para ser usadas en el enlace funcionalmente apto en casetes de expresión de genes vegetales son las secuencias de targeting, que son necesarias para dirigir el producto genético en su compartimento celular correspondiente [véase un cuadro sinóptico en Kermode, *Crit. Rev. Plant Sci* 15, 4 (1996) 285-423 y la literatura allí citada], por ejemplo, en la vacuola, el núcleo celular, todos los tipos de plastidios, tales como amiloplastos, cloroplastos, cromoplastos, el espacio extracelular, las mitocondrias, el retículo endoplasmático, cuerpos oleosos, peroxisomas y otros compartimentos en células vegetales.

La expresión de genes vegetales se puede facilitar mediante un promotor químicamente inducible [ver un cuadro sinóptico en Gatz 1997, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 48:89-108]. Promotores químicamente inducibles son especialmente apropiados, cuando se desea realizar la expresión genética de manera temporalmente específica. Ejemplos de promotores de este tipo son un promotor inducible por ácido salicílico (WO 95/19443), un promotor inducible por tetraciclina [Gatz *et al.* (1992) *Plant J.* 2, 397-404] y un promotor inducible por etanol.

También promotores que reaccionan a condiciones de estrés biótico o abiótico son promotores apropiados, por ejemplo el promotor del gen PRP1 inducido por patogenina [Ward *et al.*, *Plant. Mol. Biol.* 22 (1993) 361-366], el promotor hsp80 del tomate inducido por calor (US 5,187,267), el promotor de alfa-amilasa de la papa inducido por frío (WO 96/12814) o el promotor pinII inducido por heridas (EP-A-0 375 091).

Se prefieren, especialmente, aquellos promotores, que inducen la expresión genética en tejidos y órganos, en los que tiene lugar la síntesis de lípidos y aceite, en células de semillas, tales como células del endoespermo y del embrión que se va desarrollando. Promotores apropiados son el promotor de gen de napina a partir de colza (US 5,608,152), de promotor de USP a partir de Vicia faba [Baeumlein *et al.*, *Mol Gen Genet*, 1991, 225 (3): 459-67], el promotor de oleosina a partir de *Arabidopsis* (WO 98/45461), el promotor de faseolina a partir de *Phaseolus vulgaris* (US 5,504 r 200) (el promotor de Bce4 a partir de Brassica (WO 91/13980) o el promotor B4 de legumina [LeB4; Baeumlein *et al.*, 1992, *Plant Journal*, 2 (2): 233-9] así como promotores, que inducen expresión específica de semillas en plantas monocotiledóneas, tales como maíz, cebada, trigo, centeno, arroz, etc.

Promotores apropiados, dignos de mencionar son el promotor del gen lpt2 o lpt1 a partir de la cebada (WO 95/15389 y WO 95/23230) o los promotores descritos en la WO 99/16890 (promotores a partir del gen de hordeína de la cebada, del gen de glutéina de arroz, del gen de orizina de arroz, del gen de prolamina del arroz, del gen de gliadina del trigo, del gen de glutelina del trigo, del gen de zeína del maíz, del gen de glutelina de la avena, del gen de casirina de sorgo, del gen de secalina del centeno).

También son especialmente apropiados los promotores, que inducen la expresión específica en plástidos, ya que los plástidos son el compartimiento, en el que se sintetizan algunos precursores así como algunos productos finales de la biosíntesis de lípidos. Promotores apropiados, tales como el promotor de polimerasa de ARN viral, están descritos en la WO 95/16783 y WO 97/06250, y el promotor de clpP a partir de arábidopsis, está descrito en la WO 99/46394.

5 La invención provee además un vector de expresión recombinante que comprende una molécula de ADN de la invención que está clonada en dirección antisentido en el vector de expresión. Es decir, la molécula de ADN está unida en forma funcionalmente apta con una secuencia reguladora de tal manera, que sea posible la expresión (por transcripción de la molécula de ADN) de una molécula de ARN que está en relación “antisentido” frente al mRNA de PSE. Se pueden seleccionar secuencias de regulación, que están ligadas en forma funcionalmente apta con un ácido nucleico clonado en dirección antisentido, y que regulan la expresión continua de la molécula de ARN antisentido en un gran número de tipos de células, por ejemplo se pueden seleccionar promotores y/o enhancers virales o secuencias de regulación, que regulan la expresión constitutiva específica para tejido o para tipo celular de ARN antisentido. El vector de expresión antisentido puede estar presente en forma de un plásmido recombinante, un fagémido o el virus atenuado, en el que se producen ácidos nucleicos antisentido bajo el control de una región reguladora altamente activa, cuya actividad puede determinarse por el tipo de la célula, en la que se ha insertado el vector. Una explicación de la regulación de la expresión genética mediante genes antisentido se encuentra en Weintraub, H., *et al.*, Antisentido-ARN as a molecular tool for genetic analysis, Reviews Trends in Genetics, vol. 1(1) 1986.

20 Otro aspecto de la invención se refiere a células huésped, en las que se ha insertado un vector de expresión recombinante de la invención. Los términos “célula huésped” y “célula huésped recombinante” se utilizan en la presente en forma intercambiable. Claro está, que estos términos no se refieren únicamente a determinadas células destino, sino que también a descendientes o descendientes potenciales de esta célula. Ya que en generaciones sucesivas pueden presentarse determinadas modificaciones debido a mutaciones o influencias ambientales, estos descendientes no son necesariamente idénticos con la célula padre, pero siempre están comprendidos en el ámbito del término, tal como aquí se usa.

30 Una célula huésped puede ser una célula procariota o eucariota. Por ejemplo, se puede expresar una PSE en células de bacterias, tal como *C. glutamicum*, células de insectos, células de hongos o células de mamíferos (como células del ovario del hámster chino (CHO) o células de COS), algas, ciliatos, células vegetales, hongos u otros microorganismos, como *C. glutamicum*. Otras células huésped son conocidas a la persona técnica en la materia.

35 El vector de ADN se puede insertar en células procariotas o eucariotas mediante las técnicas de transformación o transfección convencionales. Los términos “transformación” y “transfección”, conjugación o transducción en el sentido usado aquí comprenden un gran número de procedimientos conocidos en el estado de la técnica para la inserción de ácido nucleico ajeno (por ejemplo ADN) en una célula huésped, inclusive la co-precipitación de fosfato de calcio o cloruro de calcio, la transfección inducida por DEE-dextrano, transferencia por la lipofección, competencia natural inducida químicamente, electroporación o bombardeo con partículas.

40 Procedimientos apropiados para la transformación o transfección de células huésped, inclusive células vegetales, se encuentran en: Sambrook *et al.* (Molecular Cloning: A Laboratory Manual., 2a edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) y otros manuales de laboratorio, como por ejemplo: Methods in Molecular Biology, 1995, vol. 44, Agrobacterium protocols, Eds: Gartland y Davey, Humana Press, Totowa, New Jersey.

45 De transfección estable de células de mamífero se conoce, que dependiendo del vector de expresión usado y de la técnica de transfección aplicada, sólo una parte de la células integran el ADN ajeno en su genoma. Para la identificación y selección de estos integrantes se suele insertar un gen, que codifica a un marcador selectivo (por ejemplo resistencia contra antibióticos), junto con el gen de interés en las células huésped. Marcadores selectivos preferidos abarcan aquellos que proporcionan resistencia contra medicamentos, tales como G418, higromicina y metotrexato, o aquellos que proporcionan resistencia en plantas contra un herbicida, como glifosato o glufosinato. Otros marcadores apropiados son, por ejemplo, marcadores que codifican a genes, que participan en las vías de biosíntesis de, por ejemplo, azúcares o aminoácidos tales como β -galactosidasa, ura3 o ilv2. Marcadores, que codifican a genes, tal como luciferasa, gfp u otros genes de fluorescencia, también son apropiados. Estos marcadores se pueden usar en forma de mutantes, en los que estos genes no son funcionales ya que han sido deletados, por ejemplo, mediante procesos convencionales.

60 Además, se puede insertar marcadores, que codifican a un ácido nucleico codificante de un marcador, en una célula huésped sobre el mismo vector, que aquél, que codifica a una PSE, o se pueden insertar sobre un vector separado. Células, que han sido transfectadas en forma estable con el ácido nucleico insertado, pueden ser identificadas, por ejemplo, mediante selección de medicamentos (por ejemplo, las células que contienen integrado el marcador seleccionado sobreviven, mientras que las otras células mueren).

65 Para la generación de un microorganismo recombinante homólogo se prepara un vector, que contiene por lo menos una parte del gen de PSE, en la que se ha insertado una delección, adición e sustitución, para así modificar al gen de PSE, por ejemplo interrumpirlo funcionalmente. Este gen de PSE es, preferentemente, un gen de PSE de *Physcomitrella*, *Phytophthora* o *Thraustochytrium*, pero también puede ser un homólogo o análogo a partir de otros organismos, aún de una fuente de mamíferos, hongos o insectos.

ES 2 331 228 T3

En una modalidad preferida, el vector está conformado de tal manera que el gen de PSE endógeno sea interrumpido en la recombinación homóloga (es decir, ya no codifica a una proteína funcional, también denominado vector knock-out). Alternativamente, el vector también puede estar conformador de tal manera, que el gen de PSE endógeno sea mutado en la recombinación homóloga o modificado de otra forma, pero que siga codificando para una proteína funcional (por ejemplo, la región reguladora situada corriente arriba puede estar modificada de tal forma, que por esto la expresión de la PSE endógena sea modificada). Para generar una mutación puntual por recombinación homóloga se pueden usar también híbridos de ADN-ARN conocidos como Chimeraplasty, que se conocen de ColeStrauss *et al.*, 1999, *Nucleic Acids Reseach* 27(5):1323-1330 y Kmiec, *Gene therapy*, 1999, *American Scientist*, 87(3):240-247.

En el vector para la recombinación homóloga, la sección modificada del gen de PSE está flanqueada en su extremo 5' y 3' por ácido nucleico adicional del gen de PSE, lo que permite una recombinación homóloga entre el gen de PSE exógeno, que está presente sobre el vector, y un gen de PSE endógeno en un microorganismo o una planta. El ácido nucleico de PSE flanqueante adicional es suficientemente largo para una recombinación homóloga exitosa con el gen endógeno. Normalmente, en el vector están contenidos varios cientos de pares de bases hasta kilobases de ADN flanqueante (tanto en el extremo 5', como también en el extremo 3') [para una descripción de vectores para la recombinación homóloga, ver, por ejemplo: Thomas, K.R., y Capecchi, M.R. (1987) *Cell* 51:503 o la recombinación en *Physcomitrella patens* a base de cADN ver: Strepp *et al.*, 1998, *Proc. Natl. Acad. Sci. USA* 95 (8):4368-4373]. El vector es insertado en un microorganismo o una célula vegetal (por ejemplo mediante ADN inducido por polietilenglicol), y aplicando las técnicas conocidas en el campo técnico, se seleccionan las células en las que el gen de PSE insertado se ha recombinado homológamente con el gen de PSE endógeno.

En otra modalidad, se pueden preparar organismos recombinantes, tales como microorganismos de plantas, que contienen los sistemas seleccionados, y que permiten la expresión regulada del gen insertado. La inserción de un gen de PSE en un vector, poniéndolo bajo el control del operón lac, permite la expresión del gen de PSE, por ejemplo, únicamente en presencia de IPTG. Estos sistemas de regulación son conocidos en el campo técnico.

Una célula huésped de la invención, como una célula huésped procariota o eucariota, en un cultivo o que crece en el campo, se puede usar para la producción (a saber, la expresión) de una PSE. En plantas se puede aplicar adicionalmente un procedimiento alternativo, que consiste en la transferencia directa de ADN en flores en desarrollo por medio de electroporación o por transferencia genética mediante *Agrobacterium*. Por tanto, la invención proporciona un procedimiento para la producción de PSEs, usando las células huésped de la invención. En una modalidad, el procedimiento comprende el cultivo de la célula huésped de la invención (en la que se ha insertado un vector de expresión recombinante que codifica a una PSE, o en cuyo genoma se ha insertado un gen, que codifica un tipo silvestre de PSE o una PSE modifica) en un medio apropiado, hasta que se ha producido la PSE. El procedimiento abarca en otra modalidad el aislamiento de las PSEs a partir del medio o de la célula huésped.

Células huésped, que en principio son apropiadas para recibir el ácido nucleico de la invención, el nuevo producto genético de la invención o el vector de la invención, son todos organismos procariotas o eucariotas. Los organismos huésped usados ventajosamente son organismos, tales como bacterias, hongos, levaduras, células animales o vegetales. Otros organismos ventajosos son animales o, preferentemente, plantas o partes de las mismas. Por "animal" se entiende aquí, que no están comprendidos los humanos. Hongos, levaduras o plantas se usan de preferencia, especialmente los hongos o plantas, muy preferentemente las plantas, tales como plantas oleáceas, que contienen grandes cantidades de lípidos, tales como colza, onagra, ricino, canola, cacahuete, lino, soya, cártamo, girasol, borraja, palma de aceite, coco o plantas, tales como maíz, trigo, centeno, avena, triticale, arroz, cebada, algodón, manioca, pimiento, tagetes, plantas solanáceas, tales como papas, tabaco, berenjena y tomates, especies de vicia, arvejas, alfalfa, arbustos (café, cacao, té), especies de salix, árboles (palma aceitera, coco), así como gramíneas perennes y frutas de campo. Vegetales de la invención especialmente preferidos son plantas oleáceas, tales como soya, cacahuete, colza, canola, ricino, lino, onagra, girasol, cártamo, árboles (palma aceitera, coco).

En un aspecto especialmente preferido, se refiere la invención a una célula vegetal, que comprende el polinucleótido de la invención o el vector de la invención.

Además, son preferidas las plantas transgénicas o tejidos de plantas transgénicas, que comprenden la célula vegetal según la invención.

En otro aspecto se refiere la presente invención a aquellas partes de las plantas de la invención, que se pueden cosechar, y aquél material, que es apropiado para la reproducción de las plantas transgénicas según invención y que contiene las células vegetales de la invención que expresan el ácido nucleico de la invención, o bien que contiene célula, que contienen un elevado nivel de la proteína de la invención. En principio, todas las partes de una planta pueden ser cosechadas, especialmente las flores, polen, frutas, plántulas, raíces, hojas, semillas, bulbos, tallos, etc. Material de reproducción comprende, por ejemplo, semillas, frutas, plántulas, bulbos, cortes y rizomas.

D. PSE aislada

Otro aspecto de la invención se refiere a PSEs aisladas y partes biológicamente activas de las mismas. Una proteína "aislada" o "purificada", o una parte biológicamente activa de la misma, está sustancialmente desprovista de material celular, cuando se ha producido mediante técnicas de recombinación de ADN, o de productos químicos precursores u otros químicos, cuando se ha sintetizado químicamente. El término "sustancialmente desprovisto de material celular"

comprende preparaciones de PSE, en las que la proteína ha sido separada de componentes celulares de las células, en las que es producida naturalmente o por recombinación. En una modalidad, el término “sustancialmente desprovisto de material celular” se refiere a preparaciones de PSE con menos de aproximadamente un 30% (con respecto al peso de la materia seca) de proteína que no sea PSE (que en la presente también se denomina “proteína no purificada”), más preferentemente, menos de aproximadamente un 20% de proteína no PSE, aún más preferentemente, menos de aproximadamente un 10% de no PSE y, especialmente, menos de aproximadamente un 5% de no PSE. Cuando se ha producido una PSE o una parte biológicamente activa de la misma por recombinación, entonces también es sustancialmente desprovista del medio de cultivo, es decir, el medio de cultivo constituye menos de aproximadamente un 20%, más preferentemente, menos de aproximadamente un 10% y, especialmente, menos de aproximadamente un 5% del volumen de la preparación de proteína. La expresión “sustancialmente desprovisto de productos químicos precursores u otros químicos” se refiere a preparaciones de PSE, en las que la proteína se ha separado de los químicos precursores o de otros químicos, que participan en la síntesis de la proteína. En una modalidad, el término “sustancialmente desprovisto de productos químicos precursores u otros químicos” comprende preparaciones de PSE con menos de aproximadamente un 30% (con respecto al peso de la materia seca) de químicos precursores o químicos no PSE, más preferentemente, menos de aproximadamente un 20% de productos químicos precursores o químicos no PSE, aún más preferentemente, menos de aproximadamente un 10% de productos químicos precursores o químicos no PSE, y especialmente, menos de aproximadamente un 5% de productos químicos precursores o químicos no PSE.

En modalidades preferidas, las proteínas aisladas o partes biológicamente activas de las mismas no presentan proteínas contaminantes del mismo organismo, del que proviene la PSE. En el caso de proteína de la invención, que es codificada por la secuencia representada en SEQ ID NO: 10, o codificada por un gen que comprende la SEQ ID NO: 9, hay que tener en cuenta, que la secuencia comienza con dos Met. Esto puede resultar en que en la translación de una secuencia de ácido nucleico codificante correspondiente, se expresen dos derivados de la proteína de la invención, comenzando en el primer o en el segundo Met. La relación de expresión entre los dos derivados puede oscilar, según el tipo de expresión o el organismo huésped, entre 0 y 1.

De esta manera está comprendida PSE que contiene ambos derivados nombrados o solo uno de los derivados. Los dos derivados pueden tener diferentes actividades, localizaciones, vidas medias, mecanismos de regulación, etc. Estas proteínas suelen ser obtenidas por expresión recombinante, por ejemplo, de PSE de *Physcomitrella*, *Phytophthora*, *Cryptocodium* o *Thraustochytrium* en plantas, como *Physcomitrella patens* o bien las arriba mencionadas, o microorganismos, por ejemplo, bacterias, tales como *E. coli*, *Bacillus subtilis*, *C. glutamicum*, hongos, tal como *Mortierella*, levaduras, como *Saccharomyces*, o ciliatos, tales como *Colpidium*, o algas como *Phaeodactylum*.

Una PSE aislada de la invención o una parte de la misma también puede participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en *Physcomitrella*, *Phytophthora* o *Thraustochytrium*, o en el transporte de moléculas por estas membranas. En modalidades preferidas, la proteína o parte de la misma comprende una secuencia de aminoácido que es suficientemente homóloga con una secuencia de aminoácido de la SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 para que la proteína o parte de la misma mantenga la capacidad de participar en el metabolismo de compuestos necesarios para la composición de membranas celulares en *Physcomitrella*, *Phytophthora* o *Thraustochytrium*, o en el transporte de moléculas por estas membranas.

La parte de la proteína es, preferentemente, una parte biológicamente activa, tal y como se describe en la presente. En otra modalidad preferida, una PSE de la invención tiene una de las secuencias de aminoácido mostradas en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10.

En otra modalidad preferida, la PSE tiene una secuencia de aminoácido, que es codificada por una secuencia de nucleótido que, por ejemplo, en condiciones severas hibrida con una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9.

En aún otra modalidad preferida, la PSE tiene una secuencia de aminoácido, que es codificada por una secuencia de nucleótido, que tiene una homología de por lo menos aproximadamente un 50 a 60%, preferentemente por lo menos aproximadamente un 60 a 70%, más preferentemente por lo menos aproximadamente un 70 a 80%, 80 a 90%, 90 a 95% y aún más preferentemente por lo menos aproximadamente un 96%, 97%, 98%, 99% o más, con una de las secuencias de aminoácido de SEQ ID NO:2, SEQ ID NO: 4, y SEQ ID NO:10.

La PSE de la invención preferida también posee, preferentemente, por lo menos una de las actividades de PSE aquí descritas. Una PSE de la invención preferida comprende una secuencia de aminoácido, que es codificada por una secuencia de nucleótido que hibrida, por ejemplo, en condiciones severas, con una secuencia de nucleótido de SEQ ID NO:1, SEQ ID NO:3, y SEQ ID NO:9 y puede participar en el metabolismo de compuestos necesarios para composición de membranas celulares en *Physcomitrella*, *Phytophthora*, *Cryptocodium* o *Thraustochytrium*, o en el transporte de moléculas por estas membranas, o que puede alongar a uno o varios ácidos grasos poliinsaturados con por lo menos dos dobles enlaces y una longitud de cadena de C₁₆ o C₁₈.

En otras modalidades I la PSE es sustancialmente homóloga con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 y mantiene la actividad funcional de la proteína de una de las secuencias de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, pero su secuencia de aminoácido se diferencia debido a la variación o mutagénesis naturales, como se ha descrito en detalle en el subcapítulo 1 de arriba. Por tanto, en otra modalidad, la PSE es una proteína, que comprende una secuencia de aminoácido que presenta una homología de por lo menos

ES 2 331 228 T3

aproximadamente un 50 a 60%, preferentemente por lo menos aproximadamente un 60 a 70% y más preferentemente por lo menos aproximadamente un 70 a 80%, 80 a 90%, 90 a 95% y especialmente, por lo menos aproximadamente un 96%, 97%, 98%, 99% o más con una secuencia de aminoácido completa de la SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10, y que presenta por lo menos una de las actividades de PSE aquí descritas. En otra modalidad, la invención se refiere a una proteína completa de *Physcomitrella*, *Phytophthora* o *Thraustochytrium*, que es substancialmente homóloga con una secuencia de aminoácido de SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10.

Las partes biológicamente activas de una PSE abarcan péptidos que comprenden secuencias de aminoácido que se derivan de la secuencia de aminoácido de una PSE, por ejemplo una secuencia de aminoácido o la secuencia de aminoácido de una proteína representada en SEQ ID NO:2, SEQ ID NO:4, y SEQ ID NO:10 que es homóloga con una PSE, que presenta menos aminoácidos que la PSE de longitud completa o la proteína de longitud completa, que es homóloga con una PSE, y que presentan por lo menos una actividad de una PSE. Normalmente, las partes biológicamente activas (péptidos, por ejemplo péptidos que tienen, por ejemplo, una longitud de 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos) comprenden un dominio o un motivo con por lo menos una actividad de una PSE. Además, mediante técnicas recombinantes se pueden preparar otras partes biológicamente activas en las que otras regiones de la proteína están deletadas y se pueden examinarlas con respecto a una o varias de las actividades aquí descritas. Las partes biológicamente activas de una PSE comprenden preferentemente uno o varios dominios/motivos seleccionados o partes de los mismos con actividad biológica.

Tales dominios y motivos se pueden identificar, por ejemplo, por análisis secuencial, por ejemplo con la ayuda de métodos apoyados en procesamiento electrónico de datos.

En las secuencias de la invención se encontraron, por ejemplo, unos llamados métodos de KK.

Kermode 1996, *Critical Reviews in Plant Sciences* 15 (4):285-423 describe motivos KK, una lisina doble que se encuentra predominantemente como motivo KKXX o K X K XXX y que influyen en un reciclado de ER en el aparato de Golgi y con ello el tiempo de residencia de la proteína y su actividad enzimática en un lugar determinado, especialmente, en ER.

Motivos de lisina doble se encuentran también, por ejemplo, en $\Delta 12$ -desaturasas (Arondel *et al.* 1992, *Science* 258:1353) y también están presentes en las elongasas de la invención. Tales motivos se describen, especialmente, como localizables en los terminales C. En las secuencias de la invención y aquí descritas se encuentra una acumulación ostensible de lisinas en el término C.

Elongasa Moho PSE1: término C KQKGAKTE

SEQ ID NO 2: KTKKA

SEQ ID NO 4: KKSTPAAKKTN

SEQ ID NO 6: KHLK

De esta manera, se puede tratar aquí de una posible variación genética.

Hay restos de Lys, que afectan el targeting, el direccionamiento y la localización en, o adentro de, ER. Un enmascaramiento de esta secuencia, una modificación o una modificación espacial, la cercanía al final del término C, por ejemplo por fusión con GFP "green fluorescent protein" se puede usar para influir en la compartimentación.

Las PSEs se preparan, preferentemente, mediante técnicas de recombinación de ADN. Por ejemplo, se clona una molécula de ácido nucleico codificante de una proteína en un vector de expresión (tal como se describe arriba), el vector de expresión es insertado en una célula huésped (tal como se describe arriba), y la PSE se expresa en la célula huésped. La PSE se puede aislar luego de las células mediante un esquema de purificación apropiado, aplicando técnicas de purificación de proteína estándar. Alterno a la expresión recombinante se puede sintetizar una PSE, un polipéptido de PSE, o un péptido de PSE químicamente mediante técnicas estándar de síntesis de péptidos. Además, se puede aislar PSE nativa de células (por ejemplo células de endotelio) por ejemplo usando un anticuerpo anti-PSE, que se puede producir mediante técnicas estándar, usando una PSE de la invención o un fragmento de la misma.

La invención también provee proteínas de PSE quiméricas o proteínas de fusión de PSE. Tal como aquí se usa, una "proteína de PSE quimérica" o "proteína de fusión de PSE" comprende un polipéptido de PSE, que está ligado en forma funcionalmente apta con un polipéptido no PSE. Un "polipéptido de PSE" se refiere a un polipéptido con una secuencia de aminoácido que corresponde a una PSE, mientras que un "polipéptido no PSE" es un polipéptido con una secuencia de aminoácido que corresponde a una proteína, que es substancialmente no homóloga con la PSE, por ejemplo, una proteína que se diferencia de PSE y que proviene del mismo o de otro organismo.

Dentro de la proteína de fusión, el término "ligado en forma funcionalmente apta" significará que el polipéptido de PSE y el polipéptido no PSE estén fusionados entre sí de tal manera, que ambas secuencias cumplen la función asignada a la secuencia usada. El polipéptido no PSE puede estar fusionado en el término N o en el término C del polipéptido de PSE. En una modalidad, la proteína de fusión es, por ejemplo, una proteína de fusión de GST-PSE, en

ES 2 331 228 T3

donde las secuencias de PSE están fusionadas en el término C de las secuencias GST. Estas proteínas de fusión pueden facilitar la purificación de las PSEs recombinantes.

5 En otra modalidad, la proteína de fusión es una PSE, que presenta una secuencia de señal heteróloga en su término N. En determinadas células huésped, (por ejemplo células huésped de mamíferos), la expresión y/o secreción de una PSE puede incrementarse usando una secuencia de señal heteróloga.

10 Una proteína quimérica de PSE de la invención o una proteína de fusión de PSE de la invención se prepara por técnicas estándar de recombinación de ADN. Por ejemplo, se ligan fragmentos de ADN, que codifican a diferentes secuencias de polipéptidos, según técnicas convencionales en la trama de lectura entre sí, usando, por ejemplo, extremos lisos o solapantes para la ligación, separación de enzima de restricción para suministrar extremos apropiados, el relleno de extremos cohesivos, de ser necesario, tratamiento con fosfatasa alcalina, para evitar enlaces no deseados, y ligación enzimática.

15 En otra modalidad, se puede sintetizar el gen de fusión mediante técnicas convencionales, incluyendo sintetizadores automáticos de ADN. Alternativamente, se puede realizar una amplificación de PCR de fragmentos de genes usando primers de anclaje, que generan solapas complementarias entre fragmentos de gen sucesivos, que luego pueden ser hibridados entre sí y re-amplificados, generando de este modo una secuencia de gen quimérica (ver, por ejemplo: *Current Protocols in Molecular Biology*, Eds. Ausubel *et al.*, John Wiley & Sons: 1992).

20 Además, se obtienen muchos vectores de expresión en el comercio, que ya codifican a una unidad de fusión (por ejemplo un polipéptido de GST). Un ácido nucleico codificante de PSE puede clonarse en un vector de expresión así, de manera que la unidad de fusión esté ligada en la trama de lectura con la proteína de PSE.

25 Homólogos de la PSE se pueden generar por mutagénesis, por ejemplo por mutación específica puntual o recorte de la PSE. El término "homólogos", en el sentido usado aquí, se refiere a una forma variante de la PSE, que actúa de agonista o antagonista de la actividad de PSE. Un agonista de la PSE puede mantener substancialmente la misma actividad que la PSE o una parte de la actividad biológica de la PSE. Un antagonista de la PSE puede inhibir una o varias actividades de la forma nativa de la PSE, por ejemplo, por enlace competitivo con un elemento situado corriente abajo o corriente arriba de la cascada de metabolismo para componentes de membrana celular, los cuales comprende la PSE, o por enlace con una PSE, que inicia el transporte de enlaces por membranas celulares, inhibiendo de este modo la translocación.

30 En una modalidad alternativa, se pueden identificar homólogos de la PSE, revisando bancos combinatorios de mutantes, por ejemplo mutantes de acortamiento con respecto a agonistas o antagonistas de la actividad de PSE. En una modalidad, se genera un banco variegado de variantes de PSE por mutagénesis combinatoria a nivel de ácido nucleico y se codifica mediante un banco genético variegado. Un banco variegado de variantes de PSE se puede producir, por ejemplo, por ligación enzimática de una mezcla de oligonucleótidos sintéticos en secuencias genéticas, de manera, que se pueda expresar un set degenerado de potenciales secuencias de PSE como polipéptidos individuales o, alternativamente, como set de proteínas de fusión más grandes (por ejemplo para la visualización de fago), que contienen a este set de secuencias de PSE.

35 Hay un gran número de procedimientos que se pueden aplicar para la obtención de bancos de potenciales homólogos de PSE a partir de una secuencia de oligonucleótido degenerada. La síntesis química de una secuencia genética degenerada se puede realizar en una autómatas de síntesis de ADN y el gen sintético se puede ligar entonces con un vector de expresión apropiado. El uso de un set degenerado de genes permite la provisión de todas las secuencias que codifican al set deseado de potenciales secuencias de PSE, en una mezcla. Procedimientos para la síntesis de oligonucleótidos degenerados son conocidos en el campo técnico [ver, por ejemplo, Narang, S.A. (1983) *Tetrahedron* 39: 3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.*, (1984) *Science* 198: 1056; Ike *et al.* (1983) *Nucleic Acids Res.* 11:477].

40 Adicionalmente, se pueden usar bancos de fragmentos de PSE para la obtención de una población variegada de fragmentos de PSE para revisar y luego seleccionar homólogos de una PSE. En una modalidad, se puede generar un banco de fragmentos de la secuencia codificante, tratando un fragmento de PCR bicatenario de una secuencia de PSE codificante con una nucleasa bajo condiciones, en las que tienen lugar roturas bicatenarias únicamente aproximadamente una vez por molécula, desnaturalizando el ADN bicatenario, renaturalizando el ADN con formación de ADN bicatenario, que puede comprender pares sentido/antisentido de diferentes productos con roturas bicatenarias, eliminando secciones monocatenarias de los dúplex recién formados por tratamiento con S1-nucleasa, y ligando el banco de fragmentos resultante en un vector de expresión. Con este procedimiento se puede derivar un banco de expresión, que codifica a fragmentos N-terminales, C-terminales e internos de PSE de diferente tamaño.

45 En el campo técnico se conocen numerosas técnicas para la clasificación de productos genéticos en bancos combinatorios, obtenidos por mutaciones puntuales o acortamiento, y para la clasificación de bancos de cADN con respecto a productos genéticos con una propiedad seleccionada. Estas técnicas se pueden adaptar a la clasificación rápida de bancos genéticos, generados por mutagénesis combinatoria de homólogos de PSE.

50 Las técnicas más usadas para la clasificación de grandes bancos genéticos que pueden someterse a una análisis de alto rendimiento comprenden normalmente la clonación del banco genético en vectores de expresión replicables, la

ES 2 331 228 T3

transformación de células apropiadas con el banco de vectores apropiado y la expresión de los genes combinatorios bajo condiciones que facilitan la detección de la actividad deseada, el aislamiento del vector que codifica al gen cuyo producto se ha detectado. La mutagénesis recursive-ensemble (REM), es una técnica nueva que aumenta la frecuencia de mutantes funcionales en los bancos, se puede usar en combinación con los ensayos de clasificación para la identificación de homólogos de PSE [Arkin y Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6(3):327-331]. También se pueden emplear combinaciones ventajosas de los métodos arriba mencionados.

Otra técnica conocida para la modificación de propiedades catalíticas de enzimas o bien sus genes codificantes es el llamado "Gen-Shuffling" (ver por ejemplo WO 97/20078 o WO 98/13487), que representa una combinación de fragmentos de gen, y se puede variar esta nueva combinación adicionalmente por reacciones defectuosas en cadena de polimerasa, creando de este modo una alta diversidad de secuencias a analizar. Condición para aplicar un enfoque de este tipo es, sin embargo, un sistema adecuado de screening, para comprobar la diversidad genética obtenida en cuanto a su funcionalidad.

Especialmente, para la clasificación de las actividades de elongasa, es necesario un procedimiento de clasificación que registra actividad(es) enzimática(s) en función de PUFA.

Con respecto a las actividades de elongasa con especificidad para PUFAs, es posible aprovechar la toxicidad del ácido araquidónico en presencia de un metabolito tóxico (aquí: ácido salicílico o derivados de ácido salicílico) en especies de *Mucor*, que se pueden transformar mediante técnicas de transformación conocidas con constructos de gen deseados, (Eroshin *et al.*, Mikrobiologiya, Vol. 65, No.1, 1996, páginas 31-36), para realizar una primera clasificación a base del crecimiento.

A continuación, los clones resultantes se pueden someter a un análisis de sus respectivos ingredientes lipídicos mediante cromatografía de gas y espectroscopía de masa, con el fin de detectar eductos y productos según tipo y cantidad.

En otra modalidad, se pueden usar ensayos a base de células para el análisis de un banco de PSE variegado, aplicando otros procedimientos conocidos en el campo técnico.

En otra modalidad, se refiere la presente invención a un anticuerpo, que se liga específicamente con el polipéptido de la presente invención o partes, por ejemplo epítopes, de una proteína de este tipo.

El anticuerpo de la invención se puede usar para identificar otras elongasas, especialmente, PSEs, y para aislarlas. Estos anticuerpos pueden ser anticuerpos monoclonales, anticuerpos policlonales o anticuerpos sintéticos, así como fragmentos de estos anticuerpos como, por ejemplo, fragmentos de Fab, Fv o scFV, etc. Los anticuerpos monoclonales se pueden obtener, por ejemplo, según procedimientos, tal y como se describen originalmente por Kohler y Milstein, en Nature 256 (1975), 485, y Galfrö, Meth. Enzymol. 73 (1981). Anticuerpos y fragmentos de los mismos también pueden obtenerse, por ejemplo, según Harlow & Lane, "Antibodies, a Laboratory Manual", CSH, Press, Cold Spring Harbor, 1988.

Estos anticuerpos se pueden usar, por ejemplo, para precipitar y localizar a las proteínas de la invención o para examinar la síntesis de estas proteínas, por ejemplo en organismos recombinantes, así como para identificar compuestos que interactúan con las proteínas de la invención. En numerosos casos, el enlace de anticuerpos con antígenos es equivalente al enlace de otros ligandos y antiligandos.

Además, la presente invención describe un procedimiento para identificación de un agonista o antagonista de elongasas, especialmente, de PSE, que comprende:

- a) poner en contacto las células, que expresan al polipéptido de la presente invención, con una sustancia candidato;
- b) ensayar la actividad de la PSE;
- c) comparar la actividad de PSE con una actividad estándar en ausencia de la sustancia candidato, en lo cual un incremento de la actividad de PSE en comparación con el estándar indica que la sustancia candidato es un agonista, y una menor actividad en comparación con el estándar indica que la sustancia candidato es un antagonista.

La sustancia candidato mencionada puede ser una sustancia sintetizada químicamente o ser una sustancia producida microbiológicamente y estar presente en extractos de células, por ejemplo, de plantas, animales o microorganismos. Además, la sustancia mencionada puede ser conocida como tal en el estado de la técnica, pero no es conocido hasta ahora su efecto intensificante o reprimiente de la actividad de PSEs. La mezcla de reacción puede comprender un extracto desprovisto de células o comprender una célula o cultivo celular. Métodos apropiados son conocidos a la persona técnica en la materia y se describen en términos generales, por ejemplo, en Alberts, Molecular Biology the cell, 3a Edición (1994), por ejemplo capítulo 17. Las sustancias mencionadas pueden adicionarse, por ejemplo, a la mezcla de reacción o al medio de cultivo o ser inyectadas en las células o ser pulverizadas sobre una planta.

Cuando se ha identificado una muestra, que contiene una sustancia activa según el método descrito, entonces se puede aislar la sustancia directamente de la muestra original o se puede dividir la muestra en diferentes grupos, por ejemplo cuando se compone de un gran número de diferentes componentes, para así reducir el número de las diferentes sustancias por muestra, y entonces repetir el procedimiento de la invención con una “submuestra” de la muestra original. Dependiendo de la complejidad de la muestra, se pueden repetir los pasos arriba descritos varias veces, preferentemente, hasta que la muestra identificada según los métodos descritos sólo comprenda un número reducido de sustancias o únicamente una sustancia. Preferentemente, se siguen formulando la sustancia identificada, o los derivados de la misma, según el método descrito, de manera que sea apropiada para ser usada en el cultivo de plantas o de células vegetales o de tejidos.

Las sustancias que se ensayaron e identificaron según el procedimiento descrito pueden ser: bibliotecas de expresión, por ejemplo bibliotecas de expresión de cADN, péptidos, proteínas, ácidos nucleicos, anticuerpos, pequeñas sustancias orgánicas, hormonas, PNAS o similares (Milner, *Nature Medicin* (1995), 879-880; Hupp, *Cell*. 83 (1995), 237-245; Gibbs, *Cell*. 79 (1994), 193-198 y las referencias allí citadas). Esas sustancias también pueden ser derivados funcionales o análogos de los inhibidores o activadores conocidos. Procedimientos para la obtención de derivados químicos o análogos son conocidos a la persona técnica en la materia.

Los derivados o análogos mencionados pueden ser ensayados mediante procedimientos del estado de la técnica. Además, se puede usar un diseño asistido por ordenador o peptidomimética para la obtención de derivados o análogos adecuados.

Las células o el tejido que pueden ser usados para el procedimiento de la invención, son preferentemente una célula huésped, célula vegetal o un tejido vegetal, como se describe en las modalidades arriba expuestas.

Por tanto, la presente invención describe también a una sustancia, que se ha identificado según los procedimientos de la invención arriba descritos. La sustancia es por ejemplo un homólogo de la PSE según la invención. Homólogos de las PSEs se pueden generar por mutagénesis, por ejemplo por mutación puntual o delección de las PSE. Tal como aquí se usa, el término “homólogo” es una forma variante de las PSEs, que actúa como agonista o antagonista para la actividad de las PSEs. Un agonista puede tener la actividad substancialmente idéntica o una parte de la actividad biológica de las PSEs. Un antagonista de las PSEs puede inhibir una o más actividades de las formas nativas de las PSEs, por ejemplo puede ligarse en forma competitiva con un miembro situado Downstream o Upstream de las vías metabólicas de la síntesis de ácido graso, que incluyen a las PSEs, o ligarse con PSEs y en este caso reducir o inhibir la actividad.

Por lo tanto, la presente invención se refiere también a un anticuerpo o fragmento del mismo, tal y como se describen en la presente, que inhibe a la actividad de las PSEs de la invención.

En un aspecto, la presente invención describe a un anticuerpo que reconoce o bien liga específicamente al agonista o antagonista arriba descrito.

Otro aspecto se refiere a una composición, que comprende al anticuerpo, el stop identificado según el procedimiento de la invención o la molécula antisentido.

D. Aplicaciones y procedimientos según la invención

Las moléculas de ácido nucleico, proteínas, homólogos de proteína, proteínas de fusión, anticuerpos, primers, vectores y células huésped, aquí descritos, se pueden usar en uno o varios de los procedimientos siguientes: identificación de *Physcomitrella patens*, *Cryptocodium*, *Phytophthora infestans* o *Thraustochytrium* y organismos similares, mapping de los genomas de organismos, que son similares a *Physcomitrella*, *Phytophthora*, *Cryptocodium* o *Thraustochytrium*, identificación y localización de las secuencias de *Physcomitrella*, *Phytophthora*, *Cryptocodium* o *Thraustochytrium* de interés, estudios de evolución, determinación de regiones de proteínas de PSE, que son necesarios para la función, modulación de una actividad de PSE; modulación del metabolismo de uno o varios componentes de membrana celular; modulación del transporte a través de la membrana de unos o más compuestos, así como la modulación de la producción celular de un compuesto deseado, como de un producto químico fino. Las moléculas de ácido nucleico de PSE según la invención tienen múltiples aplicaciones. Pueden ser usadas para la identificación de un organismo como *Physcomitrella*, *Phytophthora* o *Thraustochytrium* o los más relacionados a ellos. Pueden ser usadas también para la identificación de la presencia de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* o de un organismo relacionado con los mismos en una población mixta de microorganismos.

La invención proporciona las secuencias de ácido nucleico de una serie de genes de *Physcomitrella*, *Phytophthora* o *Thraustochytrium*; por sondeo del ADN genómico extraído de un cultivo de una población homogénea o mixta de microorganismos, en condiciones severas con una sonda, que abarcan una región de un gen de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* o partes de los mismos, que es única para este organismo, se puede determinar si este organismo está presente. *Physcomitrella*, *Phytophthora* o *Thraustochytrium* mismos se usan para la producción comercial de ácidos poliinsaturados. Los ácidos nucleicos de la invención son apropiados, además, para la producción de PUFA también en otros organismos, especialmente, cuando se desea lograr que los PUFAs resultantes también sean insertados en fracciones de triacilglicerol.

Además, las moléculas de ácido nucleico y las moléculas de proteína de la invención pueden servir de marcadores para regiones específicas de genoma. Estas son apropiadas no solamente para el mapping del genoma, sino también para estudios funcionales de proteínas de *Physcomitrella*, *Phytophthora*, o *Thraustochytrium*. Para la identificación de la región del genoma, en la que se liga una determinada proteína ligante de ADN de *Physcomitrella*, *Phytophthora* o *Thraustochytrium*, se podría escindir, por ejemplo, el genoma de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* e incubar los fragmentos con la proteína ligante de ADN. Aquellos, que ligan a la proteína pueden ser detectados conjuntamente con las moléculas de ácido nucleico de la invención, preferentemente con marcaciones fáciles de detectar; el enlace de una molécula de ácido nucleico de este tipo con el fragmento del genoma permite localizar el fragmento en el mapa genómico de *Physcomitrella*, *Phytophthora* o *Thraustochytrium* y facilita, cuando esto se realiza varias veces con diferentes enzimas, una determinación rápida de la secuencia de ácido nucleico a la que se liga la proteína. Las moléculas de ácido nucleico de la invención pueden ser, además, suficientemente homólogas con las secuencias de especies relacionadas, para que estas moléculas de ácido nucleico puedan servir como marcadores para la construcción de un mapa genómico en hongos o algas similares.

Las moléculas de ácido nucleico de PSE según la invención también son apropiadas para exámenes de evolución y de la estructura proteica. Los procesos metabólicos y de transporte, en los que participan las moléculas de la invención, son aprovechadas por numerosas células procariontas y eucariotas; comparando las secuencias de las moléculas de ácido nucleico de la invención con aquellas que codifican a enzimas similares de otros organismos, se puede determinar el grado de parentesco evolutivo de los organismos.

Una comparación de este tipo permite correspondientemente la determinación de cuáles regiones de secuencia están conservadas y cuáles no están conservadas, lo que es útil al determinar las regiones de la proteína, que son esenciales para la función enzimática. Este tipo de determinación es valioso para exámenes de ingeniería proteica y puede ser un indicio de cuanta mutagénesis puede ser tolerada por la proteína sin que ella pierda su función.

La manipulación de las moléculas de ácido nucleico de PSE de la invención puede resultar en la producción de PSEs con diferencias funcionales con respecto al tipo silvestre de las PSEs. La eficiencia o actividad de estas proteínas puede ser mejorada, pueden estar presentes en un mayor número que normalmente en la célula, o su eficiencia o actividad puede ser reducida. Una eficiencia o actividad mejorada significa, por ejemplo, que la enzima tiene una selectividad y/o actividad más alta, preferentemente una actividad por lo menos 10% más alta, muy preferentemente por lo menos 20% más alta, especialmente por lo menos 30% más alta, que la enzima original.

Hay una serie de mecanismos, mediante los cuales la modificación de una PSE de la invención puede influir directamente en el rendimiento, la producción y/o la eficiencia de la producción de un producto químico fino, que contiene tal proteína modificada. La obtención de compuestos de químicos finos a partir de cultivos de ciliatos, algas u hongos a gran escala es significativamente mejorada, cuando la célula segrega los compuestos deseados, ya que estos compuestos pueden ser purificados fácilmente del medio de cultivo (contrario a la extracción de la masa de las células cultivadas). Por lo demás, se puede mejorar la purificación cuando la célula almacena *in vivo* compuestos en un compartimento especializado con un tipo de mecanismo de concentración.

En plantas que expresan las PSEs, un transporte incrementado puede resultar en una mejor distribución dentro del tejido vegetal o los órganos de la planta. Al aumentar el número o la actividad de las moléculas de transporte, que exportan químicos finos fuera de la célula, puede ser posible aumentar la cantidad del producto químico fino producido que está presente en el medio extracelular, lo que puede facilitar la cosecha y purificación o una distribución más eficiente en el caso de plantas. Para la sobreproducción eficiente de uno o más químicos finos se necesitan mayores cantidades de cofactores, moléculas precursoras y compuestos intermedios para las vías de biosíntesis apropiadas.

Aumentando el número y/o la actividad de las proteínas de transporte, que participan en la importación de sustancias nutritivas, tales como fuentes de carbono (a saber, azúcares), fuentes de nitrógeno (a saber, aminoácidos, sales de amonio), fosfato y azufre, se puede mejorar la producción de un químico fino gracias a la eliminación de todas las limitaciones de la oferta de sustancias nutritivas en el proceso de biosíntesis. Los ácidos grasos, tales como PUFAs, y lípidos mismos, que contienen PUFAs son químicos finos deseables; por tanto, optimizando la actividad y aumentando el número de una o más PSEs de la invención, que participan en la biosíntesis de estos compuestos, o interfiriendo en la actividad de una o más PSEs, que participan en la degradación de estos compuestos, se puede incrementar el rendimiento, la producción y/o la eficiencia de la producción de moléculas de ácido graso y de lípidos en ciliatos, algas, hongos, levaduras u otros microorganismos.

La manipulación de uno o más genes de PSE de la invención también puede producir genes de PSE de la invención con actividades modificadas, que afectan indirectamente la producción de uno o varios químicos finos a partir de algas, plantas, ciliatos u hongos. Los procesos metabólicos bioquímicos normales resultan, por ejemplo, en la producción de múltiples productos de degradación (por ejemplo, peróxido de hidrógeno y otras especies de oxígeno reactivo), que pueden interferir activamente en estos procesos metabólicos [por ejemplo, el peroxinitrito nítrico, como se sabe, las cadenas laterales de tirosina, por lo que algunas enzimas con tirosina en el centro activo se desactivan, Groves, J.T. (1999) *Curr. Opin. Chem. Biol.* 3(2):226-235]. Estos productos de degradación suelen ser eliminados, pero las células usadas para la producción fermentativa a gran escala son optimizadas para la sobreproducción de uno o varios químicos finos, por lo que pueden producir más productos de degradación que los usuales para una célula de tipo silvestre. Optimizando la actividad de una o varias PSEs de la invención, que participan en la exportación de moléculas de degradación, se puede mejorar la viabilidad de la célula y mantener una actividad metabólica más eficiente. La

presencia de altas cantidades intracelulares del producto químico fino deseado también puede, en efecto, ser tóxica para la célula, de manera que al aumentar la capacidad de la célula para la secreción de estos compuestos puede mejorar la viabilidad de la célula.

5 Además, las PSEs de la invención pueden ser manipuladas de tal forma, que se modifiquen las cantidades relativas de diferentes moléculas de lípido y ácido graso. Esto puede afectar decisivamente la composición de lípido de la membrana celular. Ya que cada tipo de lípido tiene diferentes propiedades físicas, la modificación de la composición de lípido de una membrana puede modificar significativamente la fluidez de la membrana. Modificaciones de la fluidez de la membrana pueden influir en el transporte de moléculas por las membranas, lo que, como se explicó más arriba, puede modificar la exportación de productos de degradación o del producto químico fino producido o la importación de sustancias nutritivas necesarias. Estas modificaciones de la fluidez de la membrana también pueden influir decisivamente en la integridad de la célula; células con membranas comparablemente más débiles son más sensibles frente a condiciones de estrés abiótico y biótico, que pueden dañar o bien matar a la célula. Por manipulaciones de las PSEs, que participan en la producción de ácidos grasos y lípidos para la construcción de la membrana, de manera que la membrana resultante tenga una composición que es más sensible para las condiciones ambientales que rigen en los cultivos utilizados para la producción de químicos finos, una mayor porción de las células debería sobrevivir y reproducirse. Cantidades más altas en células productoras deberían manifestarse en un mayor rendimiento, una mayor producción o eficiencia de producción del producto químico fino a partir del cultivo.

20 Las estrategias de mutagénesis arriba mencionadas para PSEs que deben resultar en un mayor rendimiento de un químico fino, no deberían ser limitativas; variaciones de estas estrategias son fácilmente evidentes para la persona técnica en la materia. Aplicando estos mecanismos y con la ayuda de los mecanismos divulgados en la presente, se pueden usar las moléculas de ácido nucleico y moléculas de proteínas de la invención para la generación de algas, ciliatos, plantas, animales, hongos u otros microorganismos, tales como *C. glutamicum*, que expresan a moléculas de ácido nucleico de PSE y moléculas de proteínas mutadas, aumentando así el rendimiento, la producción y/o la eficiencia de producción de un compuesto deseado. Este compuesto deseado puede ser un producto natural cualquiera de algas, ciliatos, plantas, animales, hongos o bacterias, que abarca los productos finales de las vías de biosíntesis y productos intermedios de las vías metabólicas de procedencia natural, así como moléculas, que existen de manera no natural en el metabolismo de estas células, pero que son producidas por las células de la invención.

30 Otra modalidad según la invención se refiere a un procedimiento para la producción de PUFAs, y este procedimiento comprende el cultivo de un organismo no humano que contiene un ácido nucleico de la invención, un constructo génico según la invención o un vector según la invención, que codifican para un polipéptido, que alongan ácidos grasos con 16, 18 o bien 20 átomos de carbono con por lo menos dos dobles enlaces en la molécula de ácido graso en por lo menos dos átomos de carbono en condiciones, bajo las cuales los PUFAs son producidos en el organismo. PUFAs producidos mediante este procedimiento se pueden aislar por cosecha de los organismos o bien a partir del cultivo, en el que crecen, o bien del campo, del fraccionamiento y/o la extracción con un disolvente orgánico del material cosechado. A partir de este disolvente se puede aislar el aceite, que contiene lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerinas y/o ácidos grasos libres con contenido más alto de PUFAs. Por hidrólisis básica o ácida de los lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerinas se pueden aislar los ácidos grasos libres con un contenido más alto en PUFAs. Un contenido más alto en PUFAs significa por lo menos 5%, preferentemente 10%, muy preferentemente 20%, especialmente 40% de PUFAs más que en el organismo original, que no posee ningún ácido nucleico adicional, que codifica para la elongasa de la invención.

45 Preferentemente, los PUFAs obtenidos según este procedimiento son moléculas de ácido graso con 18, 20 ó 22 átomos de carbono con por lo menos dos dobles enlaces en la molécula de ácido graso, preferentemente tres, cuatro, cinco o seis dobles enlaces, muy preferentemente tres o cinco dobles enlaces. Estas moléculas de ácido graso con 18, 20 ó 22 átomos de carbono se pueden aislar del organismo en forma de un aceite, lípido o un ácido graso libre. Organismos apropiados son, por ejemplo, los arriba mencionados. Organismos preferidos son los microorganismos, animales o plantas, muy preferentemente las plantas o algas, especialmente las plantas transgénicas.

50 La invención describe aceites, lípidos o ácidos grasos o fracciones de los mismos, que han sido preparados mediante el procedimiento arriba descrito, particularmente preferible aceite, lípido o una composición de ácido graso, que comprenden los PUFAs y que se derivan de plantas transgénicas.

55 La invención describe aceites, lípidos o ácidos grasos, preparados según el procedimiento de la invención. La invención también describe composiciones de aceite, lípido o ácido graso, que contiene PUFAs preparados según el procedimiento de la invención y que provienen de plantas transgénicas que contienen un ácido nucleico de la invención, un constructo génico o un vector.

60 La invención describe el uso del aceite, lípido o de la composición de ácido graso en alimentos para animales, alimentos para humanos, cosméticos o fármacos.

65 La presente invención describe un kit, que comprende el ácido nucleico de la invención, el constructo de gen de la invención, la secuencia de aminoácido de la invención, la molécula de ácido nucleico antisentido de la invención, el anticuerpo y/o la composición de la invención, un antagonista o agonista que se ha obtenido según el procedimiento descrito, y/o aceites, lípidos y/o ácidos grasos o una fracción de los mismos. El kit puede comprender también las células huésped, organismos, plantas o partes de las mismas, según la invención, partes cosechables de las plantas

de la invención o material de reproducción, pero también el antagonista o agonista descritos. Los componentes del kit según la invención pueden estar empacados en contenedores apropiados, por ejemplo, con o en búferes u otras soluciones. Uno o más de los componentes mencionados pueden estar empacados en el mismo contenedor.

5 Adicional o alternativamente uno o más de los componentes mencionados pueden estar adsorbidos en una superficie sólida, por ejemplo un filtro de nitrocelulosa, placas de vidrio, chips, membranas de nilón o placas de microtitulación. El kit puede usarse para cada uno de los métodos o modalidades aquí descritos, por ejemplo para la producción de células huésped, plantas transgénicas, para la detección de secuencias homólogas, para la identificación de antagonistas o agonistas, etc. Además, el kit puede contener indicaciones para el uso del kit para una de las aplicaciones mencionadas.

Esta invención se ilustra en más detalle en los ejemplos siguientes, que no deben entenderse como limitativos.

15 Ejemplos

Ejemplo 1

Procedimientos generales

a) *Procedimientos generales de clonación*

Los procedimientos de clonación, como por ejemplo, escisiones de restricción, electroforesis de gel de agarosa, purificación de fragmentos de ADN, transferencia de ácidos nucleicos sobre membranas de nitrocelulosa y nilón, enlace de fragmentos de ADN, transformación de células de *Escherichia coli* y de levadura, cultivo de bacterias y análisis secuencial de ADN recombinante, se realizaron en la forma descrita en: Sambrook *et al.* [(1989), Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6] o Kaiser, Michaelis y Mitchell [(1994), "Methods in Yeast Genetics", Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3]. La transformación y el cultivo de algas, tales como *Chlorella* o *Phaeodactylum* se realizan en la forma descrita en: El-Sheekh [(1999), *Biologia Plantarum* 42:209-216] y con la ayuda de los mecanismos divulgados en la presente, se pueden usar las moléculas de ácido nucleico y moléculas de proteínas de la invención para la generación de algas, ciliatos, plantas, animales, hongos u otros microorganismos, tales como *C. glutamicum*, que expresan a moléculas de ácido nucleico de PSE y moléculas de proteínas mutadas, aumentando así el rendimiento, la producción y/o la eficiencia de producción de un compuesto deseado. Este compuesto deseado puede ser un producto natural cualquiera de algas, ciliatos, plantas, animales, hongos o bacterias, que abarca los productos finales de las vías de biosíntesis y productos intermedios de las vías metabólicas de procedencia natural, así como moléculas, que existen de manera no natural en el metabolismo de estas células, pero que son producidas por las células de la invención.

Otra modalidad según la invención se refiere a un procedimiento para la producción de PUFAs, y este procedimiento comprende el cultivo de un organismo que contiene un ácido nucleico de la invención, un constructo génico según la invención o un vector según la invención, que codifican para un polipéptido, que alongan ácidos grasos con 16, 18 o bien 20 átomos de carbono con por lo menos dos dobles enlaces en la molécula de ácido graso en por lo menos dos átomos de carbono en condiciones, bajo las cuales los PUFAs son producidos en el organismo. PUFAs producidos mediante este procedimiento se pueden aislar por cosecha de los organismos o bien a partir del cultivo en el que crecen, o bien del campo, del fraccionamiento y/o la extracción con un disolvente orgánico del material cosechado. A partir de este disolvente se puede aislar el aceite, que contiene lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerinas y/o ácidos grasos libres con contenido más alto de PUFAs. Por hidrólisis básica o ácida de los lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerinas se pueden aislar los ácidos grasos libres con un contenido más alto en PUFAs. Un contenido más alto de PUFAs significa por lo menos 5%, preferentemente 10%, muy preferentemente 20%, especialmente 40% de PUFAs más que en el organismo original, que no posee ningún ácido nucleico adicional, que codifica para la elongasa de la invención.

Preferentemente, los PUFAs obtenidos según este procedimiento son moléculas de ácido graso con 18, 20 ó 22 átomos de carbono con por lo menos dos dobles enlaces en la molécula de ácido graso, preferentemente tres, cuatro, cinco o seis dobles enlaces, muy preferentemente tres o cinco dobles enlaces. Estas moléculas de ácido graso con 18, 20 ó 22 átomos de carbono se pueden aislar del organismo en forma de un aceite, lípido o un ácido graso libre. Organismos apropiados son, por ejemplo, los arriba mencionados. Organismos preferidos son los microorganismos, animales o plantas, muy preferentemente las plantas o algas, especialmente las plantas transgénicas.

Una modalidad de la invención son aceites, lípidos o ácidos grasos o fracciones de los mismos, que han sido preparados mediante el procedimiento arriba descrito, particularmente preferible aceite, lípido o una composición de ácido graso, que comprenden los PUFAs y que se derivan de plantas transgénicas.

Una modalidad de la invención son aceites, lípidos o ácidos grasos, preparados según el procedimiento de la invención. Otras modalidades de la invención son composiciones de aceite, lípido o ácido graso, que contiene PUFAs preparados según el procedimiento de la invención y que provienen de plantas transgénicas que contienen un ácido nucleico de la invención, un constructo génico o un vector.

ES 2 331 228 T3

Otra modalidad de acuerdo con la invención es el uso del aceite, lípido o de la composición de ácido graso en alimentos para animales, alimentos para humanos, cosméticos o fármacos.

En otra modalidad, la presente invención se refiere a un kit, que comprende el ácido nucleico de la invención, el constructo de gen de la invención, la secuencia de aminoácido de la invención, la molécula de ácido nucleico antisentido de la invención, el anticuerpo y/o la composición de la invención, un antagonista o agonista que se ha obtenido según el procedimiento de la invención, y/o aceites, lípidos y/o ácidos grasos o una fracción de los mismos, según la invención. Así mismo, el kit puede comprender también las células huésped, organismos, plantas o partes de las mismas, según la invención, partes cosechables de las plantas de la invención o material de reproducción, pero también el antagonista o agonista según la invención. Los componentes del kit según la invención pueden estar empacados en contenedores apropiados, por ejemplo, con o en búferes u otras soluciones. Uno o más de los componentes mencionados pueden estar empacados en el mismo contenedor. Adicional o alternativamente uno o más de los componentes mencionados pueden estar adsorbidos en una superficie sólida, por ejemplo un filtro de nitrocelulosa, placas de vidrio, chips, membranas de nilón o placas de microtitulación. El kit puede usarse para cada uno de los métodos o modalidades aquí descritos, por ejemplo para la producción de células huésped, plantas transgénicas, para la detección de secuencias homólogas, para la identificación de antagonistas o agonistas, etc. Además, el kit puede contener indicaciones para el uso del kit para una de las aplicaciones mencionadas.

Esta invención se ilustra en más detalle en los ejemplos siguientes, que no deben entenderse como limitativos. El contenido de todas las citas bibliográficas de esta solicitud de patente, las solicitudes de patentes, las patentes y las solicitudes publicadas de patentes se incorpora aquí por referencia.

Ejemplos

Ejemplo 1

Procedimientos generales

a) *Procedimientos generales de clonación*, como por ejemplo, escisiones de restricción, electroforesis de gel de agarosa, purificación de fragmentos de ADN, transferencia de ácidos nucleicos sobre membranas de nitrocelulosa y nilón, enlace de fragmentos de ADN, transformación de células de *Escherichia coli* y de levadura, cultivo de bacterias y análisis secuencial de ADN recombinante, se realizaron en la forma descrita en Sambrook *et al.* [(1989), Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6] o Kaiser, Michaelis y Mitchell [(1994), "Methods in Yeast Genetics", Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3]. La transformación y el cultivo de algas, tales como *Chlorella* o *Phaeodactylum* se realizan en la forma descrita en El-Sheekh [(1999), *Biologia Plantarum* 42:209-216] o Apt *et al.* [(1996) *Molecular and General Genetics* 252 (5):872-9].

b) *Químicos*

Los químicos usados fueron suministrados, salvo indicación en contrario, en calidad de p. a. por las compañías Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) y Sigma (Deisenhofen). Las soluciones se prepararon usando agua pura, libre de pirógeno, en lo sucesivo denominado H₂O, de una planta purificadora de agua de sistema de agua Milli-Q (Millipore, Eschborn). Endonucleasas de restricción, enzimas modificantes de ADN y kits moleculares biológicos se consiguieron de las compañías AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen) I Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) y Stratagene (Amsterdam, Holanda). Salvo indicación en contrario, se usaron según las instrucciones del fabricante.

c) *Material de célula*

Para estos exámenes se usó una cepa de *Thraustochytrium*, suministrado por American Type Culture Collection (colección de ATCC) con el número de cepa ATCC26185 (*Thraustochytrium*), o bien en el caso de *Cryptocodium* vom Provasoli-Guillard, por National Center for Culture of Marine Phytoplankton ((CCMP) West Boothbay Harbour, ME, USA) con el No. de cultivo de cepa CCMP316. Las algas se cultivaron a 25 grados Celsius en la oscuridad en tubos de vidrio gasificados desde abajo con aire. Alternativamente se cultivó *Thraustochytrium* en la forma descrita detalladamente por Singh & Ward (1997, *Advances in Microbiology*, 45, 271-311).

Como medio de cultivo para *Cryptocodium* se usó el medio de cultivo f/2 con 10% de medio orgánico según Guillard, R.R.L. [1975; *Culture of phytoplankton for feeding marine invertebrate*. En: Smith, W.L. and Chanley, M.H. (Eds.) *Culture of marine Invertebrate animals*, NY Plenum Press, pp. 29-60.]: Este contiene:

995,5 ml agua de mar (artificial) 1 ml de NaNO₃ (75 g/l), 1 ml NaH₂PO₄ (5 g/l), 1 ml de solución con trazas de metal, 1 ml de Tris/Cl pH 8.0, 0.5 ml f/2 de solución de vitamina solución de trazas de metal: Na₂EDTA (4,36 g/l), FeCl₃ (3,15 g/l), trazas de metal primarias: CuSO₄ (10 g/l), ZnSO₄ (22 g/l), CoCl₂ (10 g/l), MnCl₂ (18 g/l), NaMoO₄ (6,3 g/l) f/2 solución de vitamina: Biotina: 10 mg/l, Tiamina 200 mg/l, Vit B12 0,1 mg/l

ES 2 331 228 T3

medio orgánico: acetato de Na (1 g/l), glucosa (6 g/l), succinato de Na (3 g/l), Bacto-Trypton (4 g/l), extracto de levadura (2 g/l)

5 d) *Material de moho (= material vegetal)*

Para este ensayo se usaron plantas de la especie *Physcomitrella patens* (Hedw.) B.S.G. de la colección de la Sección de Ensayos Genéticos de la Universidad de Hamburgo. Proviene de la cepa 16/14, coleccionada por H.L.K. Whitehouse en Gransden Wood, Huntingdonshire (Inglaterra) y subcultivado por Engel (1968, Am J Bot 55, 438-446) a partir de la spora. La proliferación de las plantas se realizó mediante spora y mediante la regeneración de los gametofitos. La protonema se desarrolló a partir de la spora haploide en cloronema rico en cloroplastos y caulonema pobre en cloroplastos, a partir de los cuales se formaron botones después de aproximadamente 12 días. Estos crecieron en gametóforos con anteridios y arquegonios. Después de la polinización se formó el esporófito diploide con seta corta y cápsula esporífera corta, en la que maduran las meiosporas.

15 e) *Cultivo de plantas*

El cultivo se realizó en una cámara climatizada con una temperatura del aire de 25°C y una intensidad de luz de 55 micromol·m⁻² (luz blanca; Phillips TL 65W/25 tubo fluorescente) y un cambio de luz-oscuridad de 16/8 horas. El moho se modificó en un cultivo líquido usando medio de Knop según Reski y Abel (1985, Planta 165, 354-358) o se cultivó sobre un medio sólido de Knop usando 1% de oxidagar (Unipath, Basingstoke, Inglaterra).

Los protonemas usados para el aislamiento de ARN y ADN se cultivaron en cultivos líquidos ventilados. Los protonemas se desmenuzaron cada 9 días y se transfirieron en medio de cultivo fresco.

25 f) *Cultivo de Phytophthora infestans*

Primero, se estableció un banco de cADN de *Phytophthora infestans*. Para ello se puede obtener la cepa de ATCC 48886 de American Type Culture Collection Rockville, USA. Desviándose del protocolo de cultivo descrito para la cepa de ATCC 48886, se lavaron las esporas de *Phytophthora* con agua fría bidestilada y se colocaron por 6 horas en la nevera para la esporulación. A continuación, se transfirió el material a un medio de arveja. Para esto, se esterilizaron 150 g de arvejas congeladas (Iglu, se consiguen en los supermercados locales) en 1 litro de agua por 20 minutos en el autoclave. El material se cultiva en matraces de 100 ml a temperatura ambiente en una sacudidora circulatoria (200 rpm). Después de dos días se filtran de a 2 matraces y en los siguientes 4 días también de a 2 matraces y se desmenuzan en el mortero en nitrógeno líquido.

40 Ejemplo 2

Aislamiento de ADN total de plantas y microorganismos, tales como Thraustochytrium y Cryptocodium para experimentos de hibridación

Los detalles respecto al aislamiento del ADN total se refieren al procesamiento de material vegetal con un peso fresco de un gramo.

Búfer de CTAB: 2% (peso/vol.) de bromuro de N-acetil-N,N,N-trimetilamonio (CTAB); 100 mM de Tris-HCl, pH 8,0; 1,4 M de NaCl; 20 mM de EDTA. Búfer de N-laurilsarcosina: 10% (peso/vol.) N-laurilsarcosina; 100 mM de Tris-HCl, pH 8,0; 20 mM de EDTA.

El material vegetal o material celular de *Cryptocodium* o *Thraustochytrium* se trituró en un mortero bajo nitrógeno líquido, de modo que se obtuvo un polvo fino, y se transfirió a recipientes de Eppendorf de 2 ml. El material vegetal congelado se cubre con una capa de 1 ml de búfer de descomposición (1 ml de búfer CTAB, 100 ml de búfer N-laurilsarcosina, 20 ml de β -mercaptoetanol y 10 ml de solución K de proteinasa, 10 mg/ml) y se incubó bajo agitación continua por una hora a 60°C. El producto homogenizado obtenido se repartió en dos recipientes de Eppendorf (2 ml) y se extrajo dos veces por agitación con el mismo volumen de cloroformo/alcohol isoamílico (24:1). Para la separación de las fases se realizó una centrifugación a 8000 x g y temperatura RT (RT = temperatura ambiente = ~ 23°C) respectivamente durante 15 minutos. Luego se precipitó el ADN por 30 min usando isopropanol helado a -70°C.

El ADN precipitado se sedimentó en 10000 g por 30 min a 4°C y se re-suspendió en 180 ml de búfer de TE (Sambrook *et al.*, 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). Para seguir purificando se trató el ADN con NaCl (1,2 M de concentración final) y se precipitó nuevamente por 30 min usando el volumen doble de etanol absoluto a -70°C. Después de la etapa de lavado con 70% etanol, se secó el ADN y luego se absorbe en 50 ml de H₂O + ARN (50 mg/ml de concentración final). El ADN se disolvió durante la noche a 4°C y luego se realizó la escisión de ARNs por una hora a 37°C. El ADN se guardó a 4°C.

ES 2 331 228 T3

Ejemplo 3

Aislamiento de ARN total y poli (A) + ARN total de plantas y microorganismos (Cryptocodinium y Thraustochytrium)

5 El aislamiento de ARN total de plantas, tales como lino y colza, etc. se realiza según el método descrito por Logemann *et al* (1987, Anal. Biochem. 163, 21). A partir de moho se puede generar el ARN total de tejido de protonema según el procedimiento de GTC (Reski *et al.*, 1994, Mol. Gen. Genet., 244:352-359).

10 *Aislamiento de ARN a partir de Cryptocodinium y Thraustochytrium*

Muestras de algas ultracongeladas (-70°C) se trituran en un mortero helado bajo nitrógeno líquido para producir un polvo fino. Se agregan 2 volúmenes de medio de homogeneización (12,024 g de sorbitol, 40,0 ml de Tris-HCl 1M, pH 9 (0,2 M); 12,0 ml de NaCl 5 M (0,3 M), 8,0 ml de EDTA 250 mM, 761,0 mg de EGTA, 40,0 ml de SDS al 10% se adiciona H₂O para dar 200 ml, y el pH se regula a 8,5) y 4 volúmenes de fenol con 0,2% de mercaptoetanol a 40-50°C, mientras que se mezcla bien para producir polvo celular congelado. A continuación, se agregan 2 volúmenes de cloroformo y se agita vigorosamente por 15 minutos. Se centrifuga durante 10 min a 10000 g y se extrae la fase acuosa con fenol/cloroformo (2 vol/2 vol) y luego finalmente con cloroformo.

20 El volumen obtenido de la fase acuosa se mezcla con 1/20 vol de acetato de Na, 4 M (pH 6) y 1 vol de isopropanol (helado) y se precipitan los ácidos nucleicos a -20°C durante la noche. Se continúa centrifugando por 30 min a 10000 g y se filtra el sobrenadante por succión. Sigue una etapa de lavado con 70% de EtOH y se centrifuga nuevamente. El sedimento se introduce en búfer de Tris borato (80 mM de búfer Tris borato, 10 mM de EDTA, pH 7,0). Luego se mezcla el sobrenadante con 1/3 vol de LiCl de 8 M y se incuba por 30 min a 4°C. Después de una centrifugación adicional, se lava el sedimento con 70% de etanol, se centrifuga y el sedimento se disuelve en agua libre de ARNs.

El poli(A)+-ARN se aísla usando Dyna Beads® (Dynal, Oslo, Finlandia), según las instrucciones en el protocolo del fabricante.

30 Después de la determinación de la concentración del ARN o poli (A)+-ARN, se precipita el ARN por adición de 1/10 volúmenes de acetato de sodio a 3 M, pH 4,6, y 2 volúmenes de etanol y se guarda a -70°C.

35 Para el análisis se separan 20 µg de ARN respectivamente en un gel de agarosa al 1,5% que contiene formaldehído y se transfieren sobre una membrana de nilón (Hybond, Amersham). La detección de transcriptos específicos se realiza según lo descrito por Amasino ((1986) Anal. Biochem. 152, 304)).

Aislamiento de ARN total y poli- (A) + ARN a partir de Phytophthora infestans

40 El ARN total se obtiene con el kit de RNeasy Plant Total RNA (Qiagen, Hilden) y el búfer de RLT allí contenido, observando las instrucciones del fabricante. Del ARN generado, se aísla el poli-(A)+ ARN con el sistema de aislamiento: Poly Attract mRNA Isolation System III de la Cia. Promega (Heidelberg), observando las instrucciones del fabricante.

45

Ejemplo 4

Construcción del banco de cADN

50

Para la construcción del banco de cADN a partir de Physcomitrella, Cryptocodinium o bien Thraustochytrium se realizó la síntesis monocatenaria usando la Reverser transcriptasa a partir del virus de leucemia del ratón (Roche, Mannheim, Alemania) y primers de oligo-d(T), la síntesis bicatenaria por incubación con ADN-polimerasa I, enzima de Klenow y escisión H de ARNs a 12°C (2 horas), 16°C (1 hora) y 22°C (1 hora). La reacción se interrumpe por incubación a 65°C (10 min) y luego se transfiere sobre hielo.

Las moléculas bicatenarias de ADN se proveen de extremos lisos mediante T4-ADN-polimerasa (Roche, Mannheim) a 37°C por 25 min). Los nucleótidos se eliminan por extracción con fenol/cloroformo y columnas de centrifugación de SephadexG50. EcoRI/XhoI -adapter (Pharmacia, Freiburg, Alemania) se ligan mediante T4-ADN-ligasa (Roche, 12°C, durante la noche) con los extremos de cADN, se recortan con XhoI y se fosforilizan con polinucleótido-quinasa (Roche, 37°C, 30 min). Esta mezcla se somete a una separación con un gel de agarosa de bajo punto de fusión. Las moléculas de ADN se eluyen mediante 300 pares de bases a partir del gel, se extraen con fenol, se concentran en columnas de Elutip-D (Schleicher y Schüll, Dassel, Alemania) y se ligan con brazos de vectores y se empaquetan en fagos de lambda-ZAPII o fagos de lambda-ZAP-Express, usando los Gigapack Gold-kits (Stratagene, Amsterdam, Holanda); se usa el material del fabricante y se observan sus instrucciones.

65

La construcción de un banco de cADN para *Phytophthora infestans* se realizó tal como se describe arriba.

ES 2 331 228 T3

Ejemplo 5

Secuenciación de ADN y análisis computarizado

5 Los bancos de cADN, tal y como se describen en el Ejemplo 4, se emplearon para la secuenciación de ADN según el proceso estándar, especialmente, aplicando el procedimiento de terminación de cadena del ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction-Kit (Perkin-Elmer, Weiterstadt, Alemania). La secuenciación de clones casuales, individuales se realizó después de la generación preparativa de plásmido a partir de bancos de cADN mediante excisión de masa *in vivo* y retransformación de DHIOB sobre placas de agar (detalles acerca del material y el protocolo, de Stratagene, Amsterdam, Holanda).

10 Plásmido de ADN se preparó a partir de cultivos de *E. coli* cultivados durante la noche en caldo de Luria con ampicilina (ver Sambrook *et al.* (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)), en un roboter de preparación de ADN de Qiagen (Qiagen, Hilden) según los protocolos del fabricante.

15 Se usan primers de secuenciación con las siguientes secuencias de nucleótido:

5'-CAGGAAACAGCTATGACC-3'

20 5'-CTAAAGGGAACAAAAGCTG-3'

5'-TGTAACACGACGGCCAGT-3'

25 Las secuencias se procesaron y anotaron usando el paquete de software estándar de EST-MAX, disponible comercialmente por Bio-Max (Munich, Alemania). Usando algoritmos comparativos y valiéndose de una secuencia de búsqueda, se buscan con la ayuda del programa BLAST genes homólogos (Altschul *et al.* (1997) "Gapped BLAST and PSIBLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.). Una secuencia a partir de *Cryptocodinium* y *Thraustochytrium* con homólogos con respecto a la secuencia de búsqueda de las elongasas de moho a partir de *Physcomitrella patens* se caracterizan más detalladamente.

30

Ejemplo 6a

35 Identificación de Tc PSE1 y TCPSE2-Gens (Tc = Thraustochytrium) (así como del gen de CcPSE1 y CcPSE2 (Cc = Cryptocodinium cohnii) por comparación con el gen de PpPSE1 a partir de Physcomitrella patens

La secuencia de longitud completa de la elongasa de moho de la invención, PpPSE1, (para la denominación ver también la Tabla 2) se usó para las comparaciones de secuencias en el algoritmo de búsqueda de TBLASTN:

40 **MEVVERFYGE LDGKVSQGVN ALLGSFGVEL TDTPTTKGLP LVDSPTPIVL GVSVYLTIVI**
GGLLWIKARD LKPRASEPFL LQALVLVHNL FCFALSLYMC VGIAYQAITW RYSLWGNAYN
PKHKEMAILV YLPYMSKYVE FMDTVIMILK RSTRQISPLH VYHSSISLI WWAIAHAPG
45 **GEAYWSAALN SGVHVLMYAY YFLAACLRSS PKLKNKYLFW GRYLTQFQMF QFMLNLVQAY**
YDMKTNAPYP QWLIKILFYY MISLLFLFGN FYVQKYIKPS DKGQKGAKTE.

50 La secuencia total de nucleótido de la elongasa de moho, PpPSE1 cADN, consta de aproximadamente 1200 bp. Contiene un marco de lectura abierto de 873 bp, que codifica a 290 aminoácidos con una masa molecular calculada de 33,4 Da. La secuencia proteica sólo posee una identidad del 38,5% y una semejanza del 48,3% con un producto genético, por ejemplo PSE1, de *Saccharomyces cerevisiae*, que es necesaria para la elongación de ácidos grasos con longitud de cadena mediana en la levadura (Toke & Martin, 1996, isolation and characterization of a gene affecting fatty acid elongation in *Saccharomyces cerevisiae*. Journal of Biological Chemistry 271, 18413-18422).

55 Las secuencias de EST, CC001042041R, así como TC002034029R y TC002014093R, se tomaron en consideración al comienzo gracias a sus débiles homologías con la elongasa de *Physcomitrella patens* (ver la Tabla 2), con el gen de PSE1, entre otros genes candidatos, como gen destino. En la Figura 5 está representado el resultado de la comparación de la secuencia peptídica de PpPSE1 con la secuencia hallada.

60 Forma parte del ácido nucleico de la invención de Seq ID NO:3. (nombre genético: Tc PSE1, No. de banco de datos propio de los inventores: TC002034029R). Las letras significan aminoácidos idénticos, mientras que el signo "más" significa un aminoácido químicamente similar. Las identidades o bien homologías de todas las secuencias halladas según la invención están resumidas en la Tabla 3.

65 La secuenciación del fragmento completo de cADN del clon de TC002034029R dio una secuencia con una longitud de 693 pares de bases, comenzando con una primera base en una trama de lectura abierta. La secuencia codifica para un polipéptido de 195 aminoácidos representado en SEQ ID 25 NO: 4 con un codon de stop en la posición del par

ES 2 331 228 T3

de bases traducido de SEQ ID NO:3 en la posición del par de bases 586-588. El clon TC002014093R comprende un polipéptido de elongasa casi completo, como puede desprenderse de la comparación de secuencia en la Figura 7.

5 Las rayas indican aminoácidos idénticos, mientras que los dobles puntos y puntos individuales representan aminoácidos químicamente intercambiables, a saber, químicamente equivalentes. La comparación se realiza con la matriz de sustitución BLOSUM62 para aminoácidos según Henikoff & Henikoff ((1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci USA 89: 10915-10919). Parámetros usados: Gap Weight: 8; Average Match: 2.912, Length Weight: 2, Average Mismatch: -2.003.

10 Además, se identificó un segundo est mediante comparación de secuencia. En la Figura 6 está representada la comparación de una secuencia peptídica de PpPSE1 con la secuencia hallada. Aunque la homología entre parámetros seleccionados se limita a unos pocos aminoácidos, esto se refiere a una región fuertemente conservada de las elongasas específicas de PUFA. Por tanto, se determinó la secuencia completa del fragmente clonado.

15 La secuenciación del fragmento completo de cADN del clon de TC002014093R dio una secuencia con una longitud de 955 pares de bases comenzando con una primera base en la trama de lectura abierta. Esta se indica según la invención con SEQ ID NR: 5. La secuencia codifica para un polipéptido de 297 aminoácidos con un codón de stop en la posición del par de bases 892-894, representado según la invención en SEQ ID NO: 6.

20 Con la ayuda de la secuencia PpPSE1 se identificó el est CC001042041R de *Cryptocodium cohnii*, que codifica para el gen CcPSE1. El est CC001042041R aislado, representado según la invención como SEQ ID NR: 7 tiene una longitud de 708 pares de bases con una trama de lectura abierta a partir de la primera base de 642 pares de bases, codificante para 214 aminoácidos y con un codon de stop en la posición 643-645. La secuencia de aminoácido hasta el codon de stop está representado según la invención en SEQ ID NR:8.

25 Además de la semejanza con el producto genético PSE1, también se puede recurrir a la similitud con la elongasa de *Saccharomyces cerevisiae* - (sce elo 1P), que es necesaria para la elongación de ácidos grasos con longitud de cadena mediana en la levadura (Toke & Martin, 1996, isolation and characterization of a gene affecting fatty acid elongation in *Saccharomyces cerevisiae*. Journal of Biological Chemistry 271, 18413-18422). En la Tabla 3 están representadas las identidades y homologías de las elongasas según la invención a partir de *Physcomitrella patens* y a partir de levadura. Los datos se obtuvieron con la ayuda del programa GAP como programa parcial del software siguiente: Wisconsin Package Version 10.0 (Genetics Computer Group (GCG), Madison, Wisc., USA).

TABLA 3

35

Identid ad/ homología	TC_ PSE1	TC_ PSE2	Pp_PSE 1	Sce elo 1P
40 Cc_PS E1	47,1 % / 40,2%	50,6 % / 43,5%	38,5% / 29,4%	45,1% / 33,5%
45 Tc_PS E1	100 / 100	n.d.	43,2% / 32,7%	41,9% / 29,9%
Tc_PS E2	41,7 % / 29,5%	100 / 100	39,2% / 30,0	35,4% / 27,8%

50 Especialmente, se pueden derivar a partir de las Figuras 5 a 10 los siguientes motivos de secuencia como regiones de alta homología y las secuencias de consenso correspondientes derivadas a partir de los mismos, que pueden llevar por traducción inversa de los aminoácidos en codones de tres pares de bases hacia oligonucleótidos, que son pueden utilizarse para el aislamiento de nuevas elongasas mediante reacción en cadena de polimerasa. Estos son especialmente los motivos de secuencia representados en la Figura 10. A partir de estos motivos se pueden derivar oligonucleótidos y se puede usar en combinación con dos oligonucleótidos en experimentos de PCR para el aislamiento de otros fragmentos de elongasa.

60 Para tal fin, se construye y sintetiza, convenientemente, un oligonucleótido que hace juego con la cadena de 5'-3' definida convencionalmente y un segundo oligonucleótido que hace juego con otra cadena de 3'-5' corriente abajo. Aquí se obtiene un número definido de combinaciones de primers, que está limitado por la permutación de posibles variantes.

65 También se pueden usar oligo-dT primers y variantes de los mismos, por ejemplo en que la última base permite una especificidad para un pool de transcritos, como por ejemplo: Oligo dT (12-20) X, donde X puede ser G, C o T. También aquí se puede usar una segunda base de Oligo dT (12-20) XY, donde X puede ser G, C o A, mientras que Y puede ser A, G, C o T.

ES 2 331 228 T3

A partir de las secuencias arriba definidas se pueden derivar secuencias de oligonucleótidos 17- a 20meros, que se pueden usar para el aislamiento de fragmentos de genes variando la combinación de los primers y los parámetros de ensayo, tales como el programa de temperatura, la concentración en iones de Mg, etc. Los fragmentos así obtenidos se pueden clonar en vectores apropiados y la secuencia de los clones obtenidos se puede determinar para identificación de nuevas elongasas según procedimiento convencionales.

Ejemplo 6b

10 *Aislamiento del clon de cADN de Pythophthora infestans*

El clon de cADN con la denominación PI001002014R de *Phytophthora infestans* se identificó a base de cAONs identificados casualmente debido a sus homologías con la elongasa de PUFA a partir del mohó *Physcomitrella patens* (ATCCC 48886). El clon contiene el motivo de consenso representado en la Figura 8: MYxYYF, donde contrario a las elongasas de PUFA identificadas hasta ahora, se encontró como aminoácido variable x un residuo de treonina. Esta variación adicional se puede usar para la derivación de primers de PCR.

Ejemplo 7

20 *Identificación de genes mediante hibridación (TC002034029R-11 iGenTc-PCE1)*

Secuencias de genes se pueden usar para la identificación de genes homólogos o heterólogos a partir de bancos de cADN o genómicos.

25 Genes homólogos (a saber, clones de cADN de longitud completa, que son homólogos, o bien homólogos) se pueden aislar mediante hibridación de ácido nucleico, usando, por ejemplo, bancos de cADN: Especialmente, para el aislamiento de genes de longitud completa, funcionalmente activos de SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO:7, SEQ ID NO:9 y SEQ ID NO:11 se puede aplicar este método.

30 Dependiendo de la frecuencia del gen de interés, se cultivan 100000 a 1000000 bacteriófagos recombinante en platos y se transfieren en una membrana de nilón. Después de la desnaturalización con álcali se inmoviliza el ADN sobre la membrana, por ejemplo por reticulación ultravioleta. La hibridación se realiza en condiciones sumamente severas.

35 La hibridación y las etapas de lavado se realizan en solución acuosa con una fuerza iónica de 1 M de NaCl y una temperatura de 68°C.

40 Sondas de hibridación se prepararon, por ejemplo por marcación mediante transcripción de nick (32P) radioactiva (High Prime, Roche, Mannheim, Alemania). Las señales se detectan mediante autoradiografía.

45 Genes parcialmente homólogos o heterólogos que son similares, pero no idénticos, se pueden identificar en forma análoga al procedimiento arriba descrito, usando condiciones de hibridación y de lavado poco severas. Para la hibridación acuosa, se suele mantener la fuerza iónica en 1M de NaCl, bajando la temperatura gradualmente de 68 a 42°C.

El aislamiento de secuencias de genes, que únicamente presentan homologías con un sólo dominio de, por ejemplo, 10 a 20 aminoácidos, se puede realizar usando sondas de oligonucleótidos sintéticos marcados radioactivamente.

50 Oligonucleótidos marcados radioactivamente se preparan fosforilando el extremo 5' de dos oligonucleótidos complementarios con T4-polinucleótido-quinasa. Los oligonucleótidos complementarios se hibridan entre sí y se ligan, de modo que se formen concatémeros. Los concatémeros bicatenarios se marcan, por ejemplo, radioactivamente mediante transcripción Nick. La hibridación se realiza, normalmente, en condiciones poco severas, usando altas concentraciones de oligonucleótidos.

55 6 x SSC

0,01 M fosfato de sodio

60 1 mM EDTA (pH 8)

0,5% SDS

100 microg/ml de ADN de esperma de salmón desnaturalizado

65 0,1% leche en polvo pobre en grasa

ES 2 331 228 T3

Durante la hibridación la temperatura se reduce gradualmente a 5 hasta 10°C por debajo de la temperatura del oligonucleótido calculada o la temperatura ambiente (RT = ~ 23°C en todos los experimentos 23, salvo indicación en contrario). A continuación, se realizan etapas de lavado y autoradiografía. El lavado se realiza bajo condiciones muy poco severas, por ejemplo, se realizan tres etapas de lavado, usando 4 x SSC. Otros detalles pueden desprenderse de Sambrook, J. I *et al.* (1989), "Molecular cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, o Ausubel, F.M., *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

El clon TC002034029R-11 con el nombre genético Te PCE11 es una elongasa de secuencia completa de *Thraustochytrium* y, por tanto, más largo que el clon TC002034029R de Seq. ID No. 3 y Seq. ID No. 4. Este clon se aísla mediante procedimientos de hibridación como los arriba descritos (Ejemplo 7). Se trata de una secuencia de ADN de una longitud de 1054 pares de bases que codifica para 271 aminoácidos con un codón de inicio (start) en la posición del par de bases 43-45 y un codón de stop en la posición del par de bases 856-858.

Ejemplo 8

Identificación de genes destino por clasificación de bancos de expresión con anticuerpos.

Se usan secuencias de cADN para la obtención de proteína recombinante, por ejemplo, en *E. coli* (por ejemplo Qiagen QIAexpress pQE-System). Las proteínas recombinantes se suelen purificar, entonces, por cromatografía de afinidad de Ni-NTA (Qiagen). Las proteínas recombinantes se usan, entonces, para la obtención de anticuerpos específicos, por ejemplo, aplicando técnicas estándar, para la inmunización de conejos.

Los anticuerpos se purifican, luego, por afinidad, usando una columna de Ni-NTA, presaturada con antígeno recombinante, en la forma descrita por Gu *et al.*, (1994) *BioTechniques* 17:257-262.

El anticuerpo puede usarse, entonces, para la revisión de bancos de expresión de cADN mediante clasificación inmunológica (Sambrook, J., *et al.* (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, o Ausubel, F. M., *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons).

Ejemplo 9

Plásmidos para la transformación de plantas

Para la transformación de plantas se pueden usar vectores binarios, tal como pBinAR (Hofgen y Willmitzer, *Plant Science* 66 (1990) 221-230). La construcción de los vectores binarios se puede alcanzar por ligación del cADN en orientación sentido o antisentido en T-ADN. 5' del cADN activa un promotor vegetal y la transcripción del cADN. Una secuencia de poliadenilación se encuentra 3' del cADN.

La expresión específica de tejido se puede alcanzar usando un promotor específico de tejido. Por ejemplo, se puede alcanzar una expresión específica de semilla, clonando el promotor de napina o de LeB4 o el promotor de USP 5' del cADN. También se puede usar cualquier otro elemento de promotor específico de semilla. Para la expresión constitutiva en la planta total, se puede usar el promotor de CaMV-35S.

La proteína expresada se puede dirigir en un compartimento celular, usando un péptido señal, por ejemplo, para pastidios, mitocondrias o el retículo endoplasmático (Kermode, *Crit. Rev. Plant Sci.* 15, 4 (1996) 285-423). El péptido señal se clona 5' en la trama de lectura con el cADN, para alcanzar la localización subcelular de la proteína de fusión.

Ejemplo 10

Transformación de Agrobacterium

La transformación de plantas inducida por *Agrobacterium* se puede realizar, por ejemplo, usando la cepa GV3101 (pMP90-) (Koncz y Schell, *Mol. Gen. Genet.* 204 (1986) 383-396) o LBA4404 (Clontech) de *Agrobacterium tumefaciens*. La transformación puede realizarse aplicando las técnicas de transformación estándar (Deblaere *et al.*, *Nucl. Acids. Res.* 12 (1984), 4777-4788).

Ejemplo 11

Transformación de plantas

La transformación de plantas inducida por *Agrobacterium* se puede realizar, aplicando las técnicas de transformación y regeneración estándar (Gelvin, Stanton B., Schilperoort, Robert A. *Plant Molecular Biology Manual*, 2ª edición, Dordrecht: Kluwer Academic Publ., 1995, en Sect., Ringbuc Centrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton: CRC Press, 1993, 360 págs., ISBN 0-8493-5164).

ES 2 331 228 T3

Por ejemplo, se puede transformar colza mediante transformación de cotiledóneas o hipocotiledóneas (Moloney *et al.*, Plant Cell Report 8 (1989) 238-242; De Block *et al.*, Plant Physiol. 91 (1989) 694-701). El uso de antibióticos para la selección de *Agrobacterium* y de las plantas viene determinado por el vector binario y la cepa de *Agrobacterium* usados. La selección de colza suele efectuarse con canamicina como marcador de planta seleccionable.

La transferencia genética inducida por *Agrobacterium* en lino (*Linum usitatissimum*) se puede efectuar usando, por ejemplo una técnica descrita por Mlynarova *et al.* (1994) Plant Cell Report 13:282-285. La soya puede ser transformada usando, por ejemplo, una técnica descrita en la EP-A-0 0424 047 (Pioneer Hi-Bred International) o en la EP-A-0 0397 681, la US 5,376,543, US 5,169,770 (University Toledo).

La transformación de plantas mediante bombardeo de partículas, la absorción de ADN inducida por polietilenglicol o mediante la técnica de fibra de carbonato de silicio se describen, por ejemplo en Freeling y Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

Ejemplo 12

Plásmidos para la transformación de plantas

Para la transformación de plantas se pueden usar vectores binarios, tal como pBinAR (Hofgen y Willmitzer, Plant Science 66 (1990) 221-230). Los vectores binarios se pueden construir por ligación de cADN en orientación sentido o antisentido en T-ADN. 5' del cADN activa un promotor vegetal y la transcripción del cADN.

Una secuencia de poliadenilación se encuentra 3' del cADN.

La expresión específica de tejido se puede alcanzar usando un promotor específico de tejido. Por ejemplo, se puede alcanzar una expresión específica de semilla, clonando el promotor de napina o de LeB4 o el promotor de USP 5' del cADN. También se puede usar cualquier otro elemento de promotor específico de semilla. Para la expresión constitutiva en la planta total, se puede usar el promotor de CaMV-35S.

Especialmente, se pueden clonar genes codificantes para elongasas y desaturasas, construyendo varios casetes de expresión en serie, en un vector binario, para replicar la vía metabólica en las plantas.

Dentro de un casete de expresión se puede dirigir la proteína a expresar mediante un péptido señal, por ejemplo para plástidos, mitocondrias o el retículo endoplasmático, en un compartimento celular (Kermode, Crit. Rev. Plant Sci. 15 1 4 (1996) 285-423). El péptido señal se clona 5' en la trama de lectura con el cADN, para alcanzar la localización subcelular de la proteína de fusión.

Ejemplo 13

Mutagénesis in vivo

La mutagénesis *in vivo* de microorganismos se puede realizar mediante pasaje del ADN del plásmido (u otro vector) por *E. coli* u otros microorganismos (por ejemplo *Bacillus spp.* o levaduras, como *Saccharomyces cerevisiae*), en los que está mermada la capacidad de mantener la integridad de su información genética. Cepas de mutantes usuales tienen mutaciones en los genes para el sistema de reparación de ADN (por ejemplo mutHLS, mutD, mutT etc.; como literatura, véase Rupp, W.D. (1996) ADN repair mechanisms, en: *Escherichia coli* and Salmonella, págs. 2277-2294, ASM: Washington). Estas cepas son conocidas a la persona técnica en la materia, El uso de estas cepas se ilustra, por ejemplo, en Greener, A., y Callahan, M. (1994) Strategies 7: 32-34. La transferencia de moléculas de ADN mutadas en plantas se realiza, preferentemente, tras selección y ensayo de los microorganismos. Plantas transgénicas son generadas según diferentes ejemplos, en la parte de los ejemplos del presente documento.

Ejemplo 14

Examen de la expresión de un producto genético recombinante en un organismo transformado

La actividad de un producto genético recombinante en un organismo huésped transformado se midió en el nivel de transcripción y/o translación.

Un procedimiento apropiado para determinar la cantidad de transcripción del gen (que es un indicio de la cantidad de ARN disponible para la translación del producto genético) es la realización de un Northern Blot como se describe más abajo (como referencia véase Ausubel *et al.* (1988) Current Protocols in Molecular Biology, Wiley: New York, o la parte de los ejemplos arriba mencionada), en los que se dota un primer, que está diseñado de tal manera, que se liga con gen de interés, de una marcación detectable (generalmente, radioactiva o quimiluminiscente), de manera que cuando el ARN total de un cultivo del organismo es extraído, separado sobre un gel, transferido sobre una matriz sólida e incubado con esta sonda, el enlace y la extensión del enlace de la sonda indica la presencia y también la cantidad

ES 2 331 228 T3

de mRNA para el gen. Esta información indica el grado de la transcripción del gen transformado. ARN celular total puede prepararse de células, tejidos u órganos con varios procedimientos, todos los cuales son conocidos en el campo técnico, como por ejemplo, los descritos por Bormann, E. R., *et al.* (1992) *Mol. Microbiol.* 6: 317-326.

5 *Hibridación Northern:*

Para la hibridación de ARN se separan 20 μ g de ARN total ó 1 μ g de poli (A) + -ARN mediante electroforesis de gel en geles de agarosa con una concentración de 1,25%, usando formaldehído en la forma descrita en Arnasio (1986, *Anal. Biochem.* 152, 304), se transfieren mediante atracción capilar, valiéndose de 10 x SSC de membranas de nilón cargadas en forma positiva (Hybond N+, Amersham, Braunschweig), se inmovilizan mediante luz ultravioleta y se prehibridan por 3 horas a 68°C, usando búfer de hibridación (10% de sulfato de dextrano peso/vol., 1 M de NaCl, 1% de SDS, 100 mg de ADN de esperma de arenque).

La marcación de la sonda de ADN con el Highprime ADN labeling- Kit (Roche, Mannheim, Alemania) se realizó durante la prehibridación, usando alpha-32P-dCTP (Amersham, Braunschweig, Alemania). La hibridación se realizó a 68°C durante la noche en el mismo búfer, adicionando la sonda de ADN marcada. Las etapas de lavado se realizaron dos veces por 15 min, usando 2 X SSC y dos veces por 30 min, usando 1 X SSC, 1% de SDS, a 68°C. La exposición de los filtros cerrados se realizó a -70°C durante un período de 1 a 14 días.

Para analizar la presencia y la cantidad relativa en proteína translatada por este mRNA se pueden aplicar técnicas estándar, como por ejemplo un Western-Blot (véase por ejemplo Ausubel *et al.* (1988) *Current Protocols in Molecular Biology*, Wiley: New York). En este procedimiento se extraen todas las proteínas celulares, se separan mediante electroforesis de gel, se transfieren sobre una matriz, como por ejemplo nitrocelulosa, y se incuban con una sonda, como por ejemplo un anticuerpo, que se liga específicamente con la proteína deseada. Esta sonda suele estar dotada de una marcación quimiluminiscente o colorimétrica, que es fácil de detectar.

La presencia y la cantidad de la marcación observada indican la presencia y la cantidad de la proteína mutada deseada, presente en la célula.

30 Ejemplo 15

Análisis del efecto de las proteínas recombinantes sobre la producción del producto deseado

El efecto de la modificación genética sobre plantas, hongos, algas, ciliatos o sobre la producción de un compuesto deseado (por ejemplo un ácido graso) puede ser determinado, cultivando los microorganismos modificados o la planta modificada bajo condiciones apropiadas (como las arriba mencionadas) y examinando el medio y/o los componentes celulares sobre la producción incrementada del producto deseado (a saber, de lípidos o de un ácido graso).

Estas técnicas de análisis son conocidas a la persona técnica en la materia y abarcan la espectroscopía, cromatografía de capa delgada, procesos de tintura de diferente índole, procedimientos enzimáticos y microbiológicos, así como cromatografía analítica, como cromatografía de líquido de alto rendimiento (véase, por ejemplo, Ullman, *Encyclopedia of Industrial Chemistry*, vol. A2, págs. 89 90 y págs. 44 613, VCH: Weinheim (1985); Fallon, A., *et al.*, (1987)

“Applications of HPLC in Biochemistry” en: *Laboratory Techniques in Biochemistry and Molecular Biology*, V. 17; Rehm *et al.* (1993) *Biotechnology*, V. 3, capítulo III: “Product recovery and purification”, pág. 469-714, VCH: Weinheim; Belter, P.A., *et al.* (1988) *Bioseparations: downstream processing for Biotechnology*, John Wiley and Sons; Kennedy, J.F., y Cabral, J.M.S. (1992) *Recovery processes for biological Materials*, John Wiley and Sons; Shaeiwitz, J.A., y Henry, J.D. (1988) *Biochemical Separations*, en: *Ullmann’s Encyclopedia of Industrial Chemistry*, V. B3; capítulo 11, S. 1-27,

VCH: Weinheim; y Dechow, F.J. (1989) *Separation and purification techniques in biotechnology*, Noyes Publications).

Además de los procedimientos arriba mencionados, se extraen lípidos vegetales a partir de material vegetal en la forma descrita por Cahoon *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96 (22):12935-12940, y Browse *et al.* (1986) *Analytic Biochemistry* 152:141-145. El análisis cuantitativo y cualitativo de lípidos y ácidos grasos está descrito en Christie, William W., *Advances in Lipid Methodology*, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., *Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press*, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); “Progress in Lipid Research”, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: *Progress in the Chemistry of Fats and Other Lipids* CODEN.

Adicionalmente a la medida del producto final de la fermentación, también se pueden analizar otros componentes de las vías metabólicas usadas para la producción del compuesto deseado, tales como productos intermedios y productos secundarios, para determinar la eficiencia total de la producción del compuesto. Los procedimientos de análisis comprenden medidas de las cantidades de sustancias nutritivas en el medio (por ejemplo azúcar, carbohidratos, fuentes de nitrógeno, fosfato y otros iones), medidas de la composición de la biomasa y del crecimiento, análisis de la pro-

ES 2 331 228 T3

ducción de metabolitos usuales de vías de biosíntesis y medidas de gases, que son generados durante la fermentación. Procedimientos estándar para estas medidas están descritos en Applied Microbial Physiology; A Practical Approach, P.M. Rhodes y P.F. Stanbury, Hrsgeb., IRL Press, S. 131-163 y 165-192 (ISBN: 0199635773) y la literatura allí citada.

Un ejemplo es el análisis de ácidos grasos (abreviaturas: FAME, metil éster de ácido graso; GC-MS, cromatografía de gas-líquido- espectroscopia de masas; TAG, triacilglicerina TLC, cromatografía de capa delgada).

La prueba inequívoca de la presencia de productos de ácido graso se puede obtener mediante análisis de organismos recombinantes según procedimientos de análisis estándar: GC, GC-MS o TLC, como se describe en varias ocasiones por Christie y en la literatura allí citada (1997, en: Advances on Lipid Methodology, cuarta edición: Christie, Oily Press, Dundee, 119-169; 1998, Métodos de cromatografía de gases y espectrometría de masas, lípidos 33:343-353).

El material a analizar se puede fraccionar por tratamiento con ultrasonido, molienda en un molino de vidrio, nitrógeno líquido y molienda o mediante otros procedimientos aplicables. El material debe centrifugarse después del fraccionamiento. El sedimento se resuspende en agua destilada, se calienta 10 min a 100°C, se enfría sobre hielo y se centrifuga otra vez, seguido de una extracción en 0,5 M de ácido sulfúrico en metanol con 2% de dimetoxipropano durante una hora a 90°C, obteniéndose compuestos de aceite y lípido hidrolizados, que resultan en lípidos transmetilados.

Estos metil ésteres de ácido graso se extraen en éter de petróleo y se someten, finalmente, a un análisis de GC, usando una columna capilar (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 microm, 0,32 mm) a un gradiente de temperatura de entre 170°C y 240°C durante 20 min y 5 min a 240°C. La identidad del metil éster de ácido graso obtenido debe definirse, aplicando los estándares que se pueden obtener por fuentes comerciales (a saber, Sigma).

Ácidos grasos, para los que no hay ningún estándar, deben ser identificados mediante derivatización y análisis de GC-MS subsiguiente. Por ejemplo, la localización de ácidos grasos con triples enlaces debe ser presentada mediante GC-MS después de la derivatización con derivados de 4,4-dimetoxiazolina (Christie, 1998, ver arriba).

Ejemplo 16

Estructuras de expresión en sistemas microbianos heterólogos

Cepas, condiciones de crecimiento y plásmidos

La cepa *Escherichia coli*: XL1 Blue MRF⁺ kan (Stratagene), se emplea para el subclonado de las nuevas elongasas, como PpPSE1 de *Physcomitrella patens*. Para la expresión funcional de este gen usamos la cepa de *Saccharomyces cerevisiae*, INVSc 1, (Invitrogen Co.).

E. coli se cultiva en caldo de Luria-Bertini (LB, Duchefa, Haarlem, Holanda) a 37°C. Si es necesario, se adiciona ampicilina (100 mg/litro), y se agrega 1,5% de agar (peso/vol.) para medios de LB sólidos. *S. cerevisiae* se cultiva a 30°C en medio de YPG o en medio mínimo completo sin uracilo (CMDum; ver en: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L. B., Coen, D.M., y Varki, A. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York), alternativamente con 2% (peso/vol.) de rafinosa o glucosa. Para medios sólidos se agregan 2% (peso/vol.) de agar de BactoTM (Difco). Los plásmidos usados para el clonado y la expresión son pUC18 (Pharmacia) y pYES2 (Invitrogen Co.).

Ejemplo 17

Clonado y expresión de elongasas específicas de PUFA a partir de Physcomitrella, Cryptocodinium y Thraustochytrium.

Los genes de longitud completa de las secuencias de la invención se pueden aislar en la forma descrita en el Ejemplo 7 y se pueden seguir tratando en la forma descrita a continuación. Como ejemplos se ilustran ejemplos de expresión concretos.

A) Elongación de ácidos grasos mediante elongasa de moho PpPSE1

Para la expresión en levadura, se modificó primero el clon de *P. patens*-cADN, PpPSE1, (nombre anterior de secuencia de banco de datos: 08ck19b07, nombre nuevo: pp001019019f), que codifica al gen de la elongasa específica de PUFA, (PSEI-), de tal manera que se obtuvieron un sitio de restricción BamHI y la secuencia de consenso de levadura para una translación altamente eficiente (Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292), junto con el codon de iniciación y un sitio de restricción BamHI, que flanqueaba al codon de iniciación.

ES 2 331 228 T3

Para la amplificación del marco de lectura abierto se sintetizó un par de primers, que era complementario a sus extremos 5' y 3'.

El gen de la secuencia de la invención representada en SEQ ID NO: 1 se clonó mediante reacción en cadena de polimerasa en pYES, y se obtuvo el plásmido pYppPSE1:

Los siguientes oligonucleótidos se usaron para el experimento de PCR:

Ppex6: ggatccacataatggaggtcgtggagagattc

Ppex6r: ggatcctcactcagtttagctcctttg

La reacción de PCR se realizó con el ADN del plásmido del clon PP001019019F como matriz en un Thermocycler (Biometra) con la Pfu-ADN-polimerasa (Stratagene) y el siguiente programa de temperatura: 3 min a 96°C, luego 25 ciclos con 30 s a 96°C, 30 s a 55°C y 1 min a 72°C, 1 ciclo con 10 min a 72°C.

El tamaño concreto del fragmento de ADN amplificado se confirmó mediante electroforesis de gel de agarosa-TBE. El ADN amplificado se extrajo del gel mediante el kit de extracción de gel QIAquick (QIAGEN) y se clonó usando el Sure Clone Ligation Kit (Pharmacia), primero en pUC18. El fragmento así clonado se cortó con BamHI, se liga en pYES, y se obtuvo pYppPSE1. La orientación del fragmento se examinó mediante HindIII. Después de la transformación de *E. coli* XLI Blue MRF' kan se preparó una minipreparación de ADN (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation. *BioTechniques* 4, 310-313) de transformados, y se identificaron clones positivos mediante análisis de restricción de BamHI. La secuencia del producto de PCR clonado se confirmó mediante resecuenciación, valiéndose del ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt).

Un clon se cultivó para la maxipreparación de ADN con el kit de extracción de ADN de plásmido, Nucleobondo® AX 500 (Macherey-Nagel, Düringen).

Saccharomyces INVSc1 fue transformado con pYppPSE1 o bien con pYES2 como control mediante un protocolo de PEG/acetato de litio modificado (Ausubel *et al.*, 1995). Después de la selección sobre placas de agar de CMdum con 2% de glucosa se seleccionaron transformantes y un transformante de pYES2 para el cultivo ulterior y la expresión funcional como ya se ha descrito, y se nutren con diferentes ácidos grasos en medio.

i) Muestras de lípidos de levaduras, que han sido transformadas con plásmido de pYES sin inserción de fragmento, o que expresan al gen de Pp-PSE1 (indicaciones en % molar) después de la nutrición con 250 micrones de ácido hexadecatriénico (16:3 Δ 7c,10c,13c).

TABLA 4

PYES2	PYES2	pYpp_PSE1	PYpp_PSE1
16: 0	11,8%	16: 0	11,1%
16: 1	28,7%	16: 1	23,9%
16: 3 Δ 7c,10c,13c	9,2%	16: 3 Δ 7c,10c,13c	12,0%
18: 0	10,6%	18: 0	8,6%
18: 1 Δ 9c	34,9%	18: 1 Δ 9c	20,6%
18: 1 Δ 11c	1,1%	18: 1 Δ 11c	1,4%
18: 3 Δ 9c,12c,15c	3,7%	18: 3 Δ 9c,12c,15c	21,4%

ii) Muestras de lípidos de levaduras, que han sido transformadas con plásmido de pYES sin inserción de fragmento o que expresan al gen de Pp-PSE1 (indicaciones en % molar) después de la nutrición con 500 microm de ácido pinolénico (18:3 Δ 6c,9c,12c).

ES 2 331 228 T3

TABLA 5

PYES2	PYES2	pYpp_PSE1	PYPp_PSE1
16: 0	18,3%	16: 0	16,9%
16: 1 Δ^{9c}	16,0%	16: 1 Δ^{69c}	15,3%
18: 0	8,6%	18: 0	8,4%
18: 1 Δ^{9c}	16,7%	18: 1 Δ^{9c}	17,5%
18: 1 Δ^{11c}	0,7%	18: 1 Δ^{11c}	2,0%
18: 3 $\Delta^{3c,9c,12c}$	39,8%	18: 3 $\Delta^{3c,9c,12c}$	32,6%
20: 3 $\Delta^{7c,11c,14c}$	0%	20: 3 $\Delta^{7c,11c,14c}$	5,1%

iii) Muestras de lípidos de levaduras, que han sido transformadas con el plásmido de pYES sin inserción de fragmento o que expresan al gen de Pp-PSE1 (indicaciones en % molar) después de la nutrición con 500 microm de ácido estearidónico (18:4 $\Delta^{6c,9c,12c,15c}$).

TABLA 6

PYES2	pYES2	pYpp_PSE1	PYPp_PSE1
16: 0	15,2%	16: 0	15,6%
16: 1 Δ^{9c}	13,1%	16: 1 Δ^{9c}	14,9%
18: 0	12,3%	18: 0	10,7%
18: 1 Δ^{9c}	12,9%	18: 1 Δ^{9c}	14,0%
18: 1 Δ^{11c}	0,7%	18: 1 Δ^{11c}	1,2%
18: 3 $\Delta^{6c,9c,12c,15c}$	45,4%	18: 3 $\Delta^{6c,9c,12c,15c}$	23,8%
20: 4 $\Delta^{8c,11c,14c,17c}$	0,4%	20: 4	19,8%

iv) Muestras de lípidos de levaduras, que han sido transformadas con plásmido de pYES sin inserción de fragmento o que expresan al gen de Pp-PSE1 Gen (indicaciones en % molar) después de la nutrición con 500 microm de ácido linólico (18:2 $\Delta^{9c,12c}$).

TABLA 7

pYES2	pYES2	pYpp_PSE1	PYPp_PSE1
16: 0	7,9%	16: 0	8,7%
16: 1 Δ^{9c}	1,2%	16: 1 Δ^{9c}	1,3%
18: 0	5,3%	18: 0	5,1%
18: 1 Δ^{9c}	1,3%	18: 1 Δ^{9c}	1,3%
18: 2 $\Delta^{9c,12c}$	83,9%	18: 2 $\Delta^{9c,12c}$	80,4%
20: 2 $\Delta^{11c,14c}$	0,5%	20: 2 $\Delta^{11c,14c}$	3,2%

B) Elongación de ácidos grasos mediante una elongasa a partir de *Thraustochytrium*

Para la expresión en levadura se modifica, primero, el clon de cADN de *Thraustochytrium* de SEQ ID NR: 3 (TcPSE2) que codifica un gen de elongasa - (PSE) específico de PUFA, de tal manera que represente un polipéptido funcionalmente activo. Para tal fin, se alarga el término N de la proteína a nivel de ADN por las pocas bases faltantes a partir de la elongasa de *Physcomitrella patens* en 42 pares de bases. Pero también es posible agregar solamente un codon de iniciación en la trama de lectura concordante con la secuencia.

ES 2 331 228 T3

Para el experimento de peR se usan los siguientes oligonucleótidos:

pTCPSE2-5' :

**aaaggatccacataatggaggtcgtggagagattctacggtgagttggatggga
agGTCATTTTCGGGCTCGACC**

pTCPSE2-3': aaggatccctgagtttagctccctttgcttcc

Los dos oligonucleótidos contienen adicionalmente un sitio de restricción de BamHI y la secuencia de consenso de levadura para una translación altamente eficiente (Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292).

La reacción de PCR se realiza con ADN de plásmido como matriz en un thermocycler (Biometra) con la polimerasa de Pfu-ADN- (Stratagene) y el siguiente programa de temperatura: 3 min a 96°C, seguido de 25 ciclos con 30 s a 96°C, 30 s a 55°C y 3min a 72°C, 1 ciclo con 10 min a 72°C y stop a 4°C.

El tamaño correcto del fragmento de ADN amplificado es confirmado mediante electroforesis de gel de agarosa-TBE.

El ADN amplificado se extrae a partir del gel con el kit de extracción de gel, QIAquick (QIAGEN), y se liga en el sitio de restricción de SmaI del vector desfosforilado de pUC18, usando el Sure Clone Ligation Kit (Pharmacia), y se obtiene pUC-híbrido-TcPSE2. Después de la transformación de *E. coli* XLI Blue MRF⁺ kan se realiza una minipreparación de ADN (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation BioTechniques 4, 310-313) en 24 transformantes resistentes a ampicilina y se identifican clones positivos mediante análisis de restricción de BamHI. La secuencia del producto de PCR clonado se confirma mediante resecuenciación usando el ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt).

El ADN de plásmido de pUC-PSEI, así como el híbrido de pUC-TcPSE2 son fraccionados conjuntamente con BamHI y los fragmentos obtenidos son ligados en el sitio de restricción de BamHI del Shuttlevector pYES2 de la levadura- *E. coli* desfosforilada, obteniéndose híbrido de pY2-Tc PSE2. Después de la transformación de *E. coli* y minipreparación de ADN a partir de los transformantes se examina la orientación del fragmento de ADN en el vector mediante escisión de HindIII.

Respectivamente se cultiva un clon para la maxipreparación de ADN con el kit de extracción de plásmido de ADN, Nucleobond[®] AX 500 plásmido, (Macherey-Nagel, Düringen).

Saccharomyces INVSc1 se transforma con pY2PSEI, pYES2, pY2-híbrido-TcPSE2 y pYES2 mediante un protocolo modificado de PEG/acetato de litio (Ausubel *et al.*, 1995). Después de la selección sobre placas de agar de CMdum con 2% de glucosa se seleccionan cada vez cuatro transformantes de pY2PSEI (pY2PSE1a-d), transformantes de pY2-híbrido-TcPSE2 (pY2-híbrido-TcPSE2 la-d) y un transformante de pYES2 para el cultivo ulterior y la expresión funcional.

Expresión funcional de una actividad de elongasa en levadura cultivo previo

20 ml de medio líquido de CMdum con 2% (peso/vol.) de rafinosa se inoculan con clones transgénicos de levadura (pY2-híbrido-TcPSE2 la-d, pYES2) y se cultivan 3 días a 30°C, 200 rpm, hasta alcanzar una densidad óptica a 600 nm (OD600) de 1,5-2.

Cultivo principal

Para la expresión se enriquecen 20 ml de medio líquido de CMdum con 2% de rafinosa y 1% (vol./vol.) de Tergitol NP-40 con sustratos de ácido graso hasta una concentración final de 0,003% (peso/vol.). Los medios se inoculan con los cultivos previos a una OD600 de 0,05. La expresión se induce a una OD600 de 0,2 con 2% (peso/vol.) de galactosa durante 16 horas, después de lo cual se cosechan los cultivos una OD600 de 0,8-1/2.

Análisis de ácido graso

Los ácidos grasos totales se extraen de cultivos de levadura y se analizan mediante cromatografía de gas. De estos se cosechan células de 5 ml de cultivo mediante centrifugación (1000 x g, 10 min, 4°C) y se lava una vez con 100 mM de NaHCO₃, pH 8,0, para eliminar los restos del medio y de los ácidos grasos. Para la obtención del metil éster de ácido graso (FAMES) se tratan los sedimentos celulares con 1 M de H₂SO₄ metanólico y 2% (vol./vol.) de dimetoxipropano durante una hora a 80 C. Los FAMES se extraen dos veces con 2 ml de éter de petróleo, se lavan una vez con 100 mM de NaHCO₃, pH 8,0, y una vez con agua destilada, y se secan con Na₂SO₄. El disolvente orgánico se evapora bajo corriente de argón, y los FAMES se disuelven en 50 microlitros de éter de petróleo. Las

ES 2 331 228 T3

muestras se separan con una columna capilar de ZEBRON-ZB-Wax (30 m, 0,32 mm, 0,25 micro m; Phenomenex) en un cromatógrafo de gas 6850 de Hewlett Packard con un detector de ionización de llamas. La temperatura del horno se programa en 70°C (mantener 1 min) hasta 200°C en incrementos de 20°C/min, luego a 250°C (mantener 5 min) en incrementos de 5°C/min, y, finalmente, a 260°C en incrementos de 5°C/min. Se usa nitrógeno como gas soporte (4,5 ml/min a 70°C). Los ácidos grasos se identifican por comparación con los tiempos de retención de estándares de FAME (SIGMA).

Las muestras de ácido graso de cinco cepas transgénicas de levadura en % molar se indican en la Tabla 1.

Las relaciones del ácido g-linolénico adicionado y absorbido se hacen resaltar mediante cifras en negrilla, las relaciones de los productos elongados por cifras rojas, y las relaciones del ácido g-linolénico elongados por cifras en negrilla (última línea).

El análisis de GC de FAMES de los lípidos totales de las levaduras transformadas con pYES2 (i/control) y pY2PSEI (ii-iv c+d/respectivamente transformadas con pY2PSEIA, pY2PSEIB, pY2PSEIC, pY2PSEIO) está representado en la Figura 2a-e. Para el análisis se cultivan las levaduras transgénicas en presencia de g-18:3. La Tab. 1 presenta sus muestras de ácidos graso en % molar. La absorción de g-18:3 está resaltada por cifras en negrilla, el producto de elongación, ácido dihomo-g-linolénico (20:308,11,14) está subrayado y la relación g-18: 3-producto de elongación (también en % molar) se hace resaltar por cifras en negrilla (última línea).

La estructura y los espectros de masa de derivado de DMOX de cis-D6, 9,12 C18: 3 están representados en la Fig. 3a+b. La estructura y los espectros de masa del derivado de OMOX de 08,11,14 C20:3 están representados en la Fig. 4a+b.

Los resultados indican, que g-18:3 se encuentra insertado en grandes cantidades en todas las levaduras transgénicas. Todos los cuatros clones transgénicos de levadura, que han sido transformados con pY2PSE1, muestran un pico (máximo) adicional en el cromatograma de gas, que se identificó por comparación de los tiempos de retención como 20: 3 DB,11,14. Se puede valer adicionalmente de una cromatografía de gas/espectroscopía de masa para confirmar esta identidad. La fracción de g-18:3 elongado asciende, como se indica en la Tabla 1, a 23, 7 hasta 40,5%. Además, no se encontró ninguna elongación adicional significativa de ácido palmítico (16:0), ácido palmitoleico (16: 1), ácido esteárico (18: 0) o ácido oleico(18:1 D9).

Los productos identificados muestran, que las secuencia de nucleótido de PpPSEI codifican una elongasa de ácido graso D6-selectiva a partir del moho *Physcomitrella patens*, que resulta en la formación de nuevos ácidos grasos en levaduras transgénicas.

Las relaciones de los sustratos de ácido graso adicionados y absorbidos pueden ser determinadas en la forma arriba descrita. De este modo se puede registrar la calidad y cantidad de la reacción de elongasa.

La estructura y los espectros de masa de los derivados de DMOX indican también la posición respectiva de un doble enlace.

Se pueden realizar otros experimentos de alimentación con otros y más diversos ácidos grasos (por ejemplo ácido araquidónico, ácido eicosapentaénico, etc.) para la confirmación detallada de la selectividad de sustrato de estas elongasas.

Ejemplo 18

50 Purificación del producto deseado a partir de los organismos transformados en general

El producto deseado puede ser generado a partir de material vegetal u hongos, algas, ciliatos, células animales o a partir del sobrenadante de los cultivos arriba descritos, aplicando diferentes procedimientos conocidos en el campo técnico. Cuando el producto deseado no es segregado a partir de células, se pueden cosechar las células a partir del cultivo mediante centrifugación lenta. Las células se pueden lisar mediante técnicas estándar, como por ejemplo fuerza mecánica o tratamiento con ultrasonido. Los órganos de plantas se pueden separar mecánicamente de otros tejidos u otros órganos. Después de la homogeneización se eliminan los detritos de las células por centrifugación. La fracción sobrenadante, que contiene las proteínas solubles se guarda para la purificación ulterior del compuesto deseado. Cuando el producto es segregado a partir de las células deseadas, entonces se eliminan las células del cultivo por centrifugación lenta, y se guarda la fracción sobrenadante para la purificación ulterior.

La fracción sobrenadante de cada procedimiento de purificación se somete a una cromatografía con una resina apropiada, reteniéndose la molécula deseada, o bien sobre una resina cromatográfica, mientras que las numerosas impurezas no son retenidas en la muestra, o las impurezas son retenidas en la resina, mientras que la muestra no es retenida en la resina. Estas etapas cromatográficas se pueden repetir, si es necesario, usando las mismas u otras resinas cromatográficas. La persona técnica en la materia está versada en la selección de resinas cromatográficas adecuadas y en su aplicación más efectiva para una determinada molécula a purificar.

ES 2 331 228 T3

El producto purificado puede ser concentrado por filtración o ultrafiltración y ser guardado a una temperatura a la que la estabilidad del producto es máxima.

5 En el campo técnico se conoce un amplio aspecto de procedimientos de purificación, y el procedimiento precedente no debe entenderse en forma limitativa. Estos procedimientos están descritos, por ejemplo, en Bailey, J.E., & Ollis, D.F., *Biochemical Engineering Fundamentals*, McGrawHill: New York (1986).

10 La identidad y pureza de los compuestos aislados se puede determinar mediante técnicas estándar en el campo técnico. Éstas incluyen la cromatografía de líquido de alto rendimiento (HPLC), procedimientos espectroscópicos, procedimientos de tintura, cromatografía de capa delgada, NIRS, tests enzimáticos o microbiológicos. Para una sinopsis de estos procedimientos de análisis, véase: Patek *et al.* (1994) *Appl. Environ. Microbiol.* 60:133-140; Malakhova *et al.* (1996) *Biotekhnologiya* 11:27-32; y Schmidt *et al.* (1998) *Bioprocess Engineer.* 19:67-70. *Ulmann's Enciclopedia of Industrial Chemistry* (1996) V. A27, VCH: Weinheim, pág. 89-90, pág. 521-540, pág. 540-547, pág. 559-566, 15 575-581 y pág. 581-587; Michal, G (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A., *et al.* (1987) *Applications of HPLC in Biochemistry en: Laboratory Techniques in Biochemistry and Molecular Biology*, Bd. 17.

Equivalentes

20 La persona técnica en la materia reconoce o puede determinar numerosos equivalentes de las modalidades específicas aquí descritas, aplicando solamente experimentos de rutina. Estos equivalentes deben estar comprendidos por las reivindicaciones de la patente.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

- 5 1. Ácido nucleico aislado que comprende una secuencia de nucleótido que codifica un polipéptido que alarga, por lo menos en dos átomos de carbono, ácidos grasos con 16, 18 ó 20 átomos de carbono; el ácido nucleico se selecciona del grupo compuesto por
- a) una de las secuencias de aminoácido representada en SEQ ID NO: 1, SEQ ID NO: 3 o SEQ ID NO: 9,
- 10 b) un ácido nucleico que, según el código genético degenerado, proviene de la secuencia representada en SEQ ID NO:1, SEQ ID NO: 3 o SEQ ID NO:9,
- c) derivados de la secuencia representada en SEQ ID NO:1, SEQ ID NO:3 o SEQ ID NO:9 que codifican para polipéptidos de las secuencias de aminoácidos representadas en SEQ ID NO:2, SEQ ID NO:4 o SEQ ID NO:10 y presentan al menos 50% de identidad a nivel de aminoácido, sin que se reduzca esencialmente la acción enzimática de los polipéptidos.
- 15 2. Ácido nucleico aislado según la reivindicación 1, donde la secuencia proviene de una planta vegetal, un alga o un hongo.
- 20 3. Secuencia aislada de ácido nucleico según la reivindicación 1 o reivindicación 2, donde la secuencia proviene de *Physcomitrella* o *Thraustochitrium*.
4. Constructo génico que comprende un ácido nucleico aislado según una de las reivindicaciones 1 hasta 3, donde el ácido nucleico está enlazado de manera funcional con una o más señales de regulación.
- 25 5. Constructo génico según la reivindicación 4, cuya expresión de gen se intensifica por las señales de regulación.
6. Constructo génico que contienen una secuencia de ácido nucleico según las reivindicaciones 1 hasta 3 o un constructo génico según la reivindicación 4 ó 5 y al menos otro ácido nucleico que codifica para un gen de la biosíntesis de ácido graso, seleccionado del grupo de $\Delta 19$ -, $\Delta 17$ -, $\Delta 15$ -, $\Delta 12$ -, $\Delta 9$ -, $\Delta 8$ -, $\Delta 6$ -, $\Delta 5$ -, $\Delta 4$ -desaturasa, las diversas hidroxilasas, la $\Delta 12$ -acetilasa, las acil-ACP-tioesterasas, β -cetoacil-ACP-sintasas o β -cetoacil-ACP-reductasas.
- 30 7. Secuencia de aminoácido que se codifica por la secuencia de aminoácido aislada según las reivindicaciones 1 hasta 3 o el constructo génico según una de las reivindicaciones 4 hasta 6.
8. Vector que comprende un ácido nucleico según una de las reivindicaciones 1 hasta 3 o un constructo génico según una de las reivindicaciones 4 hasta 6.
- 40 9. Organismo no humano transgénico que comprende al menos un ácido nucleico según una de las reivindicaciones 1 hasta 3, un constructo génico según una de las reivindicaciones 4 hasta 6 o un vector según la reivindicación 8.
10. Organismo transgénico según la reivindicación 9, donde el organismo es un microorganismo, un animal o una planta vegetal.
- 45 11. Organismo transgénico según la reivindicación 9 o 10, donde el organismo es una planta transgénica.
12. Anticuerpo que enlaza específicamente un polipéptido que se codifica por uno de los ácidos nucleicos de una de las reivindicaciones 1 hasta 3.
- 50 13. Molécula de ácido nucleico-antisentido que comprende la secuencia complementaria del ácido nucleico según una de las reivindicaciones 1 hasta 3.
- 55 14. Proceso para la obtención de PUFAs, donde el proceso comprende el cultivo de un organismo no humano que comprende un ácido nucleico según la reivindicación 1, un constructo génico según la reivindicación 5 o un vector según la reivindicación, que codifican un polipéptido que alarga, en al menos dos átomos de carbono, ácidos grasos con 16, 18 ó 20 átomos de carbono con al menos dos enlaces dobles, en condiciones en las que se forman PUFAs en el organismo.
- 60 15. Proceso según la reivindicación 14, donde los PUFAs preparados mediante el proceso son moléculas de ácido graso con 18, 20 o 22 átomos de carbono con al menos dos enlaces dobles en la molécula de ácido graso.
16. Proceso según la reivindicación 14 ó 15, donde las moléculas de ácido graso de 18, 20 ó 22 átomos de carbono se aíslan del organismo en forma de un aceite, lípido o de un ácido graso libre.
- 65 17. Proceso según una de las reivindicaciones 14 hasta 16, donde el organismo es un microorganismo, un animal o una planta vegetal.

ES 2 331 228 T3

18. Proceso según una de las reivindicaciones 14 hasta 16, donde el organismo es una planta transgénica.

19. Proceso según una de las reivindicaciones 14 hasta 18, donde el ácido graso con 16, 18 ó 20 átomos de carbono es un ácido graso con tres enlaces dobles en la molécula.

5

10

15

20

25

30

35

40

45

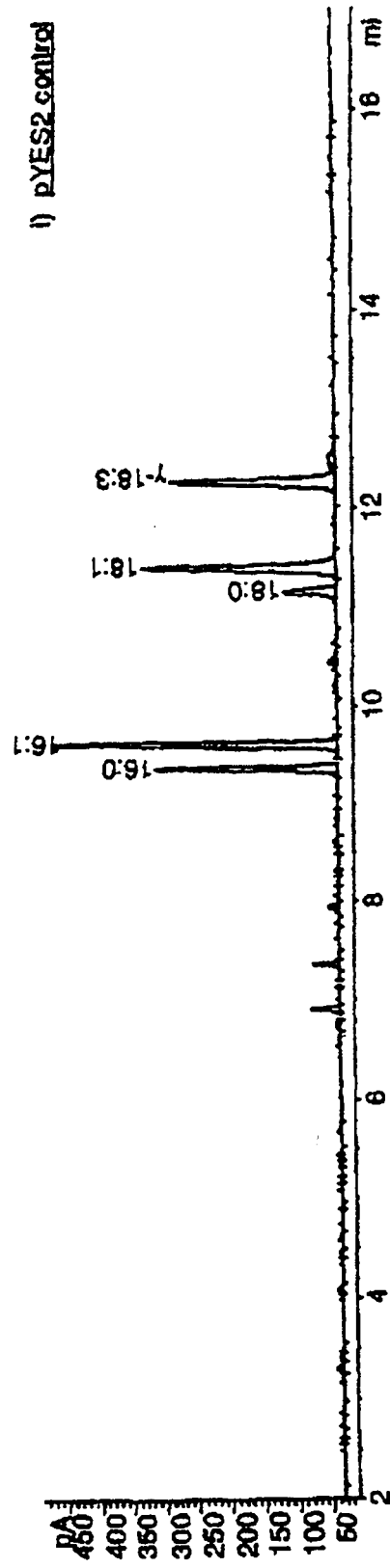
50

55

60

65

Figura 2a



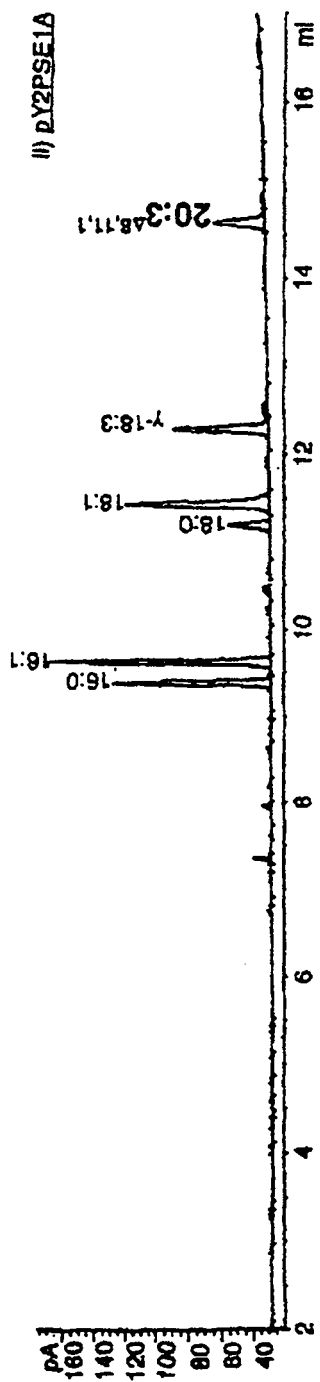


Figura 2b

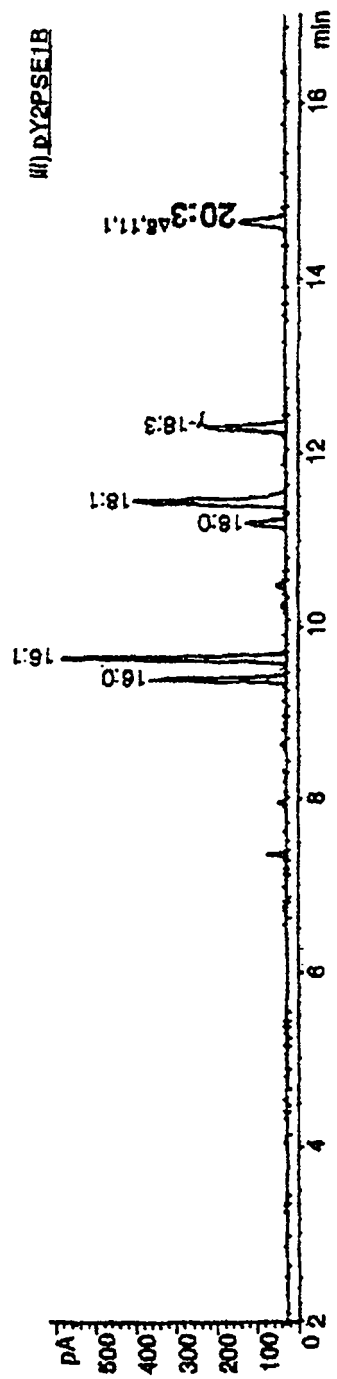
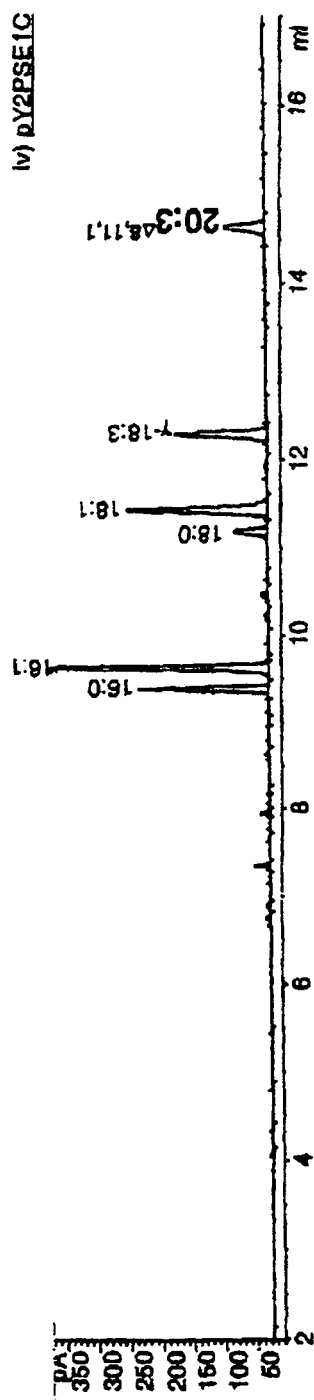


Figura 2c

Figure 2d



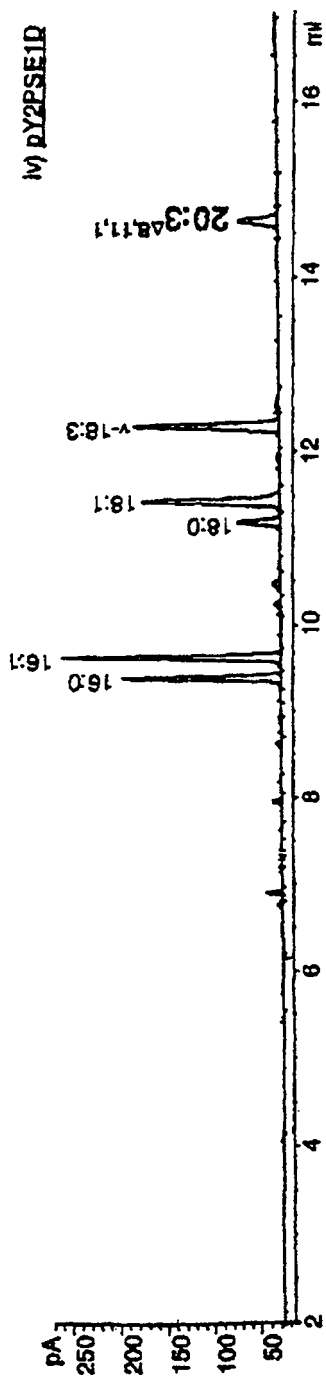


Figure 2e

Figura 3a: Derivado DMOX de cis- $\Delta^{6,9,12,18:3}$

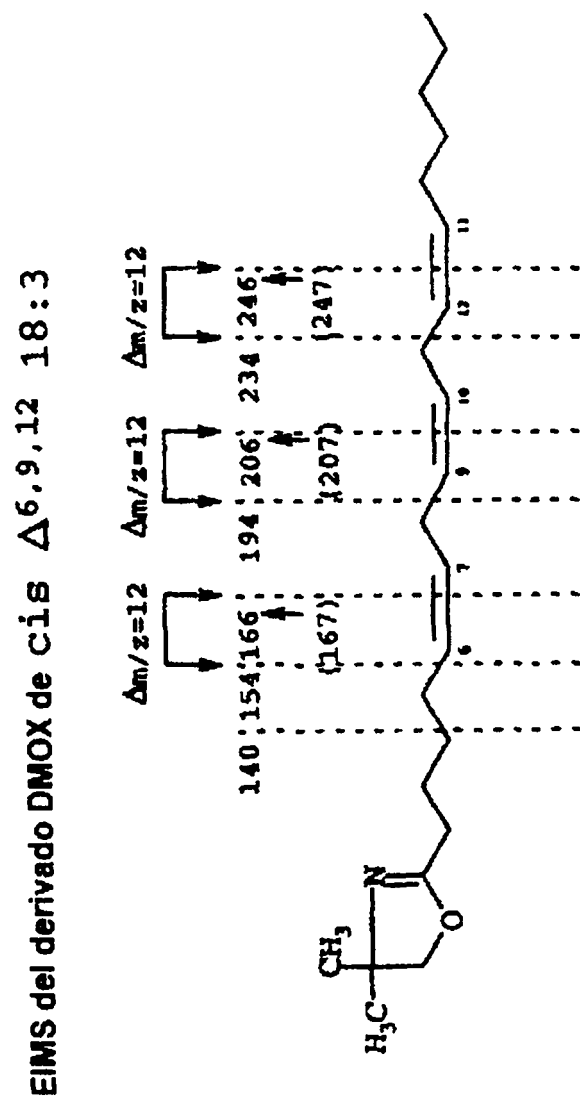


Figura 3b: Espectro de masas del derivado DMOX de cis $\Delta^{6,9,12}$ 18:3

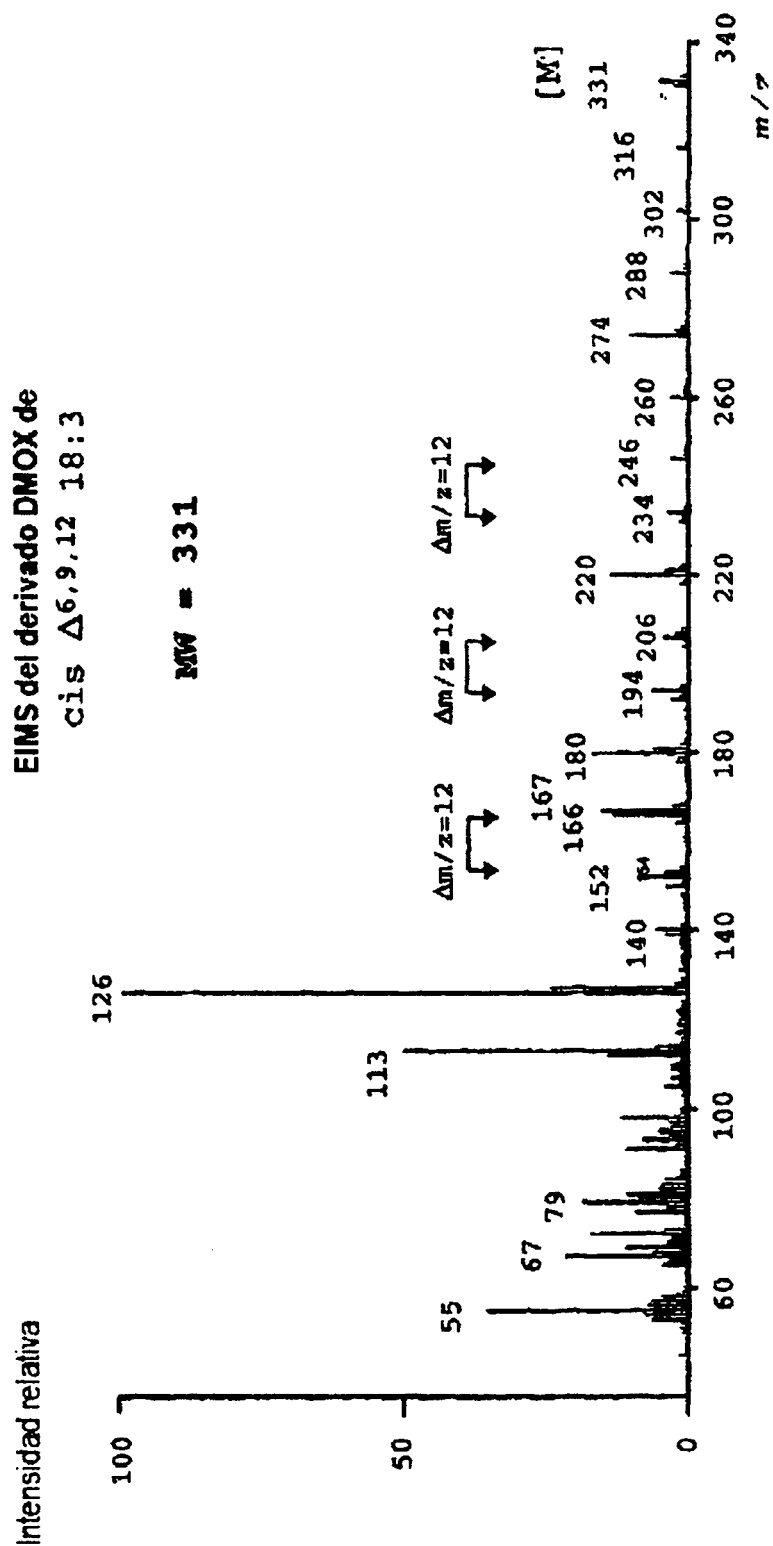
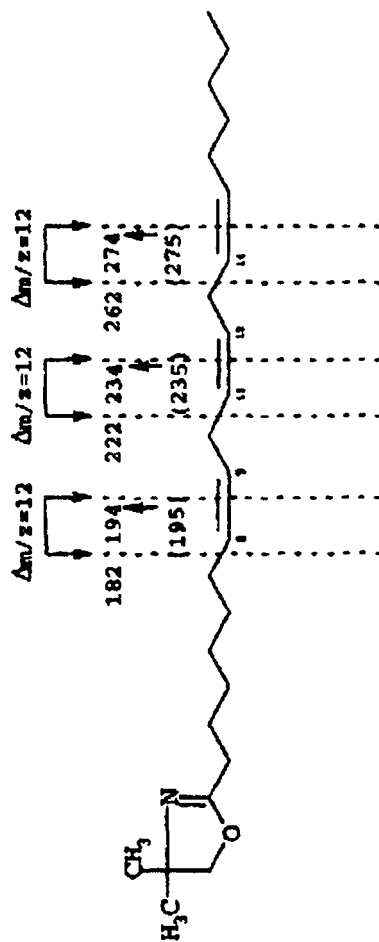


Fig. 4a: Derivado DMOX de *cis* $\Delta^{8,11,14}$ 20:3

El-MS del derivado de DMOX de

cis $\Delta^{8,11,14}$ 20:3



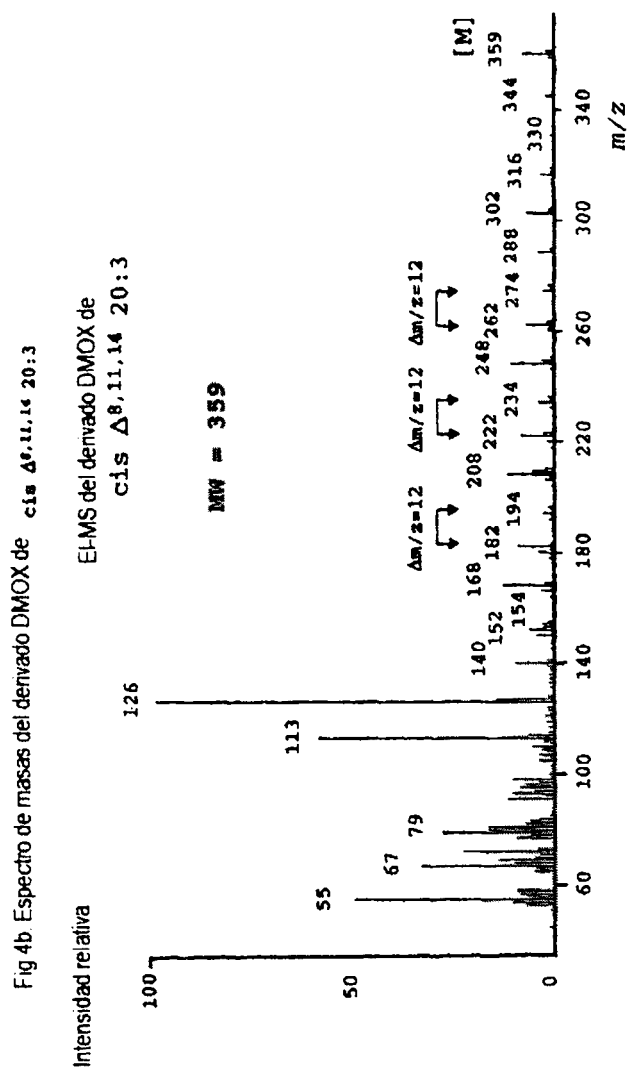


Fig. 5:

Pp_PSE1: KHKEMAILVYLFYMSKYVEFMDTVIMILKRSTRQISFLHVYHHSSISLIWVAIAHHAPGG
 KH ++ LF +SK E+ DTV++I+K + ++ FLHV HH++ W H
Tc_PSE1: KHPHFQLISLLFALSKIWEWFDTVLLIVKGN--KLRFHLVHHAHTF--FWLYAIDHIFLS

Pp_PSE1: EAYWSAALNSGVHVLMYAYYFLAACLRSSPKLKNKYLFWGRYLTQFQMFQFMLNL-----
 + A+N+ +H +MYA+YF R PK + TQ Q+ QF+ ++

Tc_PSE1: SIKYGVAVNAFIHTVMYAHYF-----RPFPGLRPLI-----TQLQIVQFIFSIGHITA

Pp_PSE1: VQAYYDMKTNAPYPQWLKILFYMISLLFLFGNFYVQKYI
 + +YD + W + +++ L LF NFY+Q+Y+

Tc_PSE1: IYWHYDCEPLVHTHFWEYVTPYLFVVPFLLFLNFYVQYV

Fig. 6:

Pp_PSE1: LNSGVHVLMYAYYFLAA
+N+VH +MYAYY A

Tc_PSE2: INASVHAIMYAYYAFTA

Fig. 7:

Pp_PSE1: AILVYLFYMSKYVEFMDTVIMILKRRSTRQISFLHVYHHSSISLIWW-AIA-HHAPGGEAY
 A+ + LF SK E MDTV +ILK +++ FL YHH+++ L W A+A + PG
 Cc_PSE1: ALALCLFCFSKIPELMDTVFLILKG--KKVRFLLQWYHHATVMLFCWLALATEYTPG----L
 Pp_PSE1: WSAALNSGVHVLMYAYYFL 203
 W AA N VH +MY Y+FL
 Cc_PSE1: WFAATNYFVHSIMMYFFL 339

ES 2 331 228 T3

Fig 8 Comparación de la secuencia polipeptídica PpPSE1 con Tc_PSE2

```

Pp_PSE1: 1 MEVVERFYGELDGKVSQGVNAL..LGSPGVELTDTPTTKGLPLVDSPTPIV 49
      | |.. | . | .      ||      . |
Tc_PSE2: 1 .....VISGLDLLPVLWETMKFDTAEVVSWWLRTHMVFYF 36

      . . . . .
50 LGVSVYLTIVIGLLWIKAR...DL.KPRASEPFLQALVLVENLFCFAL 95
      | :|| :: | :.. | || || |.      | . : . :
37 LMCFIYLVVIFGIQYIMEDRKEFDLRKPLAA....WSAFLAIFSIGASIR 82

      . . . . .
96 SLYMCVGIAYQAITWRYSIMGNAYNP..KIKEMAILVYLFYMSKYVEFMD 143
      . . . . | : | : | | . |      : | || | . |
83 TVPVLLKMLYEKGT.HHVLCGDTAMDWIDNPAGVWTEAFIPSEIPELID 131

      . . . . .
144 TVIMILKTRSTRQISFLVYHSSISLI.WMAIAHAPGGEAYWSAALNSG 192
      | . :|:: | : || || | . : | | | | | : ||:|.
132 TLFIVLRK..RKLITLHWYHVTVLLFCWHAWATPALTGIVF..AINDAS 177

      . . . . .
193 VHVLMYAYYFLAACLRSSPKLKNKYLFWGRYLTQPFQPFMLALVQAY.. 240
      || :|||| | | | | | | :| | . | .. :
178 VHAIKYAYYAPTA.LGYRP...TSYAI...YITLIQIMQMVGTAVTFYI 220

      . . . . .
241 .YDMKINAFYP.....QW..LIK.....ILFYMI...S 263
      ||| | | | | | | | | | | | :| .. |
221 GYDMAFVTPQPFRLDMKLNWDPLSRGENTEPTCKGANSSNAIFGVIMYAS 270

      . . . . .
264 LLFLFGNIFYVQKYIKPSDGRQKQKATE. 290
      |:|| | : |:| | |
271 YLYLFCLEFFYDYLKPKKSTPAACKTN. 297
    
```

| aminoácido idéntico
 /: equivalente químico

Fig. 9. Comparación de la secuencia polipeptídica CC_PSE1 con Tc_PSE2

```

      .       .       .       .
6  MTEKRG LQFTICG STGELVQNLQDGPTALALCLFCFSKIP ELMDTVFLIL 55
   |  .:|   :|| |   . | |   | ||||| |. ||. | : : |
90 MLYEKGTHHVLCGDTRN. .DWVIDNPAGVWTFMFI FFSKIP ELDITLFLIVL 137

      .       .       .       .
56 KGKQVRFLQWYHHATVMLPCWLALATEYTPGLWFAATNYFVHSIMYNYFF 105
   : :|. | ||| | | : ||| | | | | | : ||| | | :
138 RKRKLITLHWYHHVTVLLFCWHAWATFALTGIVFAADNASVHAIMYAYY. 186

      .       .       .       .
106 LMTFKTAARVVKPIAPLITIIQIAQMVGGLIVN. ....GIAITT...F.. 145
     |           | ||: ||| ||| | |   . | | |
187 ..AFTALGYRPTS YAIYITLIQIMQMVVGTAVIFPYIGYDMAFVTPQPFRL 234

      .       .       .       .
146 .....FTTGACQ. IQSVTVYSATVMYASYFYLFSQLFLEAY 180
     | |. |   : : |||| | | | | | |
235 DMKLNWDPLSKGENTEPTCKGANSSNAIFGVDMYASYLYLFCLFFYMAY 283
    
```

| aminoácido idéntico
 / equivalente químico

Fig. 10

	Motivo 1	Motivo 2	Motivo 3	Motivo 4
Pp_PSE1		FLHVVYHH	LMYAYYF	FGNFFVQ
Tc_PSE1		FLHVLHH	LMYAHYF	FLNFFLQ
Tc_PSE2	FSKIPEL	TLHWYHH	IMYAYFF	
Cc_PSE1	FSKIPEL	FLQWYHH	IMYMYFF	

Consenso:	FSKIPEL	FLHWYHH	LMYAYYF	FGNFFVQ
Variación		T QVL	I MHP	L L

ES 2 331 228 T3

LISTA DE SECUENCIAS

<110> BASF Aktiengesellschaft
<120> New Elongase gene and a process for the production of polyunsaturated fatty acids
5 <130> 0050/51159
<140> 20000081 <141> 2000-02-09
<160> 12
10 <170> PatentIn Vers. 2.0
<210> 1
<211> 1192
<212> ADN
15 <213> *Phiscomitrella patens*
<220>
<221> CDS
20 <222> (58)..(930)

<400> 1

```
25      ctgcttcgtc tcaccttggg ggtgtgattc gggagtgggt tgagttggtg gacgcga      57

      atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg      105
Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
30      1          5          10          15

      cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat      153
Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
35      20          25          30

      acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc atc      201
Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
40      35          40          45

      gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg      249
Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
45      50          55          60

      tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg      297
Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
50      65          70          75          80
```

55

60

65

ES 2 331 228 T3

5 ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc agt 345
 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95

10 ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac 393
 Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
 100 105 110

15 tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg att 441
 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
 115 120 125

20 ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat acc 489
 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
 130 135 140

25 gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc cac 537
 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
 145 150 155 160

30 gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct cat 585
 Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
 165 170 175

35 cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca gga 633
 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
 180 185 190

40 gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt cga 681
 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
 195 200 205

45 agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac ttg 729
 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
 210 215 220

50 aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac 777
 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
 225 230 235 240

55 tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att 825
 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
 245 250 255

60 ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac 873
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
 260 265 270

65

ES 2 331 228 T3

gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa 921
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
 275 280 285

5

act gag tga gctgtatcaa gccatagaaa ctctattatg ttagaacctg 970
 Thr Glu
 290

10

aagttggtgc tttcttatct ccacttatct ttttaagcagc atcagttttg aatgatgtg 1030

15

tgggcgtggt ctgcaagtag tcatcaatat aatcggcctg agcacttcag atggattggt 1090

20

agaacatgag taaaagcggg tattacggtg tttattttgt accaaatcac cgcacgggtg 1150

aattgaaata tttcagattt gatcaatttc atctgaaaaa aa 1192

<210> 2

25 <211> 290

<212> PRT

<213> *Phiscomitrella patens*

30 <400> 2

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
 1 5 10 15

35

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
 20 25 30

40

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
 35 40 45

45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
 50 55 60

50

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
 65 70 75 80

55

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95

60

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
 100 105 110

65

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
 115 120 125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr

ES 2 331 228 T3

	130		135		140	
5	Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His					
	145		150		155	160
10	Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His					
		165		170		175
15	His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly					
		180		185		190
20	Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg					
		195		200		205
25	Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu					
		210		215		220
30	Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr					
			230		235	240
35	Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile					
			245		250	255
40	Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr					
		260		265		270
45	Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys					
		275		280		285
50	Thr Glu					
	290					
55	<210> 3					
	<211> 687					
	<212> ADN					
	<213> Thraustochitrium					
	<220>					
	<221> CDS					
	<222> (1).. (588)					
	<400> 3					
60	cgc agc gtg cat aac ctc ggg ctc tgc ctc ttc tcg ggc gcc gtg tgg					48
	Arg Ser Val His Asn Leu Gly Leu Cys Leu Phe Ser Gly Ala Val Trp					
	1	5		10		15
65	atc tac acg agc tac ctc atg atc cag gat ggg cac ttt cgc agc ctc					96

ES 2 331 228 T3

Ile Tyr Thr Ser Tyr Leu Met Ile Gln Asp Gly His Phe Arg Ser Leu
20 25 30

5 gag gcg gca acg tgc gag ccg ctc aag cat ccg cac ttc cag ctc atc 144
Glu Ala Ala Thr Cys Glu Pro Leu Lys His Pro His Phe Gln Leu Ile
35 40 45

10 agc ttg ctc ttt gcg ctg tcc aag atc tgg gag tgg ttc gac acg gtg 192
Ser Leu Leu Phe Ala Leu Ser Lys Ile Trp Glu Trp Phe Asp Thr Val
50 55 60

15 ctc ctc atc gtc aag ggc aac aag ctc cgc ttc ctg cac gtc ttg cac 240
Leu Leu Ile Val Lys Gly Asn Lys Leu Arg Phe Leu His Val Leu His
65 70 75 80

20 cac gcc acg acc ttt tgg ctc tac gcc atc gac cac atc ttt ctc tcg 288
His Ala Thr Thr Phe Trp Leu Tyr Ala Ile Asp His Ile Phe Leu Ser
85 90 95

25 tcc atc aag tac ggc gtc gcg gtc aat gct ttc atc cac acc gtc atg 336
Ser Ile Lys Tyr Gly Val Ala Val Asn Ala Phe Ile His Thr Val Met
100 105 110

30 tac gcg cac tac ttc cgc cca ttc ccg aag ggc ttg cgc ccg ctt att 384
Tyr Ala His Tyr Phe Arg Pro Phe Pro Lys Gly Leu Arg Pro Leu Ile
115 120 125

35 acg cag ttg cag atc gtc cag ttc atc ttc agc atc ggc atc cat acc 432
Thr Gln Leu Gln Ile Val Gln Phe Ile Phe Ser Ile Gly Ile His Thr
130 135 140

40 gcc atc tac tgg cac tac gac tgc gag ccg ctc gtg cat acc cac ttt 480
Ala Ile Tyr Trp His Tyr Asp Cys Glu Pro Leu Val His Thr His Phe
145 150 155 160

45 tgg gaa tac gtc acg ccc tac ctc ttc gtc gtg ccc ttc ctc atc ctc 528
Trp Glu Tyr Val Thr Pro Tyr Leu Phe Val Val Pro Phe Leu Ile Leu
165 170 175

50 ttt ctc aat ttc tac ctg cag cag tac gtc ctc gcg ccc gca aaa acc 576
Phe Leu Asn Phe Tyr Leu Gln Gln Tyr Val Leu Ala Pro Ala Lys Thr
180 185 190

55 aag aag gca tag ccacgtaaca gtagaccagc agcgccgagg acgcgtgccg 628
Lys Lys Ala
195

60 cgttatcgcg aagcacgaaa taaagaagat catttgattc aaaaaaaaaa aaaaaaaaaa 687

65

ES 2 331 228 T3

<211> 195

<212> PRT

<213> Thraustochitrium

5

<400> 4

10	Arg	Ser	Val	His	Asn	Leu	Gly	Leu	Cys	Leu	Phe	Ser	Gly	Ala	Val	Trp
	1				5					10					15	
15	Ile	Tyr	Thr	Ser	Tyr	Leu	Met	Ile	Gln	Asp	Gly	His	Phe	Arg	Ser	Leu
				20					25					30		
20	Glu	Ala	Ala	Thr	Cys	Glu	Pro	Leu	Lys	His	Pro	His	Phe	Gln	Leu	Ile
			35					40					45			
25	Ser	Leu	Leu	Phe	Ala	Leu	Ser	Lys	Ile	Trp	Glu	Trp	Phe	Asp	Thr	Val
	50						55					60				
30	Leu	Leu	Ile	Val	Lys	Gly	Asn	Lys	Leu	Arg	Phe	Leu	His	Val	Leu	His
	65					70					75					80
35	His	Ala	Thr	Thr	Phe	Trp	Leu	Tyr	Ala	Ile	Asp	His	Ile	Phe	Leu	Ser
					85					90					95	
40	Ser	Ile	Lys	Tyr	Gly	Val	Ala	Val	Asn	Ala	Phe	Ile	His	Thr	Val	Met
				100					105					110		
45	Tyr	Ala	His	Tyr	Phe	Arg	Pro	Phe	Pro	Lys	Gly	Leu	Arg	Pro	Leu	Ile
		115						120					125			
50	Thr	Gln	Leu	Gln	Ile	Val	Gln	Phe	Ile	Phe	Ser	Ile	Gly	Ile	His	Thr
		130					135					140				
55	Ala	Ile	Tyr	Trp	His	Tyr	Asp	Cys	Glu	Pro	Leu	Val	His	Thr	His	Phe
	145					150					155					160
60	Trp	Glu	Tyr	Val	Thr	Pro	Tyr	Leu	Phe	Val	Val	Pro	Phe	Leu	Ile	Leu
					165					170					175	
65	Phe	Leu	Asn	Phe	Tyr	Leu	Gln	Gln	Tyr	Val	Leu	Ala	Pro	Ala	Lys	Thr
				180					185					190		
70	Lys	Lys	Ala													
			195													

ES 2 331 228 T3

<210> 5
 <211> 955
 <212> ADN
 5 <213> *Thaustochytrium*
 <220>
 <221> CDS
 10 <222> (1)..(894)
 <400> 5

gtc att tcg ggc ctc gac ctt ctc ccc gtg ctc tcg tgg gag act atg 48
 Val Ile Ser Gly Leu Asp Leu Leu Pro Val Leu Ser Trp Glu Thr Met
 1 5 10 15
 aag ttc gac act gcc gaa gtt gtc tcg gtc tgg ctg cgc acg cac atg 96
 Lys Phe Asp Thr Ala Glu Val Val Ser Val Trp Leu Arg Thr His Met
 20 25 30
 tgg gtc ccc ttc ctg atg tgc ttc atc tac ctg gtc gtc atc ttc ggc 144
 Trp Val Pro Phe Leu Met Cys Phe Ile Tyr Leu Val Val Ile Phe Gly
 35 40 45
 atc cag tac tac atg gag gac cgc aag gag ttc gat ctg cgc aag ccg 192
 Ile Gln Tyr Tyr Met Glu Asp Arg Lys Glu Phe Asp Leu Arg Lys Pro
 50 55 60
 ctg gcc gcc tgg agc gcc ttc ttg gcc att ttc agc atc ggc gcc tcc 240
 Leu Ala Ala Trp Ser Ala Phe Leu Ala Ile Phe Ser Ile Gly Ala Ser
 65 70 75 80
 atc cgc acc gtg ccc gtc ctg ctc aag atg ctc tac gaa aag ggc acg 288
 Ile Arg Thr Val Pro Val Leu Leu Lys Met Leu Tyr Glu Lys Gly Thr
 85 90 95
 cac cac gtg ctc tgc ggc gac acg cgc aac gac tgg gtc att gac aac 336
 His His Val Leu Cys Gly Asp Thr Arg Asn Asp Trp Val Ile Asp Asn
 100 105 110
 ccg gcc ggc gtc tgg acc atg gcc ttt atc ttt tcc aag att ccc gag 384
 Pro Ala Gly Val Trp Thr Met Ala Phe Ile Phe Ser Lys Ile Pro Glu
 115 120 125
 ctc atc gac acc ctc ttt atc gtg ctc cgc aag cgc aag ctc atc acc 432
 Leu Ile Asp Thr Leu Phe Ile Val Leu Arg Lys Arg Lys Leu Ile Thr
 130 135 140
 ctc cac tgg tac cac cac gtg acc gtg ctc ctg ttc tgc tgg cac gcc 480
 Leu His Trp Tyr His His Val Thr Val Leu Leu Phe Cys Trp His Ala

ES 2 331 228 T3

	145		150		155		160	
5	tgg gcc acc ttt gcg ctc acc ggc atc gtc ttt gcc gcc atc aac gcc	528						
	Trp Ala Thr Phe Ala Leu Thr Gly Ile Val Phe Ala Ala Ile Asn Ala							
	165	170	175					
10	tcg gtg cac gcc atc atg tac gcc tat tac gcc ttc acg gcc ctc ggc	576						
	Ser Val His Ala Ile Met Tyr Ala Tyr Tyr Ala Phe Thr Ala Leu Gly							
	180	185	190					
15	tac cga cca acc tcg tac gcc atc tac att acg ctc att cag atc atg	624						
	Tyr Arg Pro Thr Ser Tyr Ala Ile Tyr Ile Thr Leu Ile Gln Ile Met							
	195	200	205					
20	cag atg gtc gtc ggc acc gcc gtc acc ttt tac att ggc tac gac atg	672						
	Gln Met Val Val Gly Thr Ala Val Thr Phe Tyr Ile Gly Tyr Asp Met							
25	210	215	220					
30	gcc ttt gtc acg ccg cag ccc ttc cgc ctt gac atg aaa ctc aac tgg	720						
	Ala Phe Val Thr Pro Gln Pro Phe Arg Leu Asp Met Lys Leu Asn Trp							
	225	230	235	240				
35	gac ccg ctc agc aag ggc gag aac acc gag ccc acc tgc aag ggc gcc	768						
	Asp Pro Leu Ser Lys Gly Glu Asn Thr Glu Pro Thr Cys Lys Gly Ala							
	245	250	255					
40	aac tcc tcc aac gcc atc ttc ggc gtc atc atg tac gcc tcg tac ctc	816						
	Asn Ser Ser Asn Ala Ile Phe Gly Val Ile Met Tyr Ala Ser Tyr Leu							
	260	265	270					
45	tac ctc ttc tgc ctc ttc ttc tac atg gcc tac ctg cgc ccg aag aag	864						
	Tyr Leu Phe Cys Leu Phe Phe Tyr Met Ala Tyr Leu Arg Pro Lys Lys							
	275	280	285					
50	tcg acg ccc gcg gcc aag aag aca aac taa tcgcacacta ccaaacaatc	914						
	Ser Thr Pro Ala Ala Lys Lys Thr Asn							
	290	295						
55	ttccactcga cctagaaaaa aaaaaaaaaa aaaaactcga g	955						

60 <210> 6
 <211> 297
 <212> PRT
 <213> *Thaustochytrium*

65

ES 2 331 228 T3

<400> 6

5 Val Ile Ser Gly Leu Asp Leu Leu Pro Val Leu Ser Trp Glu Thr Met
 1 5 10 15
 Lys Phe Asp Thr Ala Glu Val Val Ser Val Trp Leu Arg Thr His Met
 20 25 30
 10 Trp Val Pro Phe Leu Met Cys Phe Ile Tyr Leu Val Val Ile Phe Gly
 35 40 45
 15 Ile Gln Tyr Tyr Met Glu Asp Arg Lys Glu Phe Asp Leu Arg Lys Pro
 50 55 60
 20 Leu Ala Ala Trp Ser Ala Phe Leu Ala Ile Phe Ser Ile Gly Ala Ser
 65 70 75 80
 25 Ile Arg Thr Val Pro Val Leu Leu Lys Met Leu Tyr Glu Lys Gly Thr
 85 90 95
 30 His His Val Leu Cys Gly Asp Thr Arg Asn Asp Trp Val Ile Asp Asn
 100 105 110
 35 Pro Ala Gly Val Trp Thr Met Ala Phe Ile Phe Ser Lys Ile Pro Glu
 115 120 125
 40 Leu Ile Asp Thr Leu Phe Ile Val Leu Arg Lys Arg Lys Leu Ile Thr
 130 135 140
 45 Leu His Trp Tyr His His Val Thr Val Leu Leu Phe Cys Trp His Ala
 145 150 155 160
 50 Trp Ala Thr Phe Ala Leu Thr Gly Ile Val Phe Ala Ala Ile Asn Ala
 165 170 175
 55 Ser Val His Ala Ile Met Tyr Ala Tyr Tyr Ala Phe Thr Ala Leu Gly
 180 185 190
 60 Tyr Arg Pro Thr Ser Tyr Ala Ile Tyr Ile Thr Leu Ile Gln Ile Met
 195 200 205
 65 Gln Met Val Val Gly Thr Ala Val Thr Phe Tyr Ile Gly Tyr Asp Met
 210 215 220
 Ala Phe Val Thr Pro Gln Pro Phe Arg Leu Asp Met Lys Leu Asn Trp
 225 230 235 240
 Asp Pro Leu Ser Lys Gly Glu Asn Thr Glu Pro Thr Cys Lys Gly Ala
 245 250 255
 Asn Ser Ser Asn Ala Ile Phe Gly Val Ile Met Tyr Ala Ser Tyr Leu

ES 2 331 228 T3

	260	265	270	
5	Tyr Leu Phe Cys Leu Phe Phe Tyr Met Ala Tyr Leu Arg Pro Lys Lys			
	275	280	285	
10	Ser Thr Pro Ala Ala Lys Lys Thr Asn			
	290	295		
	<210> 7			
	<211> 708			
15	<212> ADN			
	<213> Crypthecodinium			
	<220>			
	<221> CDS			
20	<222> (1)..(645)			
	<400> 7			
25	cgg cac gag gta cac atg acc gag aag agg gga ctg cag ttc acg atc			48
	Arg His Glu Val His Met Thr Glu Lys Arg Gly Leu Gln Phe Thr Ile			
	1	5	10	15
30	tgc ggc tct act ggt gag ttg gtg cag aat ctc cag gat ggt ccc act			96
	Cys Gly Ser Thr Gly Glu Leu Val Gln Asn Leu Gln Asp Gly Pro Thr			
	20	25	30	
35	gcc ttg gcg ttg tgc ctc ttt tgc ttc agc aaa att ccc gag ttg atg			144
	Ala Leu Ala Leu Cys Leu Phe Cys Phe Ser Lys Ile Pro Glu Leu Met			
40	35	40	45	
45	gac acg gtc ttt ctc atc ttg aag ggc aag aag gtt cgc ttt ttg cag			192
	Asp Thr Val Phe Leu Ile Leu Lys Gly Lys Lys Val Arg Phe Leu Gln			
	50	55	60	
50	tgg tac cac cac gct acc gtg atg ctc ttc tgc tgg ttg gca ctg gct			240
	Trp Tyr His His Ala Thr Val Met Leu Phe Cys Trp Leu Ala Leu Ala			
	65	70	75	80
55	acg gag tac acc ccg ggc ctc tgg ttc gcg gcc act aac tac ttc gtg			288
	Thr Glu Tyr Thr Pro Gly Leu Trp Phe Ala Ala Thr Asn Tyr Phe Val			
	85	90	95	
60	cac tcc atc atg tac atg tac ttc ttc ttg atg acc ttc aag acg gcc			336
	His Ser Ile Met Tyr Met Tyr Phe Phe Leu Met Thr Phe Lys Thr Ala			
	100	105	110	
65	gca aag gtc gtg aag ccc att gcc cct ctc atc acc atc atc cag atc			384

ES 2 331 228 T3

Ala Lys Val Val Lys Pro Ile Ala Pro Leu Ile Thr Ile Ile Gln Ile
 115 120 125

5 gcc cag atg gtc tgg ggt ctc atc gtc aac ggc atc gcg atc acc act 432
 Ala Gln Met Val Trp Gly Leu Ile Val Asn Gly Ile Ala Ile Thr Thr
 10 130 135 140

ttc ttc acc acg ggc gcc tgc cag atc cag tcc gtg acg gtc tac tcg 480
 Phe Phe Thr Thr Gly Ala Cys Gln Ile Gln Ser Val Thr Val Tyr Ser
 15 145 150 155 160

gcc att gtg atg tac gct tcg tac ttc tac ctc ttc tcc cag ctc ttc 528
 Ala Ile Val Met Tyr Ala Ser Tyr Phe Tyr Leu Phe Ser Gln Leu Phe
 20 165 170 175

ctg gag gca tac gga tcc gct ggc aag aac aag aag aag ctc gcc cgc 576
 Leu Glu Ala Tyr Gly Ser Ala Gly Lys Asn Lys Lys Lys Leu Ala Arg
 25 180 185 190

gag ctc tcc cga aag atc tcc gag gct ctc ctg aat agt ggc gac gag 624
 Glu Leu Ser Arg Lys Ile Ser Glu Ala Leu Leu Asn Ser Gly Asp Glu
 30 195 200 205

gta gcc aag cac ctc aag tga actgagcgc ctcacatcttgg tctgggtccgc 675
 Val Ala Lys His Leu Lys
 35 210 215

40 caaattgccg cgtgcatgtg catgagatgc tgt 708

45 <210> 8
 <211> 214
 <212> PRT
 <213> Crypthecodinium

55

60

65

ES 2 331 228 T3

<400> 8

5 Arg His Glu Val His Met Thr Glu Lys Arg Gly Leu Gln Phe Thr Ile
1 5 10 15

10 Cys Gly Ser Thr Gly Glu Leu Val Gln Asn Leu Gln Asp Gly Pro Thr
20 25 30

15 Ala Leu Ala Leu Cys Leu Phe Cys Phe Ser Lys Ile Pro Glu Leu Met
35 40 45

20 Asp Thr Val Phe Leu Ile Leu Lys Gly Lys Lys Val Arg Phe Leu Gln
50 55 60

25 Trp Tyr His His Ala Thr Val Met Leu Phe Cys Trp Leu Ala Leu Ala
65 70 75 80

30 Thr Glu Tyr Thr Pro Gly Leu Trp Phe Ala Ala Thr Asn Tyr Phe Val
85 90 95

35 His Ser Ile Met Tyr Met Tyr Phe Phe Leu Met Thr Phe Lys Thr Ala
100 105 110

40 Ala Lys Val Val Lys Pro Ile Ala Pro Leu Ile Thr Ile Ile Gln Ile
115 120 125

45 Ala Gln Met Val Trp Gly Leu Ile Val Asn Gly Ile Ala Ile Thr Thr
130 135 140

50 Phe Phe Thr Thr Gly Ala Cys Gln Ile Gln Ser Val Thr Val Tyr Ser
145 150 155 160

55 Ala Ile Val Met Tyr Ala Ser Tyr Phe Tyr Leu Phe Ser Gln Leu Phe
165 170 175

60 Leu Glu Ala Tyr Gly Ser Ala Gly Lys Asn Lys Lys Lys Leu Ala Arg
180 185 190

65 Glu Leu Ser Arg Lys Ile Ser Glu Ala Leu Leu Asn Ser Gly Asp Glu
195 200 205

Val Ala Lys His Leu Lys
210

<210> 9

<211> 1054

ES 2 331 228 T3

<212> ADN

<213> Thraustochitrium

5 <220>

<221> CDS

<222> (43)..(858)

10 <400> 9

```

15      gaattcggca cgagagcgcg cggagcggag acctcggccg cg atg atg gag ccg      54
                                           Met Met Glu Pro
                                           1
20      ctc gac agg tac agg gcg ctg gcg gag ctc gcc gcg agg tac gcc agc      102
      Leu Asp Arg Tyr Arg Ala Leu Ala Glu Leu Ala Ala Arg Tyr Ala Ser
           5              10              15              20

```

25

30

35

40

45

50

55

60

65

ES 2 331 228 T3

5 tcg gcg gcc ttc aag tgg caa gtc acg tac gac gcc aag gac agc ttc 150
 Ser Ala Ala Phe Lys Trp Gln Val Thr Tyr Asp Ala Lys Asp Ser Phe
 25 30 35

10 gtc ggg ccc ctg gga atc cgg gag ccg ctc ggg ctc ctg gtg ggc tcc 198
 Val Gly Pro Leu Gly Ile Arg Glu Pro Leu Gly Leu Leu Val Gly Ser
 40 45 50

15 gtg gtc ctc tac ctg agc ctg ctg gcc gtg gtc tac gcg ctg cgg aac 246
 Val Val Leu Tyr Leu Ser Leu Leu Ala Val Val Tyr Ala Leu Arg Asn
 55 60 65

20 tac ctt ggc ggc ctc atg gcg ctc cgc agc gtg cat aac ctc ggg ctc 294
 Tyr Leu Gly Gly Leu Met Ala Leu Arg Ser Val His Asn Leu Gly Leu
 70 75 80

25 tgc ctc ttc tcg ggc gcc gtg tgg atc tac acg agc tac ctc atg atc 342
 Cys Leu Phe Ser Gly Ala Val Trp Ile Tyr Thr Ser Tyr Leu Met Ile
 85 90 95 100

30 cag gat ggg cac ttt cgc agc ctc gag gcg gca acg tgc gag ccg ctc 390
 Gln Asp Gly His Phe Arg Ser Leu Glu Ala Ala Thr Cys Glu Pro Leu
 105 110 115

35 aag cat ccg cac ttc cag ctc atc agc ttg ctc ttt gcg ctg tcc aag 438
 Lys His Pro His Phe Gln Leu Ile Ser Leu Leu Phe Ala Leu Ser Lys
 120 125 130

40 atc tgg gag tgg ttc gac acg gtg ctc ctc atc gtc aag ggc aac aag 486
 Ile Trp Glu Trp Phe Asp Thr Val Leu Leu Ile Val Lys Gly Asn Lys
 135 140 145

45 ctc cgc ttc ctg cac gtc ttg cac cac gcc acg acc ttt tgg ctc tac 534
 Leu Arg Phe Leu His Val Leu His His Ala Thr Thr Phe Trp Leu Tyr
 150 155 160

50 gcc atc gac cac atc ttt ctc tcg tcc atc aag tac ggc gtc gcg gtc 582
 Ala Ile Asp His Ile Phe Leu Ser Ser Ile Lys Tyr Gly Val Ala Val
 165 170 175 180

55 aat gct ttc atc cac acc gtc atg tac gcg cac tac ttc cgc cca ttc 630
 Asn Ala Phe Ile His Thr Val Met Tyr Ala His Tyr Phe Arg Pro Phe
 185 190 195

60 ccg aag ggc ttg cgc ccg ctt att acg cag ttg cag atc gtc cag ttc 678
 Pro Lys Gly Leu Arg Pro Leu Ile Thr Gln Leu Gln Ile Val Gln Phe
 200 205 210

65

ES 2 331 228 T3

att ttc agc atc ggc atc cat acc gcc att tac tgg cac tac gac tgc 726
 Ile Phe Ser Ile Gly Ile His Thr Ala Ile Tyr Trp His Tyr Asp Cys
 5 215 220 225

 gag ccg ctc gtg cat acc cac ttt tgg gaa tac gtc acg ccc tac ctt 774
 Glu Pro Leu Val His Thr His Phe Trp Glu Tyr Val Thr Pro Tyr Leu
 10 230 235 240

 ttc gtc gtg ccc ttc ctc atc ctc ttt ttc aat ttt tac ctg cag cag 822
 Phe Val Val Pro Phe Leu Ile Leu Phe Phe Asn Phe Tyr Leu Gln Gln
 15 245 250 255 260

 tac gtc ctc gcg ccc gca aaa acc aag aag gca tag ccacgtaaca 868
 Tyr Val Leu Ala Pro Ala Lys Thr Lys Lys Ala
 20 265 270

 gtagaccagc agcgccgagg acgctgtgccg cgttatcgcg aagcacgaaa taaagaagat 928

 catttgattc aacgaggcta cttgcggcca cgagaaaaaaaa aaaaaaaaaa aaaaaaaaaa 988
 30 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1048

 ctcgag 1054
 35
 <210> 10
 <211> 271
 <212> PRT
 40 <213> Thraustochitrium

 <400> 10
 45 Met Met Glu Pro Leu Asp Arg Tyr Arg Ala Leu Ala Glu Leu Ala Ala
 1 5 10 15

 Arg Tyr Ala Ser Ser Ala Ala Phe Lys Trp Gln Val Thr Tyr Asp Ala
 50 20 25 30

 Lys Asp Ser Phe Val Gly Pro Leu Gly Ile Arg Glu Pro Leu Gly Leu
 55 35 40 45

 Leu Val Gly Ser Val Val Leu Tyr Leu Ser Leu Leu Ala Val Val Tyr
 60 50 55 60

 Ala Leu Arg Asn Tyr Leu Gly Gly Leu Met Ala Leu Arg Ser Val His
 65 65 70 75 80

 Asn Leu Gly Leu Cys Leu Phe Ser Gly Ala Val Trp Ile Tyr Thr Ser

ES 2 331 228 T3

				85					90					95		
5	Tyr	Leu	Met	Ile	Gln	Asp	Gly	His	Phe	Arg	Ser	Leu	Glu	Ala	Ala	Thr
				100					105					110		
10	Cys	Glu	Pro	Leu	Lys	His	Pro	His	Phe	Gln	Leu	Ile	Ser	Leu	Leu	Phe
				115					120					125		
15	Ala	Leu	Ser	Lys	Ile	Trp	Glu	Trp	Phe	Asp	Thr	Val	Leu	Leu	Ile	Val
				130					135					140		
20	Lys	Gly	Asn	Lys	Leu	Arg	Phe	Leu	His	Val	Leu	His	His	Ala	Thr	Thr
				145					150					155		160
25	Phe	Trp	Leu	Tyr	Ala	Ile	Asp	His	Ile	Phe	Leu	Ser	Ser	Ile	Lys	Tyr
					165									170		175
30	Gly	Val	Ala	Val	Asn	Ala	Phe	Ile	His	Thr	Val	Met	Tyr	Ala	His	Tyr
					180									185		190
35	Phe	Arg	Pro	Phe	Pro	Lys	Gly	Leu	Arg	Pro	Leu	Ile	Thr	Gln	Leu	Gln
														195		200
40	Ile	Val	Gln	Phe	Ile	Phe	Ser	Ile	Gly	Ile	His	Thr	Ala	Ile	Tyr	Trp
														210		215
45	His	Tyr	Asp	Cys	Glu	Pro	Leu	Val	His	Thr	His	Phe	Trp	Glu	Tyr	Val
														225		230
50	Thr	Pro	Tyr	Leu	Phe	Val	Val	Pro	Phe	Leu	Ile	Leu	Phe	Phe	Asn	Phe
														245		250
55	Tyr	Leu	Gln	Gln	Tyr	Val	Leu	Ala	Pro	Ala	Lys	Thr	Lys	Lys	Ala	
														260		265
60																270

<210> 11

55 <211> 421

<212> ADN

<213> *Phitophthora infestans*

<220>

60 <221> CDS

<222> (1)..(279)

65

ES 2 331 228 T3

<400> 12

5	His Thr Ile Met Tyr Thr Tyr Tyr Phe Val Ser Ala His Thr Arg Asn	1	5	10	15
10	Ile Trp Trp Lys Lys Tyr Leu Thr Arg Ile Gln Leu Ile Gln Phe Val	20	25	30	
15	Thr Met Asn Val Gln Gly Tyr Leu Thr Tyr Ser Arg Gln Cys Pro Gly	35	40	45	
20	Met Pro Pro Lys Val Pro Leu Met Tyr Leu Val Tyr Val Gln Ser Leu	50	55	60	
25	Phe Trp Leu Phe Met Asn Phe Tyr Ile Arg Ala Tyr Val Phe Gly Pro	65	70	75	80
30	Lys Lys Pro Ala Val Glu Glu Ser Lys Lys Lys Leu	85	90		
35					
40					
45					
50					
55					
60					
65					