



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 337 564**

51 Int. Cl.:
C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03704382 .5**

96 Fecha de presentación : **13.01.2003**

97 Número de publicación de la solicitud: **1472357**

97 Fecha de publicación de la solicitud: **03.11.2004**

54 Título: **Método para elaborar ácidos grasos poliinsaturados mediante un nuevo gen de elongasa.**

30 Prioridad: **30.01.2002 DE 102 03 713**
11.02.2002 DE 102 05 607

45 Fecha de publicación de la mención BOPI:
27.04.2010

45 Fecha de la publicación del folleto de la patente:
27.04.2010

73 Titular/es: **BASF Plant Science GmbH**
Carl-Bosch-Strasse 38
67056 Ludwigshafen, DE

72 Inventor/es: **Lerchl, Jens;**
Heinz, Ernst y
Zank, Thorsten

74 Agente: **Carvajal y Urquijo, Isabel**

ES 2 337 564 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

Método para elaborar ácidos grasos poliinsaturados mediante un nuevo gen de elongasa.

5 Esta invención se refiere a métodos para la producción de moléculas de ácido graso de C₂₀ o C₂₂ con al menos dos enlaces dobles donde se emplea un gen de elongasa con la secuencia SEQ ID NO:1 o sus homólogos, derivados o análogos, una construcción genética que comprende este gen o sus homólogos, derivados o análogos.

10 Determinados productos y productos secundarios que proceden naturalmente de procesos metabólicos en células microbianas o en las células de animales y ventajosamente de plantas son útiles para un amplio espectro de industrias, incluyendo la industria de alimentos para animales, productos alimenticios para humanos, cosméticos y fármacos. Estas moléculas que en conjunto se denominan "sustancias químicas finas", también incluyen lípidos y ácidos grasos poliinsaturados (polyunsaturated fatty acids, PUFAs), por ejemplo, al alimento para niños para incrementar el valor
15 nutricional de estos alimentos. Por ejemplo, los PUFAs tienen un efecto positivo en el nivel de colesterol en la sangre de los humanos y por lo tanto son adecuados para la protección contra padecimientos cardiacos. Las sustancias químicas finas tales como ácidos grasos poliinsaturados (polyunsaturated fatty acids, PUFAs) pueden aislarse de fuentes animales, por ejemplo de pescado, o producirse a gran escala utilizando microorganismos por medio del cultivo de microorganismos los cuales se han desarrollado de tal manera que los mismos produzcan y acumulen o secreten, grandes
20 cantidades de una o más moléculas deseadas.

Los microorganismos que son especialmente adecuados para preparar PUFAs son microorganismos, tales como algas de la especie *Phaeodactylum tricorutum* o *Cryptothecodinium*, Ciliatas como *Stylonichia* o *Colpidium*, hongos como *Mortierella*, *Entomophthora*, *Mucor*. Mediante la selección de cepas se ha desarrollado un número de cepas
25 mutantes de los microorganismos correspondientes que producen una serie de compuestos deseables, incluyendo los PUFAs. La selección de cepas con producción mejorada de una molécula determinada de una molécula determinada es, no obstante, un método difícil y que consume tiempo.

Como una alternativa, las sustancias químicas finas pueden producir adecuadamente a gran escala por medio de
30 la producción de plantas que se han desarrollado de tal manera que ellas producen los PUFAs antes mencionados. Plantas particularmente bien adecuadas para este propósito son plantas de frutos oleaginosos que contienen grandes cantidades de compuestos lípidos, tales como colza, canola, lino, soya, girasol, cardos, borraja y onagra. Sin embargo, otros cultivos que contienen aceites o lípidos y ácidos grasos son muy adecuados, como se mencionó en la descripción detallada de la presente invención. La reproducción de plantas convencionales por medio de la selección de plantas
35 adecuadas, ha conducido a un desarrollo de una serie de plantas mutantes las cuales producen un espectro de lípidos y ácidos grasos, co-factores y enzimas deseables. Pero otras plantas útiles que contienen aceites o lípidos y ácidos grasos también son muy adecuadas, tal como se menciona en la descripción minuciosa de esta invención. Por medio del cultivo convencional, mediante la selección de plantas adecuadas, se ha desarrollado una serie de plantas mutantes que producen un espectro de lípidos y ácidos grasos, co-factores y enzimas deseables. La selección de nuevas variedades de plantas
40 con producción mejorada de una molécula determinada es un procedimiento complicado que consume tiempo o que incluso es imposible si el compuesto no está presente de manera natural en la planta correspondiente, como en el caso de ácidos grasos de C₂₀ poliinsaturados y de aquellos con cadenas de carbono más largas.

Debido a las propiedades positivas de ácidos grasos insaturados no han faltado intentos en el pasado de poner
45 a disposición genes que participan en la síntesis de ácidos grasos o triglicéridos para producir aceites en diferentes organismos con contenido modificado de ácidos grasos insaturados. De esta manera, en WO 91/13972 y su equivalente estadounidense se describe una D-9-desaturasa. En WO 93/11245 se reivindica una D-15-desaturasa, en WO 94/11516 una D-12-desaturase. D-6-desaturasas se describen en WO 93/06712, US 5,614,393, WO 96/21022 y WO 99/27111. Otras desaturasas se describen en EP-A-0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-
50 A-0 794 250, Stukey y colaboradores, J. Biol. Chem., 265, 1990: 20144-20149, Wada y colaboradores, Nature 347, 1990: 200-203 o Huang y colaboradores, Lipids 34, 1999: 649-659. En WO 96/13591 se describe y se reivindica una D-6-palmitoil-ACP-desaturasa. La caracterización bioquímica de las diferentes desaturasas se ha efectuado hasta ahora solo de manera insuficiente ya que las enzimas como proteínas enlazantes de membrana solo pueden aislarse y caracterizarse con gran dificultad (McKeon y colaboradores, Methods in Enzymol. 71, 1981: 12141-12147, Wang y
55 colaboradores, Plant Physiol. Biochem., 26, 1988: 777-792).

En levaduras pudo detectarse tanto una desviación del espectro de ácidos grasos para producir ácidos grasos insaturados como también un incremento en la productividad (véase Huang y colaboradores, Lipids 34, 1999: 649-659, Napier y colaboradores, Biochem. J., Vol. 330, 1998: 611-614). La expresión de las diferentes desaturasas en plantas transgénicas no mostró, sin embargo, el éxito deseado. Una desviación del espectro de ácido graso para producir ácidos grasos insaturados pudo mostrarse pero simultáneamente se mostró que el desempeño de síntesis de las plantas transgénicas decreció fuertemente, lo que significa que solo pequeñas cantidades de aceites pudieron aislarse en comparación con las plantas de partida.

65

ES 2 337 564 T3

La clonación y expresión de las elongasas que alongan los ácidos grasos insaturados como sustrato de la reacción enzimática en al menos dos átomos de C hasta ahora no se había descrito ni en levaduras ni en plantas.

5 Esto significa que ni las levaduras ni las plantas de cultivo producen de manera natural ácidos grasos poliinsaturados con C_{20} y/o C_{22} con al menos dos enlaces dobles en la molécula de ácido graso como ácido araquidónico (ARA) y/o ácido eicosapentaenoico (EPA) y/o ácido docosahexaenoico (DHA).

10 Tanto ahora como antes existe aún una gran demanda de genes que codifican para enzimas que participan en la biosíntesis de ácidos grasos insaturados y hacen posible producirlos a escala industrial. Existe una demanda particularmente alta de elongasas que alongan los ácidos grasos insaturados en al menos dos átomos de C. Ninguno de los métodos biotecnológicos conocidos hasta ahora para la producción de ácidos grasos poliinsaturados proporciona los mencionados ácidos grasos en cantidades económicamente aprovechables.

15 En la expresión de genes siempre existen problemas, es decir que mediante la expresión no se llega al incremento esperado en la producción del producto valioso deseado.

20 Por esto, la tarea consistió en poner a disposición métodos para la síntesis de ácidos grasos insaturados tales como ácidos grasos poliinsaturados de C_{20} y/o C_{22} con al menos dos enlaces dobles en la molécula del ácido graso como el ácido araquidónico (ARA) y/o ácido eicosapentaenoico (EPA) y/o ácido docosahexaenoico (DHA).

Este problema se resolvió mediante el método según la reivindicación 1.

25 En este se usa un ácido nucleico que comprende una secuencia de nucleótidos que codifica un polipéptido, el cual alonga ácidos grasos de C_{16} o C_{18} con al menos dos enlaces dobles en la molécula de ácido graso, seleccionado del grupo que se compone de

a) una secuencia de ácido nucleico representada en SEQ ID NO:1,

30 b) una secuencia de ácido nucleico que se deriva de la secuencia representada en SEQ ID NO:2 según el código genético degenerado,

35 c) derivados de la secuencia representada en SEQ ID NO:1 que codifica polipéptidos con al menos 50% de homología con la secuencia que codifica las secuencias de aminoácidos en SEQ ID NO:2; la secuencia actúa como elongasa de C_{16} o C_{18} .

40 Ventajosamente, la secuencia de nucleótidos alonga ácidos grasos seleccionados del grupo previamente mencionado de C_{16} o C_{18} con al menos dos enlaces dobles en el ácido graso en al menos dos átomos de carbono, aunque los siguientes ácidos grasos no se alargan: $C_{18:3}^{\Delta 5,9,12}$, $C_{20:3}^{\Delta 8,11,14}$, $C_{20:4}^{\Delta 5,8,11,14}$ y $C_{20:5}^{\Delta 5,8,11,14,17}$. Las elongasas muestran una preferencia frente a ácidos grasos de $C_{18:3}^{\Delta 6,9,12}$, $C_{18:4}^{\Delta 6,9,12,15}$ y $C_{16:3}^{\Delta 7,10,13}$, que es superior en al menos el factor 1,5, preferible en al menos el factor 1,6, particularmente en el factor 1,7 o muy particularmente preferible en al menos el factor 1,8 en comparación con los ácidos grasos insaturados como ácidos grasos de $C_{18:2}^{\Delta 9,12}$, $C_{18:3}^{\Delta 4,7,10}$, $C_{18:3}^{\Delta 5,8,11}$, $C_{18:3}^{\Delta 7,10,13}$, $C_{18:3}^{\Delta 8,11,14}$, $C_{18:3}^{\Delta 9,12,15}$ o $C_{18:3}^{\Delta 5,9,12}$.

45 Las secuencias de ácido nucleico que alongan ácidos grasos de C_{16} o C_{18} provienen originalmente de manera ventajosa de hongos, preferible de hongos como los oomycetes, por ejemplo oomycetes como los de la especie fitofora, particularmente preferible de oomycetes de la especie y tipo *Phytophthora infestans*.

50 Los ácidos nucleicos pueden usarse en el método según la invención para la modificación de aceites, ácidos grasos, lípidos, compuestos provenientes de lípidos y lo más preferible para la producción de ácidos grasos poliinsaturados.

55 Como organismos huésped para los ácidos nucleicos son adecuados de manera ventajosa microorganismos tales como *Phaeodactylum*, *Colpidium*, *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodium* así como otras algas y hongos y plantas, en particular plantas de frutos oleaginosos que se usan a gran escala en la industria para la producción de un gran número de productos químicos finos.

60 Con los vectores de clonación y técnicas de manipulación genética de los organismos arriba mencionados y ciliados divulgados, por ejemplo, en WO 98/01572 o los métodos y vectores descritos por Falciatore y colaboradores, 1999, *Marine Biotechnology* 1(3):239-251, así como Dunahay y colaboradores, 1995, *Genetic transformation of diatoms*, *J. Phycol.* 31:10004-1012 y en las citas allí encontradas, para algas y organismos relacionados tales como *Phaeodactylum tricorutum*, las moléculas de ácido nucleico pueden usarse para la modificación por ingeniería genética de estos organismos, de tal modo que se vuelvan productores mejores o más eficientes de uno o varios productos químicos finos. Esta producción mejorada o eficiencia de la producción mejorada de un producto químico fino puede provocarse por una acción directa de la manipulación de un ácido nucleico según la invención o por una acción directa de esta manipulación.

65 Briofitos (mohos) y algas son los únicos sistemas vegetales conocidos que producen cantidades considerables de ácidos grasos poliinsaturados, como ácido araquidónico (ARA) y/o ácido eicosapentaenoico (EPA) y/o ácido docosahexaenoico (DHA). Sistemas fúngicos tales como oomycetes (eucariotas/estramenopilas/oomycetes/fitiales/fitiaceen)

producen los ácidos grasos previamente mencionados. Por esto son adecuadas moléculas de ácido nucleico que provienen de un oomizete como *Phytophthora infestans*, particularmente para la modificación del sistema de producción de lípido y de PUFA en un huésped, en particular en microorganismos como los microorganismos previamente mencionados, y en plantas tales como plantas de frutos oleaginosos, por ejemplo colza, canola, lino, soya, girasol, borraja.

5 También pueden usarse ácidos nucleicos a partir de un oomizete como *Phytophthora infestans* para la identificación de tales secuencias de ADN y enzimas en estas especies que son adecuadas para la modificación de la biosíntesis de moléculas precursoras de PUFAs en los organismos correspondientes.

10 El hongo *Phytophthora infestans* es un miembro de los oomizetes. Está relacionado con otros hongos que pueden crecer en ausencia de luz. Hongos como fitofthora son muy homólogos entre sí a nivel de secuencia de ADN y de polipéptido lo cual hace posible el uso de screening (selección) heteróloga de moléculas de ADN con sondas que provienen de otros hongos u organismos, de modo que al estar presentes otras secuencias de ácido nucleico junto con la secuencia arriba nombrada puede derivarse una secuencia de consenso que es adecuada para el screening heterólogo o para la observación funcional y la predicción de funciones de gen en terceras especies. Sin embargo, por ahora no es posible una predicción de la función de las proteínas o enzimas codificadas por las secuencias. La capacidad de identificar estas funciones, por ejemplo la predicción de la especificidad de sustrato de las enzimas puede, por lo tanto, ser de significado significativo. Estas moléculas de ácido nucleico también pueden servir como secuencias de referencia para formar mapas de otros hongos o para derivar cebadores (primers) de PCR.

20 En el marco de la invención se aisló un gen PSE funcionalmente activo a partir del oomizete *Phytophthora infestans* (eucariotas/estromatopilas, oomycetes/fitiales/pitiaceen). El gen es adecuadamente ventajoso para la producción de ácidos grasos poliinsaturados de cadena larga, preferiblemente con más de dieciséis o dieciocho átomos de carbono en la estructura fundamental de carbono del ácido graso y/o al menos dos enlaces dobles en la cadena de carbono; la enzima codificada por la secuencia tiene una preferencia frente a los ácidos grasos mencionados arriba en el proceso de alargamiento.

30 Esta molécula de ácido nucleico codifica para una proteína que se denomina aquí como elongasa específica para PUFA (PSEs o PSE en singular). Estas PSEs pueden ejercer una función, por ejemplo, que esté involucrada en el metabolismo (por ejemplo en la biosíntesis o en la degradación) de compuestos que sean necesarios para la síntesis de lípidos o de ácidos grasos, como PUFAs, o que participen en el transporte transmembrana de uno o varios compuestos de lípido o de ácido graso ya sea hacia o desde la célula.

35 En esta solicitud se representa la función de la secuencia en más detalle. Hemos aislado un gen de oomicete activo funcionalmente que es adecuado para la producción de ácidos grasos poliinsaturados de cadena larga, preferiblemente con más de dieciséis o dieciocho átomos de carbono en la estructura fundamental de carbono del ácido graso y/o al menos dos dobles enlaces en la cadena de carbono; la enzima codificada por las secuencias presentan una preferencia frente a los ácidos grasos mencionados arriba en el proceso de alargamiento. Por lo tanto, aquí se habla de un gen o proteína PSE. Otras publicaciones y patentes no divulgan ni muestran ningún gen PSE funcionalmente activo, aunque existen diferentes solicitudes de patente conocidas que muestran el alargamiento de ácidos grasos saturados con cadena corta o media (WO 98/46776 y US 5,475,099) o el alargamiento o producción de ácidos grasos de cadena larga pero que tienen no más de un enlace doble o conducen a ésteres de cera de ácidos grasos de cadena larga (véase: WO 98/54954, WO 96/13582, WO 95/15387).

45 Aunque WO 99/64616, WO 98/46763, WO 98/46764, WO 98/46765 describen la producción de PUFAs en plantas transgénicas y muestran la clonación y expresión funcional de actividades correspondientes de desaturasa, en particular a partir de hongos, no obstante no muestran ningún gen codificante PSE imprescindible y ninguna actividad PSE funcional. Se ha mostrado la preparación de un ácido trienoico con cadena de carbono C18 y reivindicado por medio de ácido gamma-linolénico, aunque hasta ahora no se ha enseñado la preparación de ácidos grasos poliinsaturados de cadena muy larga (con cadena de carbono C20 y más larga, así como de ácidos trienoicos y de tipos insaturados mayores) y la especificidad por el sustrato de la elongasa.

55 Para la preparación de PUFAs de cadena larga los ácidos grasos poliinsaturados de C₁₆ y/o C₁₈ mediante la actividad enzimática de una elongasa, se alargan al menos en dos átomos de carbono. La secuencia de ácido nucleico codifica la primera elongasa fúngica que puede alargar ácidos grasos de C₁₆ y/o C₁₈ con al menos dos enlaces dobles en el ácido graso. Después de una ronda de elongación, esta actividad enzimática conduce a ácidos grasos de C₂₀, y después de dos, tres y cuatro rondas de elongación a ácidos grasos de C₂₂, C₂₄ o C₂₆. En la actividad enzimática es ventajoso que no todos los ácidos grasos insaturados de C₂₀ se alarguen. Esto hace posible la síntesis específica de ácidos grasos insaturados deseados o de mezclas de ácidos grasos. Con la elongasa arriba mencionada también pueden sintetizarse PUFAs más largos. La actividad de las elongasas conduce preferiblemente a ácidos grasos de C₂₀ y/o C₂₂ con al menos dos enlaces dobles en la molécula de ácido graso, preferiblemente con tres o cuatro enlaces dobles, particularmente se prefieren ácidos grasos con un enlace doble en posición Δ6. Después de que ha tenido lugar el alargamiento con las enzimas, pueden efectuarse otros pasos de desaturación. Por lo tanto, los productos de la actividad de elongasa y de la posible mayor desaturación conducen a PUFAs preferidos con un alto grado de desaturación, como ácido docosadiénico, ácido araquidónico, ácido ω6-eicosatriendihomo-γ-linolénico, ácido eicosapentenoico, ácido ω3-eicosatrienoico, ácido ω3-eicosatetraenoico, ácido docosapentaenoico o ácido docosahexaenoico. Sustratos de la actividad enzimática son, por ejemplo, ácido taxólico; ácido 6,9-octadecadienoico, ácido linólico, ácido γ-linolénico, ácido pinolénico, ácido α-linolénico o ácido estearidónico. Sustratos preferidos son ácido linólico, ácido γ-linolénico

y/o ácido α -linolénico. Los ácidos grasos de C₁₆ y/o C₁₈ con al menos dos enlaces dobles en el ácido graso pueden alargarse mediante la actividad enzimática en forma del ácido graso libre o en forma del éster, tales como fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicerina, diacilglicerina o triacilglicerina.

5 Usando vectores de clonación que son adecuados para el uso en plantas y para la transformación de plantas, tales como aquellos que se han divulgado 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; in: Transgenic Plants, Vol. 1, Engineering and Utilization, editado por Kung y R. Wu, Academic Press, 1993, 15-38; B. Jené y colaboradores, Techniques for Gene Transfer, en: Transgenic Plants, vol. 1, Engineering and Utilization, editado por: Kung y R. Wu,
10 Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225)) y en las referencias nombradas en estos documentos, pueden usarse los ácidos nucleicos empleados en el método de la invención para la modificación genética recombinante de un espectro amplio de plantas, de modo que estas se vuelvan un productor mejor o más eficiente de uno o varios productos derivados de lípidos, como PUFAs. Esta producción o eficiencia mejorada de la producción de un producto derivado de lípidos, como PUFAs, puede generarse mediante
15 acción directa de la manipulación o una acción indirecta de esta manipulación.

Existente una serie de mecanismos por medio de los cuales la modificación de una proteína PSE puede influir directamente el rendimiento, producción y/o eficiencia de la producción de un producto químico fino a partir de una planta de fruto oleaginoso o un microorganismo debido a una proteína modificada. La cantidad o la actividad de la proteína,
20 ácido nucleico o gen PSE puede elevarse de modo que se produzcan grandes cantidades de estos compuestos *de novo*, porque esta actividad o capacidad le faltaba a los organismos para la biosíntesis antes de la introducción del ácido nucleico correspondiente y/o el gen correspondiente.

La introducción de uno de los ácidos nucleicos arriba mencionados y/o de un gen PSE en una planta o un microorganismo puede elevar no solo el flujo de biosíntesis hacia el producto final, sino también la composición triacilglicerina correspondiente o crearlos *de novo*. Así mismo, pueden aumentarse la cantidad o actividad de otros genes que son necesarias en la importación de sustancias alimenticias para la biosíntesis de uno o más productos químicos finos (por ejemplo, ácidos grasos, lípidos polares y neutrales) de modo que se eleva la concentración de estos precursores, co-
30 factores o compuestos intermedios dentro de las células o dentro de los compartimientos de almacenamiento, por lo cual la capacidad de las células para producir PUFAs, tal como se describe en lo sucesivo, sigue incrementándose. Los ácidos grasos y lípidos son productos químicos finos valiosos de por sí; optimizando la actividad o incrementando la cantidad de uno o varios PSEs que toman parte en la biosíntesis de estos compuestos o eliminando la actividad de uno o más PSEs que participan en la degradación de estos compuestos, puede ser posible incrementar los rendimientos, producción y/o eficiencia de la producción de moléculas de ácido graso y lípidos a partir de plantas o
35 microorganismos.

La mutagénesis del gen PSE o del ácido nucleico también puede conducir a una proteína PSE con actividades modificadas que afectan de manera directa o indirecta la producción de uno o más productos químicos finos deseados. Por ejemplo, el número o la actividad del gen PSE o del ácido nucleico puede incrementarse de modo que los residuos normales de metabolismo o los productos secundarios de la célula (cuya cantidad se eleva posiblemente debido a una sobreproducción de los químicos finos deseados) se exportan de manera eficiente antes de que destruyan otras moléculas o eliminen procesos dentro de la célula (lo cual reduciría la capacidad vital de la célula) o perturbarían las rutas de biosíntesis de los químicos finos (por lo cual el rendimiento, producción o eficiencia de la producción de los químicos deseados se reducirían). Las cantidades intracelulares relativamente grandes de los químicos finos
45 mismos también pueden ser tóxicos para la célula o perturbar los mecanismos de retroalimentación de la enzima, tales como la regulación alostérica; por ejemplo, las mismas podrían incrementar la asignación del PUFA en la fracción de triacilglicerina debido a una actividad incrementada o número de otras enzimas o enzimas desintoxicantes de la ruta de PUFA la cual continúa corriente abajo; la capacidad vital de células de semillas podría incrementar, lo cual a su vez conduce a un mejor desarrollo de células en el cultivo o a semillas que producen la sustancia química fina deseada. Pueden manipularse el gen PSE o los ácidos nucleicos de tal manera que se producen las cantidades correspondientes de las diversas moléculas de lípidos y de moléculas de ácido graso. Esto puede tener un efecto drástico en la composición de lípidos de la membrana celular y genera nuevos aceites además de la presencia de PUFAs sintetizados nuevamente. Puesto que cada tipo de lípido tiene diferentes propiedades físicas, un cambio en la composición de lípidos de una membrana puede modificar significativamente la fluidez de membrana. Los cambios
55 en la fluidez de la membrana pueden tener un efecto en el transporte de moléculas a través de la membrana y en la integridad de las células, ambos tienen un efecto decisivo en la producción de sustancias químicas finas. Además, en las plantas estos cambios también pueden tener un efecto en otras características tales como la tolerancia a las situaciones de tensión abiótica y biótica.

60 La tolerancia a las situaciones de tensión abiótica y biótica es una característica general que es deseable impartir a un amplio espectro de plantas tales como maíz, trigo, centeno, avena, triticale, arroz, cebada, soya, cacahuate, abrojo, algodón, colza y canola, yuca, pimienta, girasol y tagetes, plantas solanáceas como patata, tabaco, berenjena y tomate, especies Vicia, guisantes, alfalfa, plantas de arbusto (café, cacao, té), especies Salix, árboles (palma de aceite, coco) y pastos perennes y cultivos de forraje. Estos cultivos también son plantas destino preferidas para ingeniería genética como otra forma de realización de acuerdo con la invención. Plantas particularmente preferidas de acuerdo con la invención son plantas de frutos oleaginosos, tales como soya, maní o cacahuate, colza, canola, girasol, abrojo, árboles (palma de aceite, coco) o cultivos tales como maíz, trigo, centeno, avena, triticale, arroz, cebada, alfalfa o plantas
65 arbusto (café, cacao, té).

ES 2 337 564 T3

En el método según la invención se emplean moléculas de ácido nucleico aisladas (por ejemplo, ADNcs) que comprenden una secuencia de nucleótidos que codifica una PSE o partes biológicamente activas de las mismas. En formas particularmente preferidas de realización, la molécula de ácido nucleico comprende una de las secuencias de nucleótidos representadas en SEQ ID NO:1 o la región codificante o un complemento de una de estas secuencias de nucleótidos. En otras formas particularmente preferidas de realización, la molécula aislada de ácido nucleico comprende una secuencia de nucleótidos que es homóloga en al menos aproximadamente 50%, preferiblemente al menos alrededor de 60%, más preferible alrededor de 70%, 80% ó 90% y aún más preferible al menos alrededor de 95%, 96%, 97%, 98%, 99% o más; en el sentido de la invención se entiende por homología a la identidad. En otras formas preferidas de realización la molécula de ácido nucleico aislada codifica una de las secuencias de aminoácidos representadas en la secuencia SEQ ID NO:2. El gen PSE o la secuencia de ácido nucleico también posee al menos una de las actividades PSE aquí descritas.

En otra forma de realización, la molécula de ácido nucleico aislada codifica una proteína o una parte de la misma; la proteína o la parte de la misma contiene una secuencia de aminoácido que es suficientemente homóloga con una secuencias de aminoácido de la secuencia SEQ ID NO:2, que la proteína o la parte de la misma retiene una actividad PSE y por homología se entiende identidad en el sentido de la invención. Preferiblemente, la proteína o la parte de la misma, que se codifican por la molécula de ácido nucleico, mantienen la capacidad de participar en el metabolismo de compuestos necesarios para la estructura de membranas celulares de plantas o en el transporte de moléculas por estas membranas. En una forma de realización, la proteína codificada por la molécula de ácido nucleico presenta una homología de al menos aproximadamente 50%, preferiblemente de al menos alrededor de 60% y más preferiblemente de al menos alrededor de 70%, 80% o 90% y lo más preferible de al menos alrededor de 95%, 96%, 97%, 98%, 99% o más con una secuencia de aminoácidos de la secuencia SEQ ID NO:2. En otra forma preferida de realización, la proteína es una proteína de *Phytophthora infestans* de longitud completa que es esencialmente homóloga con una secuencia completa de aminoácidos de la SEQ ID NO:2 (la cual proviene del marco de lectura abierto mostrado en SEQ ID NO: 1).

En otra forma preferida de realización, la molécula de ácido nucleico aislada proviene de *Phytophthora infestans* y codifica una proteína (por ejemplo, una proteína de fusión PSE) que contiene un dominio biológicamente activo que es homóloga (idéntica) en al menos alrededor de 50% o más con una secuencia de aminoácido de la secuencia SEQ ID NO: 2 y mantiene la capacidad de participar en el metabolismo de hongos de compuestos necesarios para la síntesis de ácidos grasos insaturados o en el transporte de moléculas por membranas o posee al menos una de las actividades listadas en la tabla I, y comprende también secuencias heterólogas de ácido nucleico que codifican un polipéptido heterólogo o proteína regulatoria.

Los ácidos nucleicos codifican para proteínas que alongan ácidos grasos seleccionados del grupo $C_{18:2}^{\Delta 9,12}$, $C_{18:3}^{\Delta 4,7,10}$, $C_{18:3}^{\Delta 5,8,11}$, $C_{18:3}^{\Delta 6,9,12}$, $C_{18:3}^{\Delta 7,10,13}$, $C_{18:3}^{\Delta 8,11,14}$, $C_{18:3}^{\Delta 9,12,15}$, $C_{18:4}^{\Delta 6,9,12,15}$, $C_{18:3}^{\Delta 5c,9,12}$ o $C_{16:3}^{\Delta 7,10,13}$, mientras que ácidos grasos como $C_{18:3}^{\Delta 5t,9,12}$, $C_{20:3}^{\Delta 8,11,14}$, $C_{20:4}^{\Delta 5,8,11,14}$ y $C_{20:5}^{\Delta 5,8,11,14,17}$ no se alongan. El gen PSE se expresó de manera heteróloga en levadura y los diferentes ácidos grasos poliinsaturados se alimentan individualmente y se hacen reaccionar por parte de la levadura transformada. Se analizaron perfiles de ácido graso de levaduras transgénicas por medio de GLC. El porcentaje de ácidos grasos alimentados convertidos se calculó de la siguiente manera: % molar (producto) $\times 100 / [\% \text{ molar (material de partida)} + \% \text{ molar (producto)}]$.

La tabla I refleja ácidos grasos que se alongan mediante la enzima arriba mencionada.

(Tabla pasa a página siguiente)

ES 2 337 564 T3

TABLA I

Ácidos grasos (en % molar) de células de levadura transformadas con el plásmido pYES2 (control) y pY2piPSE1

5 Se alimentaron ácidos grasos 18:2, γ -18:3, α -18:3 o 18:4. La fracción de lípido de transmetilo y el perfil de ácido graso se determinó por medio de GC. Cálculo %-elongación: $100 \times \% \text{ molar (producto)} / [\% \text{ molar (material de partida)} + \text{mol \% (producto)}]$. En negrilla se imprimen sustratos y ácidos grasos elongados.

	% de ácidos grasos							
	PYES2				PY2piPSE1			
Ácidos grasos	+18:	+ γ - 18:	+ α - 18:	+18:	+18:	+ γ - 18:	+ α - 18:	+18:
	2	3	3	4	2	3	3	4
16: 0	13.7	16.4	14.3	19.5	9.8	9.8	7.3	10.6
16: 1 ^{Δ9}	8.8	6.5	3.4	7.8	9.9	4.1	1.3	2.3
18: 0	6.9	9.9	11.3	13.3	8.6	12.3	14.4	19.4
18: 1 ^{Δ9}	9.9	10.5	6.0	13.6	15.3	11.8	5.6	8.4
18: 2 ^{Δ9,12}	60.7	-	-	-	41.0	-	-	-
18: 3 ^{Δ6,9,12}	-	56.7	-	-	-	51.1	-	-
18: 3 ^{Δ9,12,15}	-	-	64.7	-	-	-	65.1	-
18:A 4 ^{Δ6,9,12,15}	-	-	-	45.7	-	-	-	48.3
20: 2 ^{Δ11,14}	-	-	-	-	2.1	-	-	-
20:3 ^{Δ8,11,14}	-	-	-	-	-	11.0	-	-
20:3 ^{Δ11,14,17}	-	-	-	-	-	-	6,2	-
20:4 ^{Δ8,11,14,17}	-	-	-	-	-	-	-	11,0
%elongación	n.d.	n.d.	n.d.	n.d.	4,9	17,7	8,7	18,5

El método según la invención puede emplearse para la modulación de la producción de sustancias químicas finas a partir de microorganismos, como ciliados, hongos, levaduras, bacterias, algas y/o plantas tales como maíz, trigo, centeno, avena, triticale, arroz, mijo, soya, maní o cacahuete, algodón, abrojo, especies de brasicas como colza, abrojo, canola y remolacha, pimienta, girasol, borraja, onagra y tagetes, plantas solanáceas como patata, tabaco, berenjena y tomate, especies de vicia, guisantes, yuca, alfalfa, plantas de arbusto (café, cacao, té), especies de salix, árboles (palma de aceite, coco) y pastos perennes y frutas de forraje, ya sea de manera directa (por ejemplo si la sobre-expresión u optimización de una proteína de biosíntesis de ácido graso tiene una influencia directa sobre el rendimiento, producción y/o eficiencia de la producción del ácido graso a partir de organismos modificados) o pueden tener un efecto indirecto que no obstante conduce a un incremento del rendimiento, producción y/o eficiencia de la producción de un compuesto deseado o una disminución de compuestos no deseados (por ejemplo, si la modulación del metabolismo de lípidos y ácidos grasos, co-factores y enzimas conduce a modificaciones del rendimiento, producción y/o eficiencia de la producción o de la composición de los compuestos deseados dentro de las células, lo cual a su vez puede afectar la producción de una o más sustancias químicas finas). A continuación se siguen ilustrando aspectos de la invención.

I. Sustancias químicas finas y PUFAs

El término "sustancias químicas finas" es conocido en el campo técnico y abarca moléculas que se han producido por un organismo y encuentran aplicaciones en diferentes industrias como, pero sin limitarse a, la industria farmacéutica, agrícola, alimenticia y cosmética. Estos compuestos abarcan lípidos, ácidos grasos, cofactores y enzimas, etc. (como, por ejemplo, se describe en Kuninaka, A. (1996) Nucleótidos y compuestos relacionados, pág. 561-612, en Biotechnology vol. 6, Rehm y colaboradores, editorial VCH: Weinheim y las referencias bibliográficas allí contenidas), lípidos, ácidos grasos saturados e insaturados (por ejemplo, ácido araquidónico), vitaminas y cofactores (tal como se describe en Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, Vitaminas, pág. 443-613 (1996) VCH: Weinheim y las citas bibliográficas contenidas allí; y las sustancias químicas descritas en 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 el 1.-3. Sept. 1994 en Penang, Malasia, AOCS Press (1995)), enzimas y todas las otras de Gutcho (1983) en Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086, y las referencias bibliográficas allí indicadas. El metabolismo y los usos de determinadas sustancias químicas se siguen ilustrando a continuación.

La combinación de diferentes moléculas precursoras y enzimas de biosíntesis conduce a la preparación de diferentes moléculas de ácidos grasos, lo cual tiene un efecto decisivo en la composición de la membrana. Puede suponerse que los PUFAs no solo se incorporan sencillamente en triacilglicerina sino también en lípidos de membrana.

5 La síntesis de membranas es un proceso bien caracterizado en el que participa un número de componentes, incluyendo lípidos, como parte de la membrana de dos capas. La producción de nuevos ácidos grasos tales como los PUFAs genera por lo tanto nuevas propiedades de funciones de la membrana dentro de una célula o un organismo.

10 Las membranas celulares prestan una gran cantidad de funciones en una célula. En primer lugar, una membrana delimita el contenido de una célula del ambiente, por lo cual imparte integridad a la célula. Las membranas también pueden actuar como barreras contra el influjo de compuestos peligrosos y no deseados o bien contra la emanación de compuestos deseados.

15 Para descripciones más detalladas e implicaciones de las membranas y los mecanismos involucrados, véase Bamberg, E., y colaboradores (1993) Charge transport of ion pumps on lipid bilayer membranes, *Q. Rev. Biophys.* 26:1-25; Gennis, R.B. (1989) Pores, Channels and Transporters, en: *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, pág. 270-322; y Nikaido, H., y Saier, H. (1992) Transport proteins in bacteria: common themes in their design, *Science* 258:936-942, y las citas bibliográficas contenidas en cada una de estas referencias.

20 La síntesis de lípidos puede dividirse en dos partes: la síntesis de ácidos grasos y su enlace al sn-glicerina-3-fosfato así como la adición o modificación de un grupo de cabeza polar. Los lípidos habituales utilizados en las membranas comprenden fosfolípidos, glicolípidos, esfingolípidos y fosfoglicéridos. La síntesis del ácido graso comienza con la conversión del acetil CoA ya sea en malonil-CoA mediante acetil - CoA carboxilasa o en acetil-ACP mediante acetil transacilasa. Después de una reacción de condensación, estas dos moléculas productoras conjuntamente forman acetoacetil-ACP, la cual se convierte a través de una serie de reacciones de condensación, reducción y deshidratación de modo que se obtiene una molécula de ácido graso con la longitud de cadena deseada. La producción de los ácidos grasos insaturados de estas moléculas se cataliza mediante desaturasas específicas, ya sea aeróbicamente por medio de oxígeno molecular o anaeróbicamente (con respecto a la síntesis del ácido graso en microorganismos, véase F.C. Neidhardt y colaboradores (1996) *E. coli* y salmonela. ASM Press: Washington, D.C., pág. 612-636 y las citas bibliográficas contenidas allí; Lengeler y colaboradores (Hrsgb.) (1999) *Biology of Prokaryotes*. Thieme: Stuttgart, New York, y las citas bibliográficas contenidas allí, así como Magnuson, K., y colaboradores (1993) *Microbiological Reviews* 57:522-542 y las citas bibliográficas contenidas allí).

35 Precusores para la biosíntesis de PUFA son, por ejemplo, ácido palmitoleico, ácido linólico y ácido linolénico. Estos ácidos grasos con C₁₆ y/o C₁₈ deben alargarse a C₂₀ y C₂₂ para que se obtengan ácidos grasos del tipo de cadena eicosa y docosa. Este alargamiento ventajosamente se efectúa con los ácidos nucleicos de conformidad con la invención o con las proteínas codificadas por estos ácidos nucleicos. Con ayuda de varias desaturasas tales como las enzimas que tienen Δ-6-desaturasa, actividad de Δ-5- y Δ-4-desaturasa, pueden obtenerse ácido araquidónico, ácido eicosapentenoico y ácido docosahexaenoico así como otros PUFAs diferentes de cadena larga, extraerse y usarse para diferentes propósitos en aplicaciones para alimentos, forraje, cosmetología o farmacología.

45 Para la preparación de PUFAs de cadena larga, los ácidos grasos poliinsaturados con C₁₆ y/o C₁₈ deben alargarse en al menos dos átomos de carbono mediante la actividad enzimática de una elongasa. Las secuencias de ácido nucleico según la invención codifican primero las elongasas vegetales que pueden alargar ácidos grasos de C₁₆ y/o C₁₈ con al menos dos enlaces dobles en el ácido graso en al menos dos átomos de carbono. Después de una ronda de alargamiento, esta actividad enzimática conduce a ácidos grasos con C₁₆ o C₁₈ y después de dos, tres y cuatro o cinco rondas de alargamiento a ácidos grasos con C₂₂, C₂₄ ó C₂₆. Con las elongasas según la invención también pueden sintetizarse PUFAs más largos. La actividad de las elongasas según la invención conduce preferiblemente a ácidos grasos con C₂₀ o C₂₂ con al menos dos enlaces dobles en la molécula de ácido graso, preferiblemente con tres o cuatro enlaces dobles, en especial preferiblemente tres enlaces dobles en la molécula de ácido graso. Después del alargamiento con las enzimas de conformidad con la invención, pueden llevarse a cabo pasos adicionales de desaturación. Por consiguiente, es posible que los productos de las actividades de elongasa y de la desaturación adicional conduzcan a PUFAs preferidos con un grado superior de desaturación, tal como el ácido docosadienoico, ácido araquidónico, ácido ω6-eicosatriendihomo-γ-linolénico, ácido eicosapentenoico, ácido ω3-eicosatrienoico, ácido ω3-eicosatetraenoico, ácido docosapentaenoico o ácido docosahexaenoico. Sustratos de esta actividad enzimática son, por ejemplo, ácido taxólico, ácido 6,9-octadecadienoico, ácido linoleico, ácido γ-linolénico, ácido pinolénico, ácido α-linolénico o ácido estearidónico. Sustratos preferidos son ácido linólico, ácido γ-linolénico y/o ácido α-linolénico. Los ácidos de C₁₆ y/o C₁₈ con al menos dos enlaces dobles en el ácido graso pueden alargarse mediante la actividad enzimática en forma del ácido graso libre o en forma del éster, como fosfolípidos, glicolípidos, esfingolípidos, fosfoglicéridos, monoacilglicerina, diacilglicerina o triacilglicerina.

65 Además, los ácidos grasos deben transportarse a continuación a diversas modificaciones e incorporarse en el lípido para el almacenamiento de triacilglicerina. Otro paso importante en la síntesis de lípidos es la transferencia de ácidos grasos a los grupos de cabeza polar, por ejemplo mediante glicerina ácido graso aciltransferasa (véase Frentzen, 1998, *Lipid*, 100(4-5):161-166).

Además, la expresión de los ácidos nucleicos según la invención en los diversos organismos huéspedes no solo provoca un cambio en la composición de los lípidos de la membrana, sino que también modifica la composición de

todos los compuestos en la célula huésped, los cuales comprenden ácidos grasos insaturados, frente a las células huéspedes originales que no contienen los ácidos nucleicos o no los contienen en estas cantidades. Estas modificaciones son más pronunciadas en los organismos huéspedes, por ejemplo en células vegetales que no comprenden de manera natural las proteínas o enzimas codificadas por los ácidos nucleicos.

5

Véanse publicaciones sobre la biosíntesis del ácido graso en plantas, la desaturación, el metabolismo de lípidos y el transporte de las membrana de compuestos grasos, la beta-oxidación, la modificación de ácidos grasos y cofactores, el almacenamiento y ensamble de triacilglicerina, incluyendo las referencias bibliográficas allí citadas, en los siguientes artículos: Kinney, 1997, Genetic Engineering, editor: JK Setlow, 19:149-166; Ohlrogge y Browse, 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, 18:111-13; Gerhardt, 1992, Prog. Lipid R. 31:397-417; Gühnemann-Schäfer & Kindl, 1995, Biochim. Biophys Acta 1256:181-186; Kunau y colaboradores, 1995, Prog. Lipid Res. 34:267-342; Stymne y colaboradores, 1993, en: 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.

10

Vitaminas, cofactores y nutraceuticos, como PUFAs, comprenden un grupo de moléculas, las cuales los animales superiores ya no pueden sintetizar por sí mismos y por lo tanto tienen que asimilarlas, o las cuales los animales superiores ya no pueden sintetizar en cantidad suficiente por sí mismos y por lo tanto deben asimilarlas de manera adicional. La biosíntesis de estas moléculas en organismos que las pueden producir, como en bacterias, ha sido caracterizada en términos generales (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, pág. 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, llevada a cabo el 1.-3. Sept. de 1994 en Penang, Malasia, AOCS Press, Champaign, IL X, 374 pp).

20

25

La moléculas arriba mencionadas son moléculas o bien biológicamente activas por sí mismas, o bien precursoras de sustancias biológicamente activas que sirven ya sea como portadoras de electrones o como productos intermedios en una cantidad de rutas metabólicas. Además de su valor nutricional, estos compuestos también tienen un valor industrial significativo como colorantes, antioxidantes y catalizadores u otros auxiliares de procesamiento (véase una sinopsis de la estructura, actividad y aplicación industrial de estos compuestos, por ejemplo, en Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", vol. A27, pág. 443-613, VCH: Weinheim, 1996). Los ácidos grasos poliinsaturados tienen una variedad de funciones y efectos promotores de la salud, por ejemplo en el caso de un padecimiento coronario cardiaco, mecanismos de inflamación, nutrición de niños, etc. Véanse publicaciones y referencias bibliográficas, incluyendo referencias bibliográficas allí citadas, en: Simopoulos, 1999, Am. J. Clin. Nutr. 70 (3. Suppl.): 560-569, Takahata y colaboradores, Biosci. Biotechnol. Biochem. 1998, 62(11):2079-2085, Willich y Winther, 1995, Deutsche Medizinische Wochenschrift (Revista alemana semanal de medicina) 120(7):229 y siguientes. Los mismos son además materiales de partida importantes para la síntesis de los compuestos que controlan los procesos biológicos importantes dentro del organismo. Los mismos se utilizan, por ejemplo, en una variedad de alimentos dietéticos o en medicamentos.

30

35

40

II. Elementos y métodos de la invención

La presente invención se basa, al menos en parte, en las moléculas de ácido nucleico PSE y moléculas de proteína PSE, las cuales ejercen un efecto en la producción de membranas celulares o ácidos grasos como en *Physcomitrella patens*, *Ceratodon purpureus* y/o *Phytophthora infestans* y afectan, por ejemplo, el movimiento de moléculas por estas membranas. En una forma de realización, las moléculas PSE participan en el metabolismo de los compuestos requeridos para la síntesis de membranas celulares y/o ácidos grasos en organismos tales como microorganismos y plantas o afectan indirectamente el transporte de moléculas a través de estas membranas. En una forma preferida de realización, la actividad de las moléculas PSE de acuerdo con la invención, la cual regula la producción de los componentes de membrana y el transporte de membrana, tienen un efecto en la producción por parte de este organismo de la sustancia química fina deseada. En una forma particularmente preferida, la actividad de las moléculas PSE se modula de tal manera que el rendimiento, producción y/o eficiencia en la producción de las rutas metabólicas de los microorganismos o plantas que regulan los PSEs se modulan y se modifica la eficiencia del transporte de compuestos a través de las membranas, lo cual modula directa o indirectamente el rendimiento, producción y/o eficiencia de la producción de una sustancia química fina deseada por parte de microorganismos o plantas.

45

50

55

El término PSE o polipéptido PSE comprende proteínas que participan en el metabolismo de los compuestos requeridos para la síntesis de membranas celulares en organismos tales como los microorganismos y plantas o en el transporte de moléculas a través de estas membranas. Ejemplos de PSEs se divulgan en la SEQ ID NO:1 o sus homólogos, derivados o análogos. Los términos PSE o secuencia(s) de ácidos nucleicos PSE comprenden secuencias de ácidos nucleicos que codifican un PSE y en las que una parte es una región codificante y también sectores correspondientes de una secuencia no traducida (traducida) 5' y 3'. Ejemplos de genes PSE son los representados en SEQ ID NO:1. Los términos producción o productividad son conocidos en el campo técnico y comprenden la concentración del producto de fermentación (por ejemplo, de la sustancia química fina deseada) la cual se forma dentro de un determinado lapso 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 que es necesario para lograr una cantidad determinada de producción (por ejemplo cuanto tiempo requiere la célula en establecer una velocidad particular de capacidad de

60

65

producción de una sustancia química fina). El concepto rendimiento o rendimiento del producto/carbono es conocido en el campo técnico y comprende la eficiencia con la cual la fuente de carbono se convierte en el producto (es decir, la sustancia química fina). Esto se expresa usualmente como, por ejemplo, kg de producto por kg de fuente de carbono. Aumentando el rendimiento o producción del compuesto se incrementa la cantidad de las moléculas obtenidas o de las moléculas adecuadas de este compuesto obtenidas en una cantidad determinada de cultivo durante un lapso de tiempo establecido. Los términos biosíntesis o ruta sintética son conocidos en la técnica y comprenden la síntesis de un compuesto, preferiblemente de un compuesto orgánico, mediante una célula a partir de compuestos intermedios, por ejemplo en un proceso de pasos múltiples y/o un proceso fuertemente regulado. Los términos catabolismo o ruta catabólica son conocidos en el campo técnico y comprenden la escisión de un compuesto, preferiblemente de un compuesto orgánico, mediante una célula en los catabolitos (en general, moléculas más pequeñas o menos complejas), por ejemplo en un proceso de pasos múltiples y/o en un proceso 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 rutas biosintéticas, de modificación y catabólicas de este compuesto en la célula, las cuales atañen a este compuesto.

En una forma de realización, las moléculas PSE pueden modular la producción de una molécula deseada, tal como una sustancia química fina, en un microorganismo o en plantas. Existe una serie de mecanismos mediante los cuales la modificación de un PSR según la invención puede afectar directamente el rendimiento, la producción y/o la eficiencia de la producción de una sustancia química fina a partir de una cepa de un microorganismo o la cepa de una planta que comprenden esta proteína modificada. Puede elevarse el número o la actividad de PSEs que participan en el transporte de moléculas de sustancias químicas finas dentro o fuera de la célula de modo que se transporten cantidades mayores de estos compuestos a través de las membranas, a partir de las cuales pueden obtenerse y convertirse unas en otras más fácilmente. Ácidos grasos, triacilglicerinas y/o lípidos por sí mismos también son sustancias químicas finas deseables; por medio de la optimización de la actividad o el incremento de la cantidad de uno o varios PSEs que participan en la biosíntesis de estos compuestos, o mediante la interferencia de la actividad de uno o varios PSEs que intervienen en el catabolismo de estos compuestos, puede ser posible elevar el rendimiento, producción y/o eficiencia de la producción de moléculas de ácido graso y lípidos a partir de organismos tales como microorganismos o plantas.

La mutagénesis de los ácidos nucleicos o genes PSE también puede dar lugar a PSEs con actividades modificadas que afectan indirectamente la producción de una o más sustancias químicas finas deseadas a partir de los microorganismos o plantas. Por ejemplo, los PSEs según la invención que intervienen en la exportación de productos residuales pueden exhibir una mayor cantidad o actividad superior de modo que los residuos metabólicos normales de la célula (cuya cantidad es elevada posiblemente debido a la sobre producción de las sustancias químicas finas deseadas) se exportan de manera eficiente antes de que los mismos dañen las moléculas en la célula (lo cual podría reducir la capacidad vital de la célula) o interferir con las rutas biosintéticas de las sustancias químicas finas (lo cual podría reducir el rendimiento, producción o eficiencia en la producción de una sustancia química fina deseada). Las cantidades intracelulares relativamente grande de la sustancia química fina deseada, por sí mismas, también pueden ser tóxicas para la célula, de modo que incrementando la actividad o el número de transportadores capaces de exportar estos compuestos de la célula puede incrementarse la capacidad vital de la célula en el cultivo, lo cual a su vez conduce a un número superior de células en el cultivo, las cuales producen la sustancia química fina deseada. Los PSEs también pueden manipularse de tal modo que se producen cantidades correspondientes de diferentes moléculas de lípido y ácidos grasos. Esto puede tener un efecto considerable en la composición de lípido de la membrana celular. Puesto que cada tipo de lípido tiene diferentes propiedades físicas, una modificación de la composición de lípidos de una membrana puede modificar significativamente la fluidez de la membrana. Las modificaciones de la fluidez de la membrana puede afectar el transporte de moléculas a través de la membrana y la integridad de las células, cada una de las cuales tiene un efecto sustancial en la producción de sustancias químicas finas de microorganismos y plantas en un cultivo de fermentación a gran escala. Las membranas vegetales confieren propiedades específicas tales como tolerancia frente a calor, frío, sal, sequedad, así como tolerancia frente a patógenos tales como bacterias y hongos. Por lo tanto, la modulación de los componentes de membrana puede tener un efecto crítico en la capacidad de las plantas para sobrevivir bajo los parámetros de tensión antes mencionados. Esto puede efectuarse mediante los cambios en cascadas de señales o directamente por la composición modificada de membrana (véase, por ejemplo: Chapman, 1998, Trends in Plant Science, 3(11):419-426) y cascadas de señales (véase Wang 1999, Plant Physiology, 120:645-651) o afectar la tolerancia al frío tal como se divulga en WO 95/18222.

Las secuencias de ácidos nucleico aislado empleadas en el método según la invención están contenidas en el genoma de una cepa de *Phytophthora infestans*, la cual se encuentra disponible en colecciones tales como ATCC o DSM, por ejemplo. La secuencia de nucleótidos del ADNc aislado de *Physcomitrella patens* y las secuencias de aminoácidos derivadas de las PSEs de la *Phytophthora infestans* se muestran en las SEQ ID NO:1 ó 2, respectivamente. Se realizan análisis por ordenador que clasifican y/o identifican estas secuencias de nucleótidos como secuencias que codifican las proteínas que participan en el metabolismo de los componentes de la membrana celular o que intervienen en el transporte de compuestos a través de las membranas celulares o ácidos grasos. Una EST con el No. De acceso de la base de datos 08ck19b07 del inventor es parte de la secuencia mostrada en la SEQ ID NO:1. Entre tanto, estas ESTs se renombran, lo cual lleva el nombre enmendado: pp001019019f führte. De manera análoga, el polipéptido parcial se nombró pp001019019f. El fragmento inserto completo de la EST pp001019019f se secuenció y dio como resultado la SEQ ID NO:1, la cual muestra la secuencia completa de pp001019019f. Esta contiene un clon completo, activo funcionalmente del cual se demostró después de la expresión específica en una levadura, que tenía la especificidad de sustrato deseada, tal como se muestra en la sección de ejemplo.

ES 2 337 564 T3

La presente invención también se refiere a métodos en los que se emplean proteínas con una secuencia de aminoácidos que es esencialmente homóloga (=idéntica) a una secuencia de aminoácidos de la SEQ ID NO:2. Tal como se usa aquí, una proteína con una secuencia de aminoácidos que es esencialmente homóloga a una secuencia seleccionada de aminoácidos, tiene al menos alrededor de 50% de homología con la secuencia de aminoácidos seleccionada, por ejemplo la secuencia completa seleccionada de aminoácidos. Una proteína con una secuencia de aminoácidos que es esencialmente homóloga con una secuencia de aminoácidos seleccionada, también puede tener al menos aproximadamente de 50 a 60%, preferiblemente al menos alrededor de 60 a 70% y más preferiblemente al menos alrededor de 70 a 80%, 80 a 90% ó 90 a 95% y lo más preferible al menos alrededor de 96%, 97%, 98%, 99% o más de homología con una secuencia seleccionada de aminoácidos.

La PSE o la parte o fragmento biológicamente biológicamente activas de la misma pueden participar en el metabolismo de los compuestos requeridos para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas o tiene una o más de las actividades requeridas para el alargamiento de los PUFAs con C₁₈ de modo que se obtengan PUFAs con C₂₂ o C₂₄ y PUFAs relacionados.

Diversos aspectos de la invención se describen con mayor detalle en los siguientes sub-secciones.

A. Moléculas de ácido nucleico aisladas

En el método según la invención se emplea un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica un polipéptido que alarga los ácidos grasos de C16 y/o C18 con al menos dos enlaces dobles en el ácido graso, seleccionado del grupo que se compone de

a) una secuencia de ácidos nucleicos representada en SEQ ID NO:1,

b) una secuencia de ácidos nucleicos que, de conformidad con la degeneración del código genético, proviene de la secuencia representada en SEQ ID NO:1,

c) Derivados de la secuencia representada en SEQ ID NO:1, la cual codifica los polipéptidos con al menos 50% de homología con la secuencia que codifica las secuencias de aminoácidos en la SEQ ID NO:2; la secuencia actúa como elongasa de C₁₆ o C₁₈, tal como se describe más exactamente en otro sitio.

El ácido nucleico arriba mencionado proviene de organismos tales como ciliados, hongos, algas o dinoflagelados, los cuales son capaces de sintetizar los PUFAs preferiblemente de plantas u hongos, en particular preferible del género *Phytophthora* y lo más preferible de *Phytophthora infestans*.

El término "molécula de ácido nucleico", tal como aquí se usa debe comprender moléculas de ADN bicatenario o monocatenario (de doble hebra o de una sola hebra) (por ejemplo ADNc o ADN genómico) y/o moléculas de ARN (por ejemplo RNAm) así como análogos de ADN o ARN que se generan por medio de análogos de nucleótidos, o híbridos de ADN/ARN. Este término comprende adicionalmente la secuencia no traslatada (traducida) en el extremo 3' y 5' de la región de codificación del gen: al menos aproximadamente 100 nucleótidos de la secuencia corriente arriba del extremo 5' de la región de codificación y al menos aproximadamente 20 nucleótidos de la secuencia corriente abajo del extremo 3' de la región de codificación del gen. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria (de una o de dos hebras), pero preferiblemente es ADN bicatenario. Una molécula "aislada" de ácido nucleico se separa de otras moléculas de ácido nucleico que están presentes en la fuente natural del ácido nucleico. Un ácido nucleico "aislado" preferiblemente no tiene secuencias que flanquean por naturaleza el ácido nucleico en el ADN genómico del organismo a partir del cual proviene el ácido nucleico (por ejemplo, secuencias localizadas en los extremos 5' y 3' del ácido nucleico). En diversas formas de realización, la molécula aislada de ácido nucleico PSE puede contener, por ejemplo, secuencias de nucleótidos de aproximadamente menos de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ó 0,1 kb, las cuales flanquean por naturaleza la molécula de ácido nucleico en el ADN genómico de la célula a partir de la cual proviene el ácido nucleico (por ejemplo, una célula de *Physcomitrella patens*). Una molécula "aislada" de ácido nucleico, tal como una molécula de ADNc, puede además estar esencialmente libre de otro material celular o medio de cultivo si la misma se genera mediante técnicas recombinantes, o libre de precursores químicos u otras sustancias químicas si la misma se sintetiza químicamente.

Una molécula de ácido nucleico nombrada arriba, por ejemplo una molécula de ácido nucleico con una secuencia de nucleótidos de la SEQ ID NO:1 o una parte de la misma, puede aislarse utilizando técnicas estándar de biología molecular y la información de secuencia proporcionada en la misma. Por ejemplo, puede aislarse un ADNc de *Phytophthora infestans* a partir de un banco (genoteca) de *P. infestans*, utilizando la SEQ ID NO:1 completa o parte de la misma como sonda de hibridación y técnicas de hibridación estándar (como, por ejemplo, se describe en Sambrook y colaboradores, Molecular Cloning: A Laboratory Manual. 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Además, una molécula de ácido nucleico que comprende una secuencia completa de la SEQ ID NO:1 o una parte de la misma, puede aislarse mediante una reacción en cadena de polimerasa, y se usan cebadores de oligonucleótidos que se han generado sobre la base de esta secuencia o de partes de la misma, en particular de las regiones alrededor de los motivos His-Box, véase Shanklin y colaboradores (1994) Biochemistry 33, 12787-12794, (por ejemplo, puede aislarse una molécula de ácido nucleico que comprende la secuencia completa de la SEQ ID NO:1 o una parte de la misma mediante reacción en cadena de polimerasa usando cebadores de oligonucleótidos que se han generado sobre la base de esta secuencia igual de la SEQ ID NO:1). Por

ES 2 337 564 T3

ejemplo, el ARNm puede aislarse de las células de oomicetes (por ejemplo, mediante el método de extracción con guanidinio-tiocianato de Chirgwin y colaboradores (1979) *Biochemistry* 18:5294-5299) y ADNc puede prepararse por medio de transcriptasa inversa (por ejemplo, transcriptasa inversa de Moloney-MLV, disponible a partir de Gibco/BRL, Bethesda, MD, o transcriptasa inversa AMV disponible de Seikagaku America, Inc., St. Petersburg, FL). Cebadores sintéticos de oligonucleótidos para la amplificación por medio de la reacción en cadena de polimeras puede generarse con base a una de las secuencias de nucleótidos mostradas en la SEQ ID NO:1. Un ácido nucleico según la invención puede amplificarse usando ADNc o, de manera alternativa, usando el ADN genómico como matriz y cebadores de oligonucleótidos adecuados, según las técnicas estándar de amplificación con PCR. El ácido nucleico amplificado de esta manera puede clonarse en un vector adecuado y caracterizarse por medio de un análisis de secuencia de ADN. Los oligonucleótidos que corresponden a una secuencia de nucleótidos PSE pueden generarse mediante métodos de síntesis estándar, por ejemplo con un sintetizador de ADN automático.

El ADNc mostrado en SEQ ID NO:1 comprende secuencias que codifican los PSEs (es decir, la “región codificante”) así como las secuencias 5’ no traducidas y secuencias 3’ no traducidas. De manera alterna, la molécula de ácido nucleico puede comprender solo la región de codificación de una de las secuencias en la SEQ ID NO:1 ó puede comprender los fragmentos genómicos, completos, aislados del ADN genómico.

En otra forma preferida de realización, una molécula aislada de ácido nucleico comprende una molécula de ácido nucleico que es un complemento de una de las secuencias de nucleótidos mostradas en la SEQ ID NO:1 o una parte de la misma. Una molécula de ácido nucleico que es complementaria a una de las secuencias de nucleótidos mostradas en SEQ ID NO:1 es una que es suficientemente complementaria para una de las secuencias de nucleótidos mostradas en SEQ ID NO:1 como para que pueda hibridizarse con una de las secuencias indicadas en SEQ ID NO:1, por lo cual se genera un dúplex estable.

Homólogos de la nueva secuencia de ácido nucleico elongas con la secuencia SEQ ID NO:1 significa, por ejemplo, variantes alélicas con al menos alrededor de 50 a 60%, preferiblemente al menos alrededor de 60 a 70%, más preferible al menos alrededor de 70 a 80%, 80 a 90% ó 90 a 95% y aún más preferible al menos alrededor de 95%, 96%, 97%, 98%, 99% o más homología con una de las secuencias mostradas en SEQ ID NO:1 o sus homólogos, derivados o análogos o partes de las mismas; homología en el sentido de la invención significa identidad. Otra forma preferida de realización, una molécula de ácido nucleico aislada según la invención comprende una secuencia de nucleótidos que se hibridiza con una de las secuencias de nucleótidos mostrada en SEQ ID NO:1 o una parte de la misma, por ejemplo bajo condiciones severas. Las variantes alélicas comprenden, en particular, variantes funcionales que pueden obtenerse mediante la eliminación, inserción o sustitución de nucleótidos de/en la secuencia representada en SEQ ID NO:1; pero el propósito es que la actividad enzimática de las proteínas sintetizadas provenientes de la misma se mantenga de manera ventajosa para la inserción de uno o varios genes. Proteínas que aún poseen la actividad enzimática de la elongasa significan proteínas con al menos 10%, preferiblemente 20%, particularmente preferible 30%, muy particularmente preferible 40% de la actividad enzimática original, comparada con la proteína codificada por SEQ ID NO: 2.

Homólogos de la SEQ ID NO:1 también significan, por ejemplo, homólogos bacterianos, fúngicos y vegetales, secuencias abreviadas, ADN o ARN monocatenarios de la secuencia de ADN codificante y no codificante.

Homólogos de la SEQ ID NO:1 significan también derivados como, por ejemplo, variantes de promotores. Los promotores corriente arriba de las secuencias de nucleótidos indicadas pueden modificarse por uno o más intercambios de nucleótidos, por inserción(ones) y/o delección(ones), aunque sin esto la funcionalidad o la actividad de los promotores se vea interferida. Adicionalmente es posible que la actividad de los promotores se incremente mediante la modificación de su secuencia o que los mismos sean reemplazados completamente por promotores más activos, incluso de organismos heterólogos.

Las moléculas de ácido nucleico arriba nombradas codifican una proteína o una parte de la misma, la cual abarca una secuencia de aminoácidos que es suficientemente homóloga a una secuencia de SEQ ID NO:2, que la proteína o la parte de la misma mantiene la capacidad de participar en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas o en la síntesis de ácidos grasos. Tal como aquí se usa, el término “suficientemente homólogo” se refiere a proteínas o a partes de las mismas cuyas secuencias de aminoácidos tienen un número mínimo de residuos de aminoácidos idénticos o equivalentes (por ejemplo, un residuo de aminoácido con una cadena lateral similar, tal como un residuo de aminoácido en una de las secuencias de la SEQ ID NO:2) a una secuencia de aminoácidos de la SEQ ID NO:2, de modo que la proteína o la parte de la misma puede intervenir en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas. Tal como aquí se describe, los componentes de proteínas de estas rutas metabólicas para los componentes de la membrana o sistemas de transporte de la membrana pueden desempeñar un papel en la producción y secreción de una o más sustancias químicas finas. Los ejemplos de estas actividades también se describen aquí. Por consiguiente, la “función de una PSE” contribuye ya sea directa o indirectamente al rendimiento, producción y/o eficiencia en la producción de una o más sustancias químicas finas. Los ejemplos de especificidades de sustrato de PSE de la actividad catalítica se establecen en la tabla I.

En otra forma de realización del método según la invención, los derivados de la molécula de ácido nucleico codifican proteínas con al menos alrededor de 50 a 60%, preferiblemente al menos alrededor de 60 a 70% y muy

ES 2 337 564 T3

preferiblemente al menos alrededor de 70 a 80%, 80 a 90%, 90 a 95% y lo más preferible al menos alrededor de 96%, 97%, 98%, 99% o más homología con una secuencia completa de aminoácido de la SEQ ID NO:2. 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 y colaboradores, CABIOS, 5, 1989:151-153).

Partes de proteínas que se codifican mediante las moléculas de ácido nucleico de PSE son preferiblemente partes biológicamente activas de uno de los PSEs. Tal como aquí se ha utilizado, el término “parte biológicamente activa de un PSE” significa un segmento, por ejemplo un dominio/porción de un PSE que puede participar en el metabolismo de los compuestos necesarios para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas o que interviene en la síntesis de ácidos grasos o que tiene una actividad establecida en la tabla 1. Puede llevarse a cabo un ensayo de la actividad enzimática con el fin de determinar si un PSE o una parte biológicamente activa del mismo pueden participar en el metabolismo de los compuestos requeridos para la síntesis de membranas celulares en microorganismos o plantas o en el transporte de moléculas a través de estas membranas. Este método de prueba, tal como se describe en detalle en el ejemplo 8 de la sección de ejemplos, son corrientes para el experto en la materia.

Fragmentos adicionales de ácido nucleico que codifican segmentos activos biológicamente activos de un PSE, pueden generarse aislando una parte de la secuencia representada en SEQ ID NO:2 dargestellten Sequenz, expresando el segmento codificado del PSE o del péptido (por ejemplo mediante expresión recombinante *in vitro*) y determinando la actividad de la parte codificada del PSE o del péptido.

La invención comprende además la aplicación de moléculas de ácido nucleico que se distinguen de una de las secuencias de nucleótidos mostradas en SEQ ID NO: 1 (y partes de la misma) debido al código genético degenerado y de esta manera codifican el PSE igual a aquel que se codifica por las secuencias de nucleótidos mostradas en la SEQ ID NO:1. En otra forma de realización una molécula de ácido nucleico aislada tiene una secuencia de nucleótidos que codifica una proteína con una secuencia de aminoácidos mostrada en SEQ ID NO:2. En otra forma de realización la molécula de ácido nucleico codifica una proteína de longitud completa de *Phytophthora infestans* que es esencialmente homóloga a una secuencia de aminoácidos de SEQ ID NO:2 (que se codifica por una estructura de lectura abierta mostrada en SEQ ID NO:1).

Adicionalmente a las secuencias de nucleótidos PSE de *Phytophthora infestans*, mostradas en SEQ ID NO:1, el experto en la materia reconoce que dentro de una población (por ejemplo, de la población de *Phytophthora infestans*) pueden existir polimorfismos secuenciales de ADN que conducen a modificaciones en las secuencias de aminoácidos de los PSEs. Estos polimorfismos genéticos en el gen PSE pueden existir entre individuos de una población debido a variación natural. Tal como se utiliza aquí, el término “gen” y “gen recombinante” se refiere a moléculas de ácidos nucleicos con un marco abierto de lectura que codifica un PSE, preferiblemente un PSE de *Phytophthora infestans*. Estas variantes naturales provocan una varianza del 1 al 5% en la secuencia de nucleótidos del gen PSE. Todas estas variaciones del nucleótido y polimorfismos resultantes del aminoácido en el PSE que son resultado de variación natural y que no modifican la actividad funcional de PSEs, deben estar contenidos en el alcance del método según la invención.

Las moléculas de ácido nucleico que corresponden a las variantes naturales y no a los homólogos, derivados y análogos de *Phytophthora infestans* del ADNc del PSE de la *Phytophthora infestans* pueden aislarse según técnicas estándar de hibridación en condiciones estrictas de hibridación con base en su homología con el ácido nucleico -PSE- DE *Phytophthora infestans* aquí divulgado usando el ADNc de *Phytophthora infestans* o una parte del mismos como sondas de hibridación. En otra forma de realización una molécula de ácido nucleico aislada tiene una longitud de al menos 15 nucleótidos y se hibridiza en condiciones estrictas con la molécula de ácido nucleico que comprende una secuencia de nucleótidos de la SEQ ID NO:1. En otras formas de realización el ácido nucleico tiene una longitud de al menos 25, 50, 100, 250 o más nucleótidos. La expresión “se hibridiza en condiciones estrictas”, tal como se usa aquí, debe describir condiciones de hibridación y de lavado en las que las secuencias de nucleótidos que son homólogos en al menos 60% entre sí permanecen usualmente hibridizadas entre sí. Las condiciones son preferiblemente de tal tipo que secuencias que sean homólogas entre sí en al menos alrededor de 65%, más preferible en al menos alrededor de 70% y aún más preferible en al menos alrededor de 75% o más, habitualmente permanecen hibridizadas entre sí. Estas condiciones estrictas son conocidas por el experto en la materia y pueden encontrarse en Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6. Un ejemplo preferido, no limitante, de condiciones estrictas de hibridación son hibridaciones en 6 x cloruro de sodio/citrato de sodio (sodium chloride/sodiumcitrate = SSC) a alrededor de 45°C, seguidas de uno o varios pasos de lavado en 0,2 x SSC, 0,1% SDS a 50 hasta 65°C. El experto en la materia conoce que estas condiciones de hibridación difieren dependiendo del tipo de ácido nucleico y cuando están presentes solventes orgánicos, por ejemplo, difieren respecto de la temperatura y de la concentración del búfer. Por ejemplo, la temperatura difiere en “condiciones estándar de hibridación” 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). Si está presente un solvente orgánico en el búfer arriba mencionado, por ejemplo formamida al 50%, la temperatura en condiciones estándar es de alrededor de 42°C. Las condiciones de hibridación para los híbridos de ADN:ADN son preferiblemente, por ejemplo, de 0,1 x SSC y 20°C a 45°C, preferiblemente entre 30°C y 45°C. Preferiblemente, las condiciones de hibridación para híbridos ADN:ARN son, por ejemplo, de 0,1 x SSC y 30°C a 55°C, preferiblemente entre 45°C y 55°C. Las temperaturas de hibridación nombradas previamente se determinaron, por ejemplo, para un ácido nucleico con alrededor de 100 bp (= pares de bases) de longitud y el contenido de G + C en ausencia de formamida. El experto en la materia sabe cómo pueden determinarse las condiciones de hibridación requeridas por medio de los libros de enseñanza, tales como alguno de los mencionados anteriormente o de los siguientes libros de enseñanza: Sambrook y colaboradores, “Molecular

Cloning”, Cold Spring Harbor Laboratory, 1989; Hames und Higgins (editorial) 1985, “Nucleic Acids Hybridization: A Practical Approach”, IRL Press at Oxford University Press, Oxford; Brown (editorial) 1991, “Essential Molecular Biology: A Practical Approach”, IRL Press at Oxford University Press, Oxford. El experto en la materia conoce que el tipo y el tiempo del desarrollo de los resultados de hibridación tienen una influencia en el resultado de la hibridación. En experimentos sencillos está en capacidad de optimizar las condiciones de desarrollo de tal modo que la ya mencionada hibridación suministre resultados seguros y confiables.

Preferiblemente, una molécula de ácido nucleico aislada como la mencionada arriba, la cual se hibridiza en condiciones estrictas con una secuencia de la SEQ ID NO:1, corresponde a una molécula de ácido nucleico de procedencia natural. Tal como se usa aquí, una molécula de ácido nucleico “de procedencia natural” se refiere a una molécula de ARN o ADN con una secuencia de nucleótidos que existe en la naturaleza (que codifica una proteína natural, por ejemplo). En una forma de realización el ácido nucleico codifica un PSE de *Phytophthora infestans* de procedencia natural.

Adicionalmente a las variantes de procedencia natural de la secuencia de PSE que puede existir en la población, el experto en la materia reconoce además que los cambios por medio de mutación también pueden introducirse en una secuencia de nucleótidos de la SEQ ID NO:1, la cual conduce a cambios en la secuencia de aminoácidos del PSE codificado sin afectar adversamente la funcionalidad de la proteína PSE. Por ejemplo, pueden generarse sustituciones de nucleótidos que conducen a las sustituciones de aminoácidos en residuos de aminoácidos “no esenciales” en una secuencia de la SEQ ID NO:1. Un residuo de aminoácido “no esencial” es un residuo que puede alterarse en una secuencia de tipo silvestre de uno de los PSEs (SEQ ID NO:2) sin alterar la actividad del PSE, mientras que se necesita un residuo de aminoácido “esencial” para la actividad del PSE. Otros residuos de aminoácidos (por ejemplo, aquellos que no conservaron, o solamente se semiconservaron, en el dominio con la actividad del PSE) pueden, sin embargo, no ser esenciales para la actividad y por lo tanto pueden alterarse probablemente sin alterar la actividad del PSE.

En consecuencia, otro aspecto del método de la invención se refiere a moléculas de ácido nucleico que codifican PSEs que contienen residuos modificados de aminoácido que no son esenciales para la actividad PSE. Estos PSEs se diferencian en la secuencia de aminoácidos de una secuencia en SEQ ID NO:2 y aún así mantienen al menos una de las actividades PSE aquí descritas. La molécula aislada de ácido nucleico comprende en una forma de realización una secuencia de nucleótidos que codifica una proteína; la proteína comprende una secuencia de aminoácidos con al menos alrededor de 50% de homología con una secuencia de aminoácidos de la SEQ ID NO:2 e interviene en el metabolismo de compuestos necesarios para la síntesis de membranas celulares en *Phytophthora infestans* o en el transporte de moléculas a través de estas membranas o interviene en el metabolismo de los ácidos grasos o tienen una o más de las actividades listadas en la tabla I. La proteína codificada por la molécula de ácido nucleico preferiblemente tiene al menos alrededor de 50 a 60% de homología con una de las secuencias en SEQ ID NO:2, más preferible al menos alrededor de 60 a 70% de homología con una de las secuencias en SEQ ID NO:2, aún más preferible al menos alrededor de 70 a 80%, 80 a 90%, 90 a 95% de homología con una de las secuencias en SEQ ID NO:2 y lo más preferible al menos alrededor de 96%, 97%, 98% ó 99% de homología con una de las secuencias en SEQ ID NO:2.

Para determinar el porcentaje de homología de dos secuencias de aminoácidos (por ejemplo, una de las secuencias de la SEQ ID NO:2 y una forma mutada de la misma) o de dos ácidos nucleicos, con el propósito de la comparación óptima se escriben las secuencias una debajo de la otra (por ejemplo, pueden introducirse huecos o espacios en la secuencia de una proteína o de un ácido nucleico con el fin de generar un alineamiento óptimo con la otra proteína o el otro ácido nucleico). Luego se comparan los residuos de aminoácidos o nucleótidos en las posiciones correspondientes de aminoácidos o posiciones de nucleótidos. Si una posición en una secuencia (por ejemplo, una de las secuencias de la SEQ ID NO:2) se ocupa por el mismo residuo de aminoácido o el mismo nucleótido como la posición correspondiente en la otra secuencia (por ejemplo, una forma mutada de la secuencia seleccionada de SEQ ID NO:2), entonces las moléculas son homólogas en esta posición (es decir, “homología” de aminoácido o de ácido nucleico, tal como aquí se utiliza, corresponde a “identidad” de aminoácido o de ácido nucleico). El porcentaje de homología entre las dos secuencias es una función del número de posiciones idénticas que comparten la secuencia (es decir, el % de homología = número de posiciones idénticas/número de posiciones x 100).

Una molécula de ácido nucleico que codifica un PSE, el cual es homólogo a una secuencia de proteína de la SEQ ID NO:2, puede generarse introduciendo una o más sustituciones, adiciones o deleciones (eliminaciones) de nucleótidos en una secuencia de nucleótidos de la SEQ ID NO:1, de modo que se introduzcan una o más sustituciones, adiciones o eliminaciones de aminoácidos en la proteína codificada. Pueden introducirse mutaciones en una de las secuencias de la SEQ ID NO: 1 mediante técnicas estándar, tales como mutagénesis específica para un sitio (= site directed) y mutagénesis mediada por PCR. Preferiblemente se generan sustituciones de aminoácidos conservadoras en uno o más de los residuos de aminoácidos no esenciales pronosticados. En una “sustitución de aminoácidos conservadora” se intercambia el residuo de aminoácido por un residuo de aminoácido con una cadena lateral similar. Se han definido familias de residuos de aminoácidos con cadenas laterales similares en el campo técnico. 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, no cargadas (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 pronosticado en un PSE se intercambia preferiblemente por otro residuo de

aminoácido de la misma familia de cadenas laterales. Como una alternativa en otra forma de realización, pueden introducirse mutaciones al azar sobre la totalidad o parte de la secuencia que codifica el PSE mediante la mutagénesis por saturación, por ejemplo, y pueden seleccionarse los mutantes resultantes para la actividad del PSE descrita aquí con el fin de identificar los mutantes que retengan la actividad del PSE. Después de la mutagénesis de una de las secuencias de la SEQ ID NO:1 puede expresarse la proteína codificada recombinantemente y puede determinarse la actividad de la proteína, usando por ejemplo el ensayo aquí descrito (véase la sección de ejemplos).

B. Construcción génica

En otra forma de realización de los métodos de la invención se emplea una construcción génica que significa un ácido nucleico aislado que proviene de un vegetal o un hongo y codifica un polipéptido que alarga ácidos grasos de C₁₆ y/o C₁₈ con al menos dos enlaces dobles en el ácido graso en al menos dos átomos de carbono, o la secuencia genética de la SEQ ID NO:1, sus homólogos, derivados o análogos que se enlazan funcionalmente a una o más señales reguladoras, para incrementar ventajosamente la expresión genética. Los ejemplos de estas secuencias reguladoras son secuencias que se enlazan a inductores o represores y de esta manera regulan la expresión del ácido nucleico. Además de estas nuevas secuencias reguladoras, aún puede estar presente la regulación natural de estas secuencias antes de los propios genes estructurales y, si es adecuado, se modifica genéticamente de modo que se interrumpa la regulación natural y se mejore la expresión de los genes. Sin embargo, la construcción génica también puede tener una estructura simple, es decir que no se han insertado señales reguladoras adicionales antes de la secuencia SEQ ID NO:1 o sus homólogos y no se ha eliminado (deletado) el promotor natural con su regulación. En lugar de esto, la secuencia reguladora natural se ha mutado de tal manera que la regulación ya no tiene lugar y se refuerza la expresión génica. La construcción génica puede comprender además, ventajosamente, uno o más de las llamadas secuencias mejoradas que se enlazan funcionalmente al promotor y las cuales permiten incrementar la expresión de la secuencia de ácido nucleico. También es posible insertar adicionalmente secuencias ventajosas en el extremo 3' de las secuencias de ADN, por ejemplo elementos reguladores o terminadores. Los genes de elongasa pueden estar presentes en una o más copias en la construcción génica. Es ventajoso para la inserción de otros genes en organismos si se encuentran presentes otros genes en la construcción génica.

Existen secuencias de regulación ventajosas para el método nuevo, por ejemplo en promotores tales como el promotor *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIq*, T7, T5, T3, *gal*, *trc*, *ara*, SP6, I-PR o I-PL y se utilizan ventajosamente en las bacterias gran-negativas. Además, existen secuencias reguladoras ventajosas, por ejemplo, en los promotores Gram-positivo *amy* y SPO2, en los promotores de levaduras u hongos ADC1, MFa, AC, P-60, CYC1, GAPDH, TEF, *rp28*, ADH o en los promotores vegetales CaMV/35S [Franck y colaboradores, Cell 21 (1980) 285-294], PRP1 [Ward y colaboradores, 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, tales como los promotores descritos en EP-A-0 388 186 (inducibles con bencilsulfonamida), Plant J. 2, 1992:397-404 (Gatz y colaboradores, inducibles con tetraciclina), EP-A-0 335 528 (inducibles con ácido abscísico) o WO 93/21334 (inducibles con etanol o ciclohexenol). Otros promotores vegetales adecuados son el promotor de FBPasa citosólica o el promotor ST-LSI de la patata (Stockhaus y colaboradores, EMBO J. 8, 1989, 2445), el promotor de fosforibosilpirofosfatamido-transferasa de glicina max (No. de acceso de Genbank U87999) o el promotor de nódulos específicos descrito en EP-A-0 249 676. Los promotores especialmente ventajosos son los promotores que permiten una expresión en los tejidos que están involucrados en la biosíntesis de los ácidos grasos. Son especialmente ventajosos los promotores de semillas específicos, tales como el promotor *usp*, LEB4, faseolina o napina. Además, los promotores especialmente ventajosos son promotores de semillas específicos que pueden usarse para las monocotiledóneas o dicotiledóneas y se describen en US 5,608,152 (promotor napina de colza), WO 98/45461 (promotor faseolina de arábido), US 5,504,200 (promotor de faseolina de *Faseolus vulgaris*), WO 91/13980 (promotor Bce4 de brassica), por parte de Baeumlein y colaboradores, Plant J., 2, 2, 1992:233-239 (promotor LEB4 de leguminosas); estos promotores son adecuados para dicotiledóneas. Los siguientes promotores son adecuados, por ejemplo, para las monocotiledóneas: promotor *lpt-2* o *lpt-1* de cebada (WO 95/15389 y WO 95/23230), promotor hordeína y otros promotores adecuados se describen en WO 99/16890.

En teoría es posible usar todos los promotores naturales con sus secuencias reguladoras, tales como aquellos arriba para el método nuevo. También es posible y ventajoso usar adicionalmente promotores sintéticos.

Tal como se ha descrito arriba, la construcción génica también puede comprender genes adicionales que se introducirán en los organismos. Es posible y ventajoso introducir en los organismos huéspedes y expresar en los mismos genes reguladores 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 ruta biosintética. Estos genes pueden ser de origen heterólogo u homólogo. Los genes insertados pueden tener su propio promotor o bien pueden estar bajo el control del promotor de la secuencia SEQ ID NO:1 o sus homólogos.

Para expresar los demás genes que están presentes, la construcción génica comprende ventajosamente además secuencias reguladoras con terminales 3' y/o 5' para mejorar la expresión y estas se seleccionan para su expresión óptima como una función del organismo huésped elegido y del gen o de los genes.

Como se mencionó arriba, estas secuencias de regulación deben hacer posibles la expresión específica de los genes y la expresión de proteína. Esto puede significar, por ejemplo, según el organismo huésped el gen se expresa o sobreexpresa solo después de la inducción o que el mismo se expresa y/o sobreexpresa inmediatamente.

Además, las secuencias reguladoras o factores reguladores pueden tener preferiblemente un efecto ventajoso en la expresión de los genes que se han introducido, mejorándolos por consiguiente. De esta manera, es posible que los elementos reguladores se mejoren ventajosamente en el nivel de transcripción usando fuertes señales de transcripción, tales como promotores y/o intensificadores. Sin embargo, además también es posible fortalecer la traducción (translación), por ejemplo mediante el mejoramiento de la estabilidad del ARNm.

C. Vectores de expresión recombinantes y células huéspedes

Otro aspecto del método de la invención se refiere al empleo de vectores, preferiblemente vectores de expresión que comprenden un ácido nucleico que codifica un PSE (o parte del mismo). Tal como se usa aquí, el término “vector” se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al cual el mismo está unido. Un tipo de vector es un “plásmido”, el cual representa un bucle circular de ADN bicatenario al cual pueden ligarse segmentos adicionales de ADN. Un tipo de vector adicional es un vector viral y los segmentos adicionales de ADN pueden ligarse al genoma viral. Determinados vectores son capaces de una replicación autónoma en una célula huésped en la cual se han introducido (por ejemplo, vectores bacterianos con origen bacteriano de replicación y vectores episomales de mamífero). Otros vectores (por ejemplo, los vectores no episomales de mamífero) se integran en el genoma de una célula huésped durante la introducción en la célula huésped y se replican así conjuntamente con el genoma huésped. Además, determinados vectores pueden controlar la expresión de genes a los cuales los mismos se enlazan funcionalmente. Estos vectores se denominan aquí como “vectores de expresión”. Habitualmente, los vectores de expresión que son adecuados para las técnicas de ADN recombinante tienen la forma de plásmidos. En la presente descripción “plásmido” y “vector” pueden usarse de manera intercambiable puesto que el plásmido es la forma más frecuentemente usada de un vector. Sin embargo, la invención debe abarcar estas otras formas de vector de expresión como vectores virales (por ejemplo, retrovirus de replicación deficiente, adenovirus y virus adeno-relacionados), que ejercen funciones similares. Además, el término vector también debe abarcar otros vectores que son conocidos por el experto en la materia, como fagos, virus tales como SV40, CMV, baculovirus, adenovirus, transposons, elementos IS, fásmidos, fagémidos, cósmidos, ADN lineal o circular.

Los vectores de expresión recombinantes arriba mencionados comprenden un ácido nucleico o una construcción génica en una forma que sea adecuada para expresar el ácido nucleico en una célula huésped, lo que significa que los vectores de expresión recombinantes comprenden una o más secuencias reguladoras, seleccionadas con base en las células huéspedes que se utilizan para la expresión, las cuales se enlazan funcionalmente a la secuencia de ácido nucleico que se va a expresar. En un vector de expresión recombinante, “enlazado funcionalmente” significa que la secuencia de nucleótidos de interés está unida con la(s) secuencia(s) reguladora(s) de tal modo que es posible la expresión de la secuencia de nucleótidos y que las mismas están enlazadas entre sí de modo que ambas secuencias cumplan la función prevista que se ha atribuido a la secuencia (por ejemplo, en un sistema *in vitro* de transcripción/traducción o en una célula huésped, cuando el vector se introduce en la célula huésped). El término “secuencia de regulación” debe comprender promotores, enhancer (potenciadores) y otros elementos de control de la expresión (por ejemplo, señales de poliadenilación). Estas secuencias reguladoras se describen, por ejemplo, en Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), o véase: Gruber y Crosby, en: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, editor: Glick y Thompson, capítulo 7, 89-108, incluyendo las referencias bibliográficas allí citadas. Las secuencias reguladoras comprenden aquellas que controlan la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células huéspedes y aquellas que controlan la expresión directa de la secuencia de nucleótidos solo en ciertas células huéspedes en ciertas condiciones. El experto en la materia sabe que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped que se va a transformar, la extensión de la expresión de la proteína deseada, etc. Los vectores de expresión según la invención pueden introducirse en células huéspedes con el fin de producir proteínas o péptidos, inclusive proteínas de fusión o péptidos de fusión, los cuales se codifican mediante los ácidos nucleicos tal como se describe aquí (por ejemplo PSEs, formas mutantes de PSEs, proteínas de fusión, etc.).

Los vectores de expresión recombinantes pueden diseñarse para la expresión de PSEs en células procarióticas o eucarióticas. Por ejemplo, los genes PSE pueden expresarse en células bacterianas, tales como *C. glutamicum*, células de insectos (usando vectores de expresión de baculovirus), células de levaduras y otros hongos (véase Romanos, M.A., y colaboradores (1992) “Foreign gene expression in yeast: a review”, *Yeast* 8:423-488; van den Hondel, C.A.M.J.J., y colaboradores (1991) “Heterologous gene expression in filamentous fungi”, en: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, editores, páginas 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., y colaboradores, editores, páginas 1-28, Cambridge University Press: Cambridge), Algen (Falciatore y colaboradores, 1999, *Marine Biotechnology*, 1, 3: 239-251), ciliados de los tipos: Holotrichia, Peritrichia, Spirotrichia, Suctorina, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Pseudocohnilembus, Euplotes, Engelmaniella y Stylonychia, en particular del género *Stylonychia lemnae*, con vectores según un método 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, capítulo 6/7, pág. 71-119 (1993); F.F. White, B. Jenes y colaboradores, *Techniques for Gene Transfer*, en: *Transgenic Plants*, vol. 1, Engineering and Utilization, editores: Kung y R. Wu, Academic Press (1993), 128-43; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225 (y las referencias bibliográficas allí citadas)) o células de mamífero no humano. Células huéspedes adecuadas se describen también en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA

ES 2 337 564 T3

(1990). De manera alterna, el vector de expresión recombinante puede transcribirse y traducirse *in vitro*, utilizando por ejemplo las secuencias reguladoras del promotor T7 y la polimerasa T7.

La expresión de proteínas en procariotas se efectúa la mayoría de veces con vectores que contienen promotores consecutivos o inducibles que controlan la expresión de proteínas de fusión o proteínas de no fusión. Los vectores de fusión agregan una serie de aminoácidos a una proteína codificada en los mismos, usualmente en el terminal amino de la proteína recombinante, pero también en el terminal C o se fusionan dentro de regiones adecuadas en las proteínas. Estos vectores de fusión usualmente tienen tres tareas: 1) reforzar la expresión de la proteína recombinante; 2) incrementar la solubilidad de la proteína recombinante y 3) apoyar la purificación de la proteína recombinante actuando como ligando en la purificación por afinidad. En el caso de los vectores de expresión de fusión se introduce frecuentemente un sitio de escisión proteolítica en el sitio donde se enlazan la unidad de fusión y la proteína recombinante, de modo que la proteína pueda separarse de la unidad de fusión después de la purificación de la proteína de fusión. Estas enzimas y sus correspondientes secuencias de reconocimiento comprenden el factor Xa, la trombina y las enterocinasas.

Vectores de expresión de fusión típicos son, entre otros, pGEX (Farmacia Biotech Inc; Smith, D.B., y Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Farmacia, Piscataway, NJ), en los que la glutatióna-S- transferasa (GST), la proteína de enlace E de maltosa o la proteína A se fusiona a la proteína recombinante destino. En una forma de realización, la secuencia codificante PSE se clona en un vector de expresión pGEX de modo que se genera un vector que codifica una proteína de fusión que comprende la proteína X del sitio de escisión de GST-trombina desde el terminal N hasta el terminal C. La proteína de fusión puede purificarse mediante cromatografía por afinidad usando resina de glutatióna-agarosa. El PSE recombinante que no se fusiona con GST puede obtenerse escindiendo la proteína de fusión con trombina.

Ejemplos de vectores de expresión inducibles adecuados de la *E. coli*, que no son de fusión, son entre otros pTrc (Amann y colaboradores (1988) *Gene* 69:301-315) y pET 11d (Studier y colaboradores, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). La expresión genética destino del vector pTrc se basa en la transcripción mediante la polimerasa de ARN huésped de un promotor de fusión híbrido, trp-lac. La expresión genética destino del vector pET 11d se basa en la transcripción de un promotor de fusión T7-f10-lac el cual está mediado por una polimerasa de ARN viral co-expresada (T7 gn1). Esta polimerasa viral es suministrada por las cepas huéspedes BL21 (DE3) o HMS174 (DE3) de un λ -profago residente, el cual hospeda un gen T7 f1 bajo el control de transcripción del promotor lacUV 5.

Otros vectores adecuados en organismos procarióticos son conocidos para el experto en la materia; estos vectores están, por ejemplo, en *E. coli* pLG338, pACYC184, la serie pBR, como pBR322, la serie pUC, como pUC18 o pUC19, la serie M113mp, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1, λ gt11 ó pBd-CI, en *Streptomyces* pIJ101, pIJ364, pIJ702 ó pIJ361, en *Bacillus* pUB110, pC194 o pBD214, en *Corynebacterium* pSA77 o pAJ667.

Una estrategia para maximizar la expresión de la proteína recombinante es expresar la proteína en una bacteria huésped cuya capacidad de escindir la proteína recombinante de manera proteolítica se deteriora (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 que se va a insertar en un vector de expresión, de modo que los codones individuales para cada aminoácido sean aquellos que se usan preferiblemente en una bacteria seleccionada para la expresión, tal como *C. glutamicum* (Wada y colaboradores (1992) *Nucleic Acids Res.* 20:2111-2118). Esta modificación de las secuencias de ácido nucleico de la invención se efectúa mediante técnicas estándar de síntesis de ADN.

En otra forma de realización, el vector de expresión del PSE es un vector de expresión de levadura. Los ejemplos de vectores para la expresión en la levadura *S. cerevisiae* comprenden pYepSec1 (Baldari y colaboradores (1987) *Embo J.* 6:229-234), pMFa (Kurjan y Herskowitz (1982) *Cell* 30:933-943), pJRY88 (Schultz y colaboradores (1987) *Gene* 54:113-123) así como pYES2 (Invitrogen Corporation, San Diego, CA). Los vectores y métodos para la construcción de vectores que sean adecuados para su uso en otros hongos, tales como el hongo filamentoso, incluyen aquellos que se describen con 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 y colaboradores, editores, páginas 1-28, Cambridge University Press: Cambridge, o en: *More Gene Manipulations in Fungi* [J.W. Bennet & L.L. Lasure, editores, páginas 396-428: Academic Press: San Diego]. Otros vectores de levadura adecuados son, por ejemplo, 2 α M, pAG-1, YE6, YE13 o pEMBLye23.

Alternamente, los PSEs pueden expresarse en células de insectos usando los vectores de expresión del baculovirus. Los vectores del baculovirus que están disponibles para expresar las proteínas en las células de insecto cultivadas (por ejemplo, células Sf9), comprenden la serie pAc-Reihe (Smith y colaboradores (1983) *Mol. Cell Biol.* 3:2156-2165) y la serie pVL (Lucklow y Summers (1989) *Virology* 170:31-39).

Los vectores arriba mencionados ofrecen solo una pequeña revisión de los vectores adecuados posibles. Los plásmidos adicionales son conocidos por el experto en la materia y se describen, por ejemplo, en: *Cloning Vectors* (Hrsgeb. Pouwels, P.H., y colaboradores, Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018).

ES 2 337 564 T3

En otra forma de realización, los PSEs de la invención pueden expresarse en células de plantas unicelulares (como algas), véase Falciatore y colaboradores, 1999, *Marine Biotechnology* 1 (3):239-251 y los datos bibliográficos citados allí, y en células vegetales de plantas superiores (por ejemplo, espermatofitos tales como frutos del campo). Ejemplos de vectores de expresión vegetales abarcan aquellos que se describen con detalle 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*, vol. 1, Engineering and Utilization, Editores: Kung y R. Wu, Academic Press, 1993, páginas 15-38.

Un casete de expresión vegetal contiene preferiblemente secuencias de regulación que pueden controlar la expresión genéticas en las células vegetales y las cuales se enlazan funcionalmente, de modo que cada secuencia pueda cumplir su función, tal como la terminación de la transcripción, por ejemplo las señales de poliadenilación. Señales de poliadenilación preferidas son aquellas provenientes del t-ADN de *Agrobacterium tumefaciens*, tal como el gen 3 conocido como octopinsintasa del plásmido Ti pTiACH5 (Gielen y colaboradores, *EMBO J.* 3 (1984) 835 y siguientes) o equivalentes funcionales del mismo, aunque también son adecuados todos los demás terminadores que son funcionalmente activos en las plantas.

Puesto que la expresión genética de las plantas frecuentemente no se limita al nivel de transcripción, un casete de expresión vegetal contiene preferiblemente otras secuencias enlazadas funcionalmente, tal como los potenciadores de traslación (traducción), por ejemplo la secuencia de sobremultiplicación (overdrive), la cual contiene la secuencia líder del virus mosaico del tabaco, 5' no traducida, la cual incrementa la relación proteína/ARN (Gallie y colaboradores, 1987, *Nucl. Acids Research* 15:8693-8711).

La expresión genética de las plantas debe enlazarse funcionalmente a un promotor adecuado que realiza la expresión del gen oportunamente de manera específica para una célula o un tejido. Se prefieren promotores que producen una expresión constitutiva (Benfey y colaboradores, *EMBO J.* 8 (1989) 2195-2202), como aquellos que provienen de virus vegetales, tales como 35S CAMV (Franck y colaboradores, *Cell* 21 (1980) 285-294), 19S CaMV (véase también US 5352605 y WO 84/02913) o promotores vegetales, como el promotor de la pequeña subunidad Rubisco, descrito en US 4,962,028.

Otras secuencias preferidas para el uso en enlace funcional en casetes de expresión genética en vegetales son secuencias objetivo (targeting), las cuales se requieren para localizar o dirigir el objetivo al producto genético en su correspondiente compartimiento celular (véase una sinopsis en Kermode, *Crit. Rev. Plant Sci.* 15, 4 (1996) 285-423 y las citas bibliográficas allí citadas), por ejemplo en la vacuola, el núcleo, todos los tipos de plástidos tales como amiloplastos, cloroplastos, cromoplastos, el espacio extracelular, la mitocondria, el retículo endoplasmático, oleoplastos, peroxisomas y otros compartimientos de las células vegetales.

La expresión genética vegetal también puede facilitarse por medio de un promotor inducible químicamente (véase una sinopsis en Gatz 1997, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 48:89-108). Promotores químicamente inducibles son adecuados en particular cuando se desea que una expresión genética se efectúe de manera oportuna. Ejemplos de tales promotores son: un promotor inducible con ácido salicílico (WO 95/19443), un promotor inducible con tetraciclina (Gatz y colaboradores (1992) *Plant J.* 2, 397-404) y un promotor inducible con etanol.

Promotores que reaccionan a condiciones de tensión bióticas o abióticas también son promotores adecuados, por ejemplo el promotor del gen PRP1 inducible con patógenos (Ward y colaboradores, *Plant. Mol. Biol.* 22 (1993) 361-366), el promotor hsp80 inducible por calor de tomate (US 5,187,267), el promotor de alfaamilasa inducible por frío a partir de patata (WO 96/12814) o el promotor pinII inducible por heridas (EP-A-0 375 091).

En particular se prefieren aquellos promotores que producen la expresión genética en tejidos y órganos, en los cuales tiene lugar la biosíntesis de lípidos y aceites, en células de semillas, tales como las células del endosperma y células del embrión en desarrollo. Promotores adecuados son: el promotor del gen napina de colza (US 5,608,152), el promotor USP de *Vicia faba* (Baeumlein y colaboradores, *Mol Gen Genet*, 1991, 225 (3):459-67), el promotor oleosina de *Arabidopsis* (WO 98/45461), el promotor faseolina de *Faseolus vulgaris* (US 5,504,200), el promotor Bce4 de brassica (WO 91/13980) o el promotor B4 de legumina (LeB4; Baeumlein y colaboradores, 1992, *Plant Journal*, 2 (2):233-9) así como promotores que provocan la expresión específica de semillas en plantas monocotiledóneas como maíz, cebada, trigo, centeno, arroz, etc. Promotores notables adecuados son el promotor del gen lpt2 ó lpt1 de cebada (WO 95/15389 y WO 95/23230) o los promotores descritos en WO 99/16890 del gen de hordeína - cebada, del gen de glutelina - arroz, del gen de orizina - arroz, del gen de prolamina - arroz, del gen de gliadina - trigo, gen glutelina - trigo, del gen de zeina - maíz, del gen de glutelina - avena, del gen kasirina - sorgo, del gen de secalina - centeno).

Así mismo son promotores particularmente adecuados aquellos que provocan la expresión específica de plástidos, puesto que los plástidos son el compartimiento en el que se sintetizan los precursores y algunos productos finales de la biosíntesis de lípidos. Promotores adecuados como el promotor de la polimerasa de ARN se describen en WO 95/16783 y WO 97/06250, y el promotor clpP de *Arabidopsis*, descrito en WO 99/46394.

Un vector de expresión recombinante puede introducirse en una célula huésped. Los términos "célula huésped", "célula huésped recombinante" y "célula huésped transgénica" se usan aquí de manera mutuamente intercambiable. Obviamente, estos términos se refieren no solo a determinadas células destino sino también a la progenie o progenie

ES 2 337 564 T3

potencial de esta célula. Ya que pueden ocurrir modificaciones específicas en las subsecuentes generaciones debido a la mutación o efectos ambientales, esta progenie no necesariamente es idéntica a la célula parenteral, aunque todavía está comprendida por el alcance del término como se utiliza aquí.

5 Una célula huésped puede ser una célula procariótica o eucariótica. Por ejemplo, un PSE puede expresarse en células bacteriales tales como *C. glutamicum*, células de insectos, células fúngicas o células mamíferas (tales como células de ovario de hámster chino (CHO) o células COS), algas, ciliados, células vegetales, hongos u otros microorganismos, como *C. glutamicum*. Otras células huéspedes adecuadas son corrientes para el experto en la materia.

10 Puede introducirse el ADN de vector en células procarióticas o eucarióticas por medio de técnicas convencionales de transformación o transfección. Los términos “transformación” y “transfección”, conjugación y transducción, tal como aquí se usan, deben abarcar un gran número de métodos conocidos en el estado de la técnica para introducir ácidos nucleicos ajenos (por ejemplo, ADN) a una célula huésped, incluyendo coprecipitación de fosfato de calcio o de cloruro de calcio, transfección mediada por DEAE Dextrano, lipofección, competencia natural, transferencia modificada
15 químicamente, electroporación o bombardeo de partículas. Métodos adecuados para la transformación o transfección de células huéspedes, incluyendo células vegetales, pueden encontrarse en Sambrook y colaboradores (Molecular Cloning: A Laboratory Manual., 2. Edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) y otros manuales de laboratorio, como Methods in Molecular Biology, 1995, vol. 44, Agrobacterium protocols, editors: Gartland y Davey, Humana Press, Totowa, New Jersey.

20 Para la generación de un microorganismo recombinado de manera homóloga, se produce un vector que contenga al menos un segmento de un gen PSE en el cual se ha introducido una eliminación, adición o sustitución con el fin de modificar de esa manera el gen PSE, por ejemplo para deteriorarlo funcionalmente. Este gen PSE es preferiblemente un gen PSE de *Physcomitrella patens* o de *Phytophthora infestans*, aunque puede usarse un homólogo o análogo de
25 otros organismos, incluso de origen mamífero, fúngico o de insectos. En una forma preferida de realización se diseña el vector de tal manera que el gen PSE endógeno se deteriora funcionalmente durante la recombinación homóloga (es decir, que ya no codifica una proteína funcional, también llamado knock-out-vector o agénico). De manera alterna, el vector puede diseñarse de tal manera que el gen PSE endógeno mute o se modifique de otra manera durante la recombinación homóloga mientras que todavía codifica una proteína funcional (por ejemplo, la región reguladora
30 corriente arriba puede modificarse de tal modo que por este medio se modifique la expresión del PSE endógeno). Para generar una mutación puntual mediante recombinación homóloga también pueden usarse híbridos de ADN-ARN, lo cual también se conocen como quimeroplastia y los cuales se conocen a partir de Cole-Strauss y colaboradores, 1999, Nucleic Acids Research 27(5):1323-1330 y Kmiec, Gene therapy, 1999, American Scientist, 87(3):240-247.

35 En el vector para la recombinación homóloga, el segmento modificado del gen PSE se flanquea en su extremo 5' y 3' mediante el ácido nucleico adicional del gen PSE, de modo que sea posible una recombinación homóloga entre el gen PSE exógeno que está presente en el vector y un gen PSE endógeno en un microorganismo o una planta. El ácido nucleico del PSE flanqueante adicional es lo suficientemente largo para la recombinación homóloga exitosa con el gen endógeno. Usualmente, varios cientos de pares de bases de hasta kilobases de ADN flanqueante (tanto en el extremo
40 5' como en el 3') están presentes en el vector (véase una descripción de vectores para la recombinación homóloga, por ejemplo en Thomas, K.R., y Capecchi, M.R. (1987) Cell 51:503 o para la recombinación en *Physcomitrella patens* en la base de ADNc, en Strepp y colaboradores, 1998, Proc. Natl. Acad. Sci. USA 95 (8):4368-4373). El vector se introduce en un microorganismo o célula vegetal (por ejemplo, mediante un ADN mediado con polietilenglicol) y se seleccionan las células en las que el gen PSE introducido ha padecido una recombinación homóloga con el gen PSE
45 endógeno utilizando técnicas conocidas en el campo técnico.

En otra forma de realización, pueden obtenerse microorganismos recombinantes que contienen sistemas seleccionados que hacen posible una expresión regulada del gen introducido. La inclusión del gen PSE en un vector, cuando el mismo se coloca bajo el control del lac-operon, permite, por ejemplo, la expresión del gen PSE solo en presencia
50 de IPTG. Estos sistemas de regulación son conocidos en el campo técnico.

En plantas puede aplicarse adicionalmente un método alterno mediante transferencia directa de ADN a las flores que se desarrollan mediante electroporación o transferencia de gen por medio de *Agrobacterium*.

55 Particularmente se prefieren plantas como las plantas con frutos oleaginosos que contienen grandes cantidades de compuestos lipídicos tales como colza, onagra, canola, maní o cacahuete, lino, soya, abrojo, girasol, borraja, o plantas como maíz, trigo, centeno, avena, triticale, arroz, cebada, algodón, yuca, pimienta, tagetes, plantas solanáceas como patata, tabaco, berenjena y tomate, especies de *Vicia*, guisantes, alfalfa, plantas de arbusto (café, cacao, té), especies *Salix*, árboles (palma de aceite, coco) así como pastos perennes y frutos del campo para forraje. Particularmente se
60 prefieren plantas con frutos oleaginosos como soya, maní, colza, canola, girasol, flor de alazor, árboles (palma de aceite, coco).

“Transgénico” significa, con respecto, por ejemplo, de una secuencia de ácido nucleico, de un casete de expresión (= construcción génica) o de un vector que contiene dicha secuencia de ácido nucleico o de un organismo transformado
65 con las secuencias de ácido nucleico, casete de expresión o vector arriba mencionados, todas aquellas construcciones que se producen mediante aquellos métodos de ingeniería genética, en las que ya sea

a) la secuencia de ácido nucleico según la invención, o

b) una secuencia de control genético conectada funcionalmente con la secuencia de ácido nucleico según la invención, por ejemplo un promotor, o

c) (a) y (b)

5

no están en su ambiente genético natural o se han modificado mediante métodos de ingeniería genética y la modificación puede ser por ejemplo sustituciones, adiciones, eliminaciones, inversión o inserciones de uno o varios residuos de nucleótidos. Ambiente genético natural significa el locus o sitio cromosómico natural en el organismo fuente o la presencia de una biblioteca genómica. En el caso de una biblioteca genómica, el ambiente genético natural de la secuencia de ácido nucleico se retiene preferiblemente al menos en parte. El ambiente flanquea la secuencia de ácido nucleico al menos en un lado y tiene una longitud de secuencia de al menos 50 bp, preferible de al menos 500 bp, particularmente preferible de al menos 1000 bp, muy particularmente preferible de al menos 5000 bp. Un casete de expresión de origen natural, por ejemplo la combinación de origen natural del promotor natural de la secuencia de ácido nucleico de la invención con el gen PSE correspondiente, se convierte en un casete de expresión transgénico si éste se modifica mediante métodos no naturales, sintéticos (“artificiales”), como por ejemplo un tratamiento mutagénico. Métodos correspondientes se han descrito, por ejemplo, en US 5,565,350 o WO 00/15815.

10

15

D. Usos y métodos según la invención

20

Las moléculas de ácido nucleico, proteínas, homólogos de proteína, proteínas de fusión, vectores y células huéspedes aquí descritos pueden usarse en un método según la reivindicación 1.

25

30

35

40

45

50

55

60

De esta manera la invención se refiere a un método para la producción de PUFAs, y el método comprende el cultivo de un microorganismo o una planta que comprende un ácido nucleico, una construcción génica o un vector, los cuales codifican un polipéptido que alarga ácidos grasos de C_{16} y/o C_{18} con al menos dos enlaces dobles en la molécula de ácido graso en al menos dos átomos de carbono en condiciones bajo las cuales se producen PUFAs en el organismo. PUFAs producidas por este método pueden aislarse mediante cosechando los organismos ya sea del cultivo en que crecen o del campo, mediante eclosión y/o extracción del material cosechado con un solvente orgánico. De este solvente puede aislarse el aceite que contiene lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerina y/o ácidos grasos libres con alto contenido de PUFAs. Mediante hidrólisis básica o ácida de los lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerina, pueden aislarse los ácidos grasos libres con contenido más alto de PUFAs. Un contenido más alto de PUFAs significa al menos 5%, preferiblemente 10%, particularmente preferible 20%, muy particularmente preferible 40% más PUFAs que el organismo original, por ejemplo un oomicete como *Phytophthora* o una planta como una planta con fruto oleaginoso, el/la cual no posee ácido nucleico adicional que codifique la elongasa de la invención. Además, los aceites, lípidos, fosfolípidos, esfingolípidos, glicolípidos, triacilglicerina y/o ácidos grasos libres previamente mencionados con contenido superior de PUFAs poseen otra composición que la composición del organismos de partida. Esto aplica en especial para vegetales que no contienen de manera natural ácidos grasos de C_{20} o C_{22} poliinsaturados de cadena larga, tales como DHA, EPA o ARA.

Se prefieren las moléculas de ácido graso de C_{20} o C_{22} de PUFAs, producidas por este método, con al menos dos enlaces dobles en la molécula de ácido graso, preferiblemente tres o cuatro enlaces dobles, particularmente preferible tres enlaces dobles. Estas moléculas de ácido graso de C_{20} o C_{22} pueden aislarse del organismo en forma de un aceite, lípido o de un ácido graso libre.

45

Ejemplos

Ejemplo 1

50

Métodos generales

a) Métodos generales de clonación

55

60

Métodos de clonación como, por ejemplo escisiones de restricción, electroforesis en gel de agarosa, purificación de fragmentos de ADN, transferencia de ácidos nucleicos a membranas de nitrocelulosa y nylon, enlace de fragmentos de ADN, transformación de células de *Escherichia coli* y de levadura, cultivo de bacterias y análisis de secuencia de ADN recombinante, se realizaron tal como se describe en Sambrook y colaboradores (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).

b) Sustancias químicas

65

A menos que se indique algo diferente en el texto, las sustancias químicas utilizadas se obtuvieron en una calidad de grado analítico de las empresas Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) y Sigma (Deisenhofen). Se prepararon soluciones usando agua pura, desprovista de pirógenos, denominada en el texto a continuación como H_2O , de una unidad purificadora de agua del sistema de agua Milli-Q (Millipore, Eschborn). Las endonucleasas de restricción, las enzimas para la modificación del ADN y los equipos de biología molecular se

ES 2 337 564 T3

obtuvieron de AGS (Heidelberg), Amersham (Brunswick), Biometra (Göttingen), 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). Si no se indica algo diferentes se usaron según las instrucciones del productor.

5

Ejemplo 2

Construcción del banco de ADNc

10

Para la construcción del banco de ADNc se logró la síntesis de la primera hebra usando transcriptasa inversa de virus de leucemia de ratón (Roche, Mannheim, Deutschland) y cebadores de oligo-d(T), la síntesis de la segunda hebra por incubación con polimerasa I de ADN, enzima Klenow y escisión con ARNasa H a 12°C (2 horas), 16°C (1 hora) y 22°C (1 hora). LA reacción se detuvo por incubación a 65°C (10 min) y a continuación se transfirió a hielo. Se elaboraron moléculas de ADN bicatenario provistas con extremos planos, con polimerasa de ADN T 34 (Roche, Mannheim) a 37°C (30 min). Se retiraron los nucleótidos mediante extracción con fenol/cloroformo y columnas de centrifugación Sephadex-G50. Se ligaron adaptadores EcoRI (Pharmacia, Freiburg, Alemania) a los extremos de ADNc por medio de ligasa de ADN T4 (Roche, 12°C, por una noche) y se fosforilaron por incubación con polinucleótido cinasa (Roche, 37°C, 30 min). Esta mezcla se sometió a la separación sobre gel de agarosa con bajo punto de fusión (low-melting). Las moléculas de ADN con más de 300 pares de bases se eluyeron del gel, se extrajeron con fenol, se concentraron en las columnas Elutip D (Schleicher y Schüll, Daßel, Alemania) y se ligaron con los brazos del vector y se empaquetaron en los fagos lambda-ZAPAI o fagos lambda-ZAP-express utilizando el kit Gigapack Gold (Stratagene, Amsterdam, Holanda); se utilizó el material del productor y se siguieron sus instrucciones.

25

Ejemplo 3

Secuenciación de ADN y análisis en ordenador

30

Se usaron bancos de ADNc tal como se describe en el ejemplo 4 para secuenciar el ADN mediante métodos estándar, en particular el método de terminación de cadenas usando el kit de reacción listo para secuenciación ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer, Weiterstadt, Alemania). La secuenciación al azar se realizó a continuación de la obtención preparativa de plásmido a partir de bancos de ADNc por un corte de masas *in vivo* y retransformación de DH10B sobre placas de agar (particularidades sobre el material y el protocolo de Stratagene, Amsterdam, Holanda). ADN del plásmido se preparó a partir de los cultivos de *E. coli* que habían crecido durante una noche en caldo de Luria suplementado con ampicilina (véase Sambrook y colaboradores (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)) en un robot para la preparación de ADN de Quiagen (Quiagen, Hilden) según los protocolos del productor. Se usaron cebadores de secuenciación con las siguientes secuencias de nucleótidos:

40

5'-CAGGAAACAGCTATGACC-3'

5'-CTAAAGGGAACAAAAGCTG-3'

45

5'-TGTA AACGACGGCCAGT-3'

Las secuencias se suministraron usando el paquete de software estándar EST-MAX, que se suministra comercialmente por Bio Max (Munich, Alemania), se procesaron y comentaron. Se caracterizó con mayor detalle un clon con homologías débiles a elongasas conocidas.

50

Ejemplo 4

*Identificación del gen PSE1 de *P. infestans* y análisis del clon de ADNc PiPSE1*

55

Una secuencia EST (acceso de la base de datos: PI001002014r) se consideró como un gen destino, entre otros genes candidatos, debido a su débil homología a las elongasas conocidas.

60

Para la comparación de secuencias se usó el programa BESTFIT, es decir las matrices de sustitución de aminoácidos BLOSUM; se hace referencia a Henikoff, S., y Henikoff, J.G. (1992), Amino acid substitution matrices from protein blocks, Proc. Natl. Acad. Sci. USA 89:10915-10919.

65

La secuencia del clon con el No. Del banco de datos PI001002014r se usó para la comparación con la secuencia de péptido elol de levadura. Ya que el nuevo clon de *P. infestans* no estaba completo, se aisló el correspondiente clon de longitud completa (full-length) (PiPSE1) partiendo del banco de ADNc de *P. infestans*. Para este propósito se generó una sonda etiquetada con digoxigenina mediante PCR por medio del kit de síntesis PCR DIG (Roche); PI001002014 sirvió de plantilla. Los siguientes cebadores se usaron para la PCR:

ES 2 337 564 T3

PI-DIGf: cacaccatcatgtacacttactac

PI-DIGr: caactctctcttcgattcctccac

- 5 La sonda etiquetada aislada se usó para seleccionar (screening) el banco de ADNc de la *P. infestans* (correspondiente a los datos del productor, Stratagene). Un fragmento de 1046 bp de longitud fue aislado y se denominó como PiPSE1. El marco de lectura abierto es de 837 bp de longitud y codifica para una proteína de 278 aminoácidos con una masa molecular calculada de 32,1 kDa. Las comparaciones de secuencia dieron como resultado las siguientes identidades o similitudes de secuencia: 26%/43% con el PSE1p de *Physcomitrella patens*, 23%/37% con el HELOp humano, 21%/41% con el GLELOp de *Mortierella alpina* y 17%/36% con la elongasa de *C. elegans*.
- 10

Ejemplo 5

15 Identificación de genes por medio de hibridación

Pueden usarse secuencias de genes para identificar genes homólogos o heterólogos de los bancos de ADNc o genómicos.

- 20 Los genes homólogos (es decir, clones de ADNc de longitud completa que son homólogos, u homólogos) pueden aislarse por medio de hibridación usando, por ejemplo, bancos de ADNc. Dependiendo de la frecuencia del gen de interés, se plantaron 100000 hasta 1000000 bacteriófagos recombinantes y se transfirieron a una membrana de nylon. Después de la desnaturalización con álcali, se inmovilizó el ADN en la membrana, por ejemplo mediante reticulación con UV. La hibridación se realiza bajo condiciones fuertemente severas. La hibridación y los pasos de lavado se llevan a cabo en solución acuosa a una fuerza iónica de NaCl de 1 M y una temperatura de 68°C. Se produjeron sondas de hibridación mediante marcación, por ejemplo, por medio de transcripción con mella radioactiva (³²P-) (High Prime, Roche, Mannheim, Alemania). Las señales se detectaron por medio de auto-radiografía.
- 25

- 30 Los genes parcialmente homólogos o heterólogos que se relacionan pero no son idénticos pueden identificarse análogamente con el proceso descrito anteriormente usando condiciones hibridación y de lavado de baja severidad. Para la hibridación acuosa, la fuerza iónica se mantuvo usualmente a NaCl de 1 M y la temperatura se disminuyó paulatinamente de 68 a 42°C.

- 35 El aislamiento de secuencias de genes que solo exhiben homologías con un dominio individual de, por ejemplo, 10 a 20 aminoácidos, puede llevarse a cabo usando sondas de oligonucleótidos radioetiquetadas sintéticas. Los oligonucleótidos radioetiquetados se generan fosforilando el extremo 5' de dos oligonucleótidos complementarios con el polinucleótido cinasa T4. Los oligonucleótidos complementarios se hibridaron y se ligaron entre sí para dar lugar a concatémeros. Los concatémeros bicatenarios se radioetiquetaron por ejemplo mediante transcripción de mellado. La hibridación usualmente se llevó a cabo en condiciones de baja severidad usando concentraciones elevadas de oligonucleótidos.
- 40

Solución de hibridación de oligonucleótidos:

SSC 6 x

45

Fosfato de sodio 0,01 M

1 mM de EDTA (pH 8)

50

SDS al 0,5%

100 µg/ml de esperam de salmón desnaturalizado

Leche seca magra al 0,1%

55

- 60 Durante la hibridación, la temperatura se disminuyó en etapas de 5 a 10°C por debajo de la temperatura calculada del oligonucleótido o a temperatura ambiente (a menos que se indique otra cosa, RT = ~ 23°C en todos los experimentos), seguido por pasos de lavado y autoradiografía. El lavado se llevó a cabo a una severidad extremadamente baja, por ejemplo 3 etapas de lavado usando SSC 4x. Se describen otros detalles en Sambrook, J., y colaboradores (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, o Ausubel, F.M., y colaboradores (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

Preparación de anticuerpos específicos, por ejemplo según

65

ES 2 337 564 T3

Ejemplo 6

Hibridación Northern

5 Para la hibridación de ARN se separaron 20 mg de ARN total ó 1 mg de poli(A)⁺-ARN mediante electroforesis con gel en geles de agarosa con una fuerza de 1,25% usando formaldehído, tal como se describe en Amasino (1986, Anal. Biochem. 152, 304), se transfirieron a membranas de nylon cargadas positivamente (Hybond N+, Amersham, Brunswick) mediante atracción capilar usando SSC 10 x, se inmovilizaron por medio de luz UV y se pre-hibridaron durante 3 horas a 68°C usando búfer de hibridación (sulfato de dextrano al 10% peso/vol., NaCl de 1 M, SDS al 10
10 1%, 100 mg de ADN de esperma de arenque). El etiquetado o marcación de la sonda de ADN se efectuó con el kit de etiquetado (labeling kit) de ADN Highprime (Roche, Mannheim, Alemania) durante la etapa de prehibridación usando alfa-³²P-dCTP (Amersham, Brunswick, Alemania). La hibridación se realizó después de agregar la sonda de ADN etiquetada en el mismo amortiguador a 68°C durante una noche. Los pasos de lavado se llevaron a cabo dos veces durante 15 minutos usando SSC 2x y dos veces durante 30 minutos usando SSC 1x SDS al 1%, a 68°C. Los
15 filtros sellados se expusieron a -70°C durante un período de 1 a 14 días.

Ejemplo 7

20 *Plásmidos para la transformación de las plantas*

Para la transformación de las plantas pueden usarse vectores binarios tales como pBinAR (Höfgen y Willmitzer, Plant Science 66 (1990) 221-230). La construcción de los vectores binarios puede efectuarse ligando el ADNc en orientación sense o antisense en el T-ADN. En el 5' del ADNc, un promotor vegetal activa la transcripción de ADNc.
25 La secuencia de poliadenilación se localiza en 3' del ADNc.

La expresión del tejido específico puede lograrse usando un promotor de tejido específico. Por ejemplo, la expresión específica de semillas puede lograrse clonando en el promotor 5' napina o LeB4 o USP del ADNc. También puede usarse cualquier elemento promotor de semillas específicas. El promotor CaMV 35S puede usarse para una expresión
30 constitutiva en la planta completa.

La proteína expresada puede dirigirse a un compartimento celular usando un péptido de señal, por ejemplo para los plástidos, mitocondrias o retículo endoplasmático (Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423). El péptido de señal se clona 5' en la estructura de lectura con el ADNc con el fin de lograr una localización subcelular de la proteína
35 de fusión.

Ejemplo 8

40 *Transformación de Agrobacterium*

La transformación de plantas mediada por Agrobacterium puede llevarse a cabo usando la cepa GV3101 de la *Agrobacterium tumefaciens* (pMP90-) (Koncz y Schell, Mol. Gen. Genet. 204 (1986) 383-396) o LBA4404 (Clontech). La transformación se realiza mediante técnicas estándar de transformación (Deblaere y colaboradores, Nucl. Acids.
45 Tes. 13 (1984), 4777-4788).

Ejemplo 9

50 *Transformación de plantas*

La transformación de plantas mediada con Agrobacterium puede llevarse a cabo usando técnicas estándar de transformación y regeneración (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Edición, Dordrecht: Kluwer Academic Publ., 1995, en Sect., Ringbuc Zentrale 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 S., ISBN 0-8493-5164-2).
55

Por ejemplo la colza puede transformarse por medio de la transformación de cotiledóneas o hipocotiledóneas (Moloney y colaboradores, Plant Cell Report 8 (1989) 238-242; De Block y colaboradores, Plant Physiol. 91 (1989) 694-701). El uso de antibióticos para la selección de agrobacterias y plantas depende del vector binario y la cepa agrobacteriana usada para la transformación. La selección de colza se realiza habitualmente usando canamicina como marcador de plantas seleccionables.
60

La transferencia de genes mediada con Agrobacterium en lino puede realizarse usando, por ejemplo, una técnica descrita por Mlynarova y colaboradores (1994) Plant Cell Report 13:282-285.
65

La transformación de soya puede realizarse usando, por ejemplo, una técnica descrita en EP-A-0 0424 047 (Pioneer Hi-Bred International) o en EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo).

ES 2 337 564 T3

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

5 Ejemplo 10

Mutagénesis in vivo

10 La mutagénesis de microorganismos *in vivo* puede realizarse pasando ADN del plásmido (o cualquier otro ADN del vector) a través de *E. coli* u otros microorganismo (por ejemplo, *Bacillus* spp. o levaduras tales como *Saccharomyces cerevisiae*), en los que se deteriora la capacidad de retener la integridad de su información genética. Las cepas de mutación convencionales tienen mutaciones en los genes para el sistema de reparación del ADN (por ejemplo mutHLS, mutD, mutT, etc.; como referencia bibliográfica véase Rupp, W.D. (1996) DNA repair mechanisms, en: 15 *Escherichia coli* and Salmonella, páginas 2277-2294, ASM: Washington). Estas cepas son conocidas para el experto en la materia. El uso de estas cepas se describe, por ejemplo, en Greener, A., y Callahan, M. (1994) Strategies 7:32-34. Las moléculas de ADN mutadas se transfieren preferiblemente a las plantas después de que los micororganismos se hayan seleccionado y probado. Las plantas transgénicas se generan según los diversos ejemplos en la sección de 20 ejemplos de este documento.

20 Ejemplo 11

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

25 La actividad de un producto genético recombinante en los organismos huéspedes se mide al nivel de transcripción y/o traducción.

Un método adecuado para determinar la cantidad de transcripción del gen (el cual indica la cantidad de ARN 30 disponible para la traducción del producto genético) para llevar a cabo un northern blot (como referencia véase Ausubel y colaboradores (1988) Current Protocols in Molecular Biology, Wiley: New York, o la sección de ejemplos ya mencionada arriba) en el cual se diseña un cebador de tal manera que se enlaza al gen de interés, se etiqueta con una etiqueta detectable (usualmente radioactividad o quimioluminiscencia) de modo que cuando se extrae el ARN total de un cultivo del organismo, se separa en un gel, se transfiera a una matriz estable y se incuba con esta sonda, el enlace 35 y el grado de enlace de la sonda indica la presencia así como la cantidad de ARNm para este gen. Esta información indica el grado de transcripción del gen transformado. El ARN total de la célula puede prepararse de células, tejidos u órganos con varios métodos, todos los cuales son conocidos en el campo técnico, tal como, por ejemplo, el método de Bormann, E.R., y colaboradores (1992) Mol. Microbiol. 6:317-326.

40 Para el estudio de la presencia o la cantidad relativa de proteína traducida por este ARNm pueden usarse técnicas estándar, tales como un northern blot (véase, por ejemplo, Ausubel y colaboradores (1988) Current Protocols in Molecular Biology, Wiley: New York). En este método se extraen las proteínas de células totales, se separan por medio de electroforesis con gel, se transfieren a una matriz tal como nitrocelulosa y se incuban con una sonda tal como un anticuerpo que se enlaza específicamente a la proteína deseada. Esa sonda usualmente se provee de una etiqueta quimioluminiscente o colorimétrica que puede detectarse fácilmente. La presencia y la cantidad de la etiqueta observada 45 indican la presencia y la cantidad de la proteína mutada deseada, la cual está presente en la célula.

50 Ejemplo 12

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

El efecto de la modificación genética en plantas, hongos, algas, ciliados o en la producción de un compuesto deseado (tal como ácido graso) puede determinarse desarrollando los microorganismos modificados o la planta modificada 55 en condiciones adecuadas (tales como aquellas descritas anteriormente) y analizando el medio y/o los componentes de la célula para la producción incrementada del producto deseado (es decir, de lípidos o de ácido graso). Estas técnicas analíticas son conocidas para el experto en la materia e incluyen la espectroscopía, cromatografía de capa delgada, métodos de tincura de diferentes tipos, métodos enzimáticos y microbiológicos, y cromatografía analítica tal como cromatografía líquida de alta resolución (véase, por ejemplo, Ullman, Encyclopedia of Industrial Chemistry, vol. A2, 60 páginas 89-90 y páginas 443-613, VCH: Weinheim (1985); Fallon, A., y colaboradores, (1987) "Applications of HPLC in Biochemistry" en: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm y colaboradores (1993) Biotechnology, vol. 3, capítulo III: "Product recovery and purification", páginas 469-714, VCH: Weinheim; Belter, P.A., y colaboradores (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, vol. B3; 65 capítulo 11, páginas 1-27, VCH: Weinheim; y Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

Junto a los métodos arriba mencionados, se extraen lípidos vegetales a partir del material vegetal como se describe en Cahoon y colaboradores (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, y Browse y colaboradores (1986) Analytic Biochemistry 152:141-145. El análisis cualitativo y cuantitativo de lípidos o de ácidos grasos se describe 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 pág. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) bajo el título: Progress in the Chemistry of Fats and Other Lipids CODEN.

Adicional a la medición del producto final de la fermentación, también es posible analizar otros componentes de rutas metabólicas que se utilizan para producir el compuesto deseado, tal como productos intermedios y productos secundarios, con el fin de determinar la eficiencia total en la producción del compuesto. Los métodos analíticos incluyen mediciones de las cantidades de nutrientes en el medio (por ejemplo, azúcares, carbohidratos, fuentes de nitrógeno, fosfato y otros iones), mediciones de la composición de biomasa y mediciones de crecimiento, análisis de la producción de metabolitos habituales de rutas biosintéticas y mediciones de gases que se generan durante la fermentación. Los métodos estándar para estas mediciones se describen en Applied Microbial Physiology; A Practical Approach, P.M. Rhodes y P.F. Stanbury, Editores, IRL Press, páginas 103-129; 131-163 y 165-192 (ISBN: 0199635773) y las citas bibliográficas allí indicadas.

Un ejemplo es el análisis de ácidos grasos (abreviaturas: FAME, éster metílico de ácido graso; GC-MS, cromatografía de gases-líquidos/espectrometría de masas; TAG, triacilglicerina; TLC, cromatografía de capa delgada).

La detección inequívoca de la presencia de productos de ácido graso puede obtenerse analizando organismos recombinantes mediante métodos analíticos estándar: GC, GC-MS ó TLC, tal como se describen de manera diversa por Christie y en las referencias bibliográficas allí citadas (1997, en: Advances on Lipid Methodology, cuarta edición: Christie, Oily Press, Dundee, 119-169; 1998, método de cromatografía de gases - espectrometría de masas, Lípidos 33:343-353).

El material a analizar puede quebrarse por tratamiento con ultrasonido, molienda en molino de vidrio, nitrógeno líquido y molienda o por otros métodos aplicables. El material debe centrifugarse después de quebrarse. El sedimento se resuspende en agua destilada, se calienta 10 min a 100°C, se enfría sobre hielo y se centrifuga de nuevo, sigue una extracción en ácido sulfúrico de 0,5 M en metanol con dimetoxipropano al 2% por 1 hora a 90°C, lo cual conduce a compuestos hidrolizados de aceite y lípidos, que dan como resultado lípidos transmetilados. Estos ésteres metílicos de ácido graso se extraen en éter de petróleo y finalmente se someten a análisis GC usando una columna capilar (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 m, 0,32 mm) a un gradiente de temperatura entre 170°C y 240°C por 20 min y 5 min a 240°C. La identidad de los ésteres metílicos del ácido graso resultante deben definirse usando estándares que estén disponible comercialmente (es decir, Sigma).

En el caso de ácidos grasos para los cuales no estén disponibles los estándares, debe demostrarse la identidad a través de la derivación seguida por el análisis GM-MS. Por ejemplo, la localización de ácidos grasos con triple enlace debe demostrarse por GM-MS después de derivación con derivados de 4,4-dimetoxioxazolina (Christie, 1998, véase arriba).

Ejemplo 13

Construcciones de expresión en sistemas microbianos heterólogos

Cepas, condiciones de crecimiento y plásmidos

La cepa de *Escherichia coli* XL1 Blue MRF^r kan (Stratagene) se usó para subclonar el nuevo piPSEI de elongasa de *Phytophthora infestans*. Para la expresión funcional de este gen nosotros usamos la cepa INVSc 1 de *Saccharomyces cerevisiae* (Invitrogen Co.). Se hizo crecer *E. coli* en caldo de Luria-Bertini (LB, Duchefa, Haarlem, Holanda) a 37°C. Cuando fue necesario, se adicionó ampicilina (100 mg/litro), y se agregó agar al 1,5% (peso/volumen) para los medios sólidos LB. *S. cerevisiae* se cultivó a 30C tanto en medio YPG o en medio mínimo completo sin uracilo (CMdum; véase 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) o bien con rafinosa o glucosa al 2% (peso/volumen). Para medios sólidos se adicionó Bacto™-Agar al 2% (peso/volumen) (Difco). Los plásmidos usados para la clonación y expresión fueron pUC18 (Farmacia) y pYES2 (Invitrogen Co.).

Clonación y expresión de una elongasa específica de PUFA a partir de Phytophthora infestans

Para la expresión en levadura, el clon piPSEI del ADNc de *Phytophthora infestans* que codifica el gen de elongasa específica de PUFA (PSEI) se modificó primero de tal modo que un sitio de restricción KpnI y la secuencia de consenso de levadura para una traducción altamente efectiva (Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292) se obtiene después del codón de inicio y se obtiene un sitio de restricción XbaI que flanquea el codón de terminación. Para amplificar la estructura de lectura abierta se sintetiza un par cebador que es complementario de sus extremos 5' y 3'.

ES 2 337 564 T3

ppex1f: cggggtaccacataatgtcgactgagctactgcag

ppex1r: cactagtctagattccaactcttcttcgattcc

5 La reacción PCR se llevó a cabo con un ADN de plásmido como matriz en un formador de ciclos térmicos (Biomet) con polimerasa de ADN Pfu (Stratagene) y el siguiente programa de temperatura: 3 minutos a 96° seguido por 30 ciclos con 30 segundos a 96°C, 30 segundos a 55°C y 2 minutos a 72°C, 1 ciclo con 10 minutos a 72°C y detención a 4°C.

10 El tamaño correcto de 883 bp del fragmento de ADN amplificado se confirmó mediante electroforesis con gel TBE de agarosa. EL ADN amplificado se extrajo del gel usando el kit de extracción de gel QIAquick (QIAGEN) y se ligó en el sitio de restricción SmaI del vector pUC18 desfosforilado usando el kit Sure Clone Ligation (Pharmacia); se obtuvo pUCPSE1. Después de la transformación del XL1 Blue MRF' kan de *E. coli*, se llevó a cabo una mini-preparación del ADN (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation. *BioTechniques* 4, 310-313) de 24 transformantes resistentes a ampicilina, y se identificaron los clones positivos por medio de análisis de restricción de BamHI. La secuencia del producto PCR clonado se confirmó mediante re-secuenciación usando el kit ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt).

20 El ADN de plásmido de pUC-PSE1 se escindió adicionalmente con KpnI/XbaI y el fragmento resultante de ~900 bp se ligó en el sitio de restricción KpnI/XbaI del vector defosforilado shuttle (transportador) levadura - *E. coli* pYES2; se obtuvo pY2PSE1. Después de la transformación de la minipreparación de *E. coli* y ADN de los transformantes, la orientación del fragmento de ADN en el vector se verificó mediante la escisión con HindIII. Se hizo crecer un clon con el kit de extracción de ADN del plásmido Nucleobond® AX 500 (Macherey-Nagel, Düringen).

25 El INVSc1 de *Saccharomyces* se transformó con pY2PSE1 y pYES2 por medio de un protocolo modificado de PEG/acetato de litio (Ausubel y colaboradores, 1995). Después de la selección de las placas de agar Cmdum suplementadas con glucosa al 2%, en cada caso se seleccionaron cuatro transformantes pY2PSE11 (pY2PSE1a-d) y un transformante pYES2 para el cultivo adicional y la expresión funcional.

30 *Expresión funcional de una actividad de elongasa en levadura*

Precultivo

35 Se inocularon 20 ml de medio Cmdum con rafinosa al 2% (Peso/volumen) con los clones de levadura transgénica (pY2PSE1a-d, pYES2) y se cultivaron durante 3 días 30°C, 200 rpm, hasta haber alcanzado una densidad óptica a 600 nm (OD₆₀₀) de 1,5-2.

Cultivo principal

40 Para la expresión se suplementaron 20 ml del medio líquido Cmdum con rafinosa al 2% y tergitol NP-40 al 1% (Vol./Vol.) con ácido γ -linoleico (g-18:3) a una concentración final de 0,003% (peso/volumen). El medio se inoculó con los precultivos a una OD₆₀₀ de 0,05. La expresión se indujo durante 16 horas a una OD₆₀₀ de 0,2 con galactosa al 2% (peso/volumen), después de lo cual los cultivos habían alcanzado una OD₆₀₀ de 0,8-1,2.

45 *Análisis de ácido graso*

50 Los ácidos grasos totales se extrajeron de cultivos de levadura y se analizaron por medio de cromatografía de gas. Se aquí se recogieron células de 5 ml de cultivo por medio de centrifugación (1000 x g, 10 min, 4°C) y se lavaron una vez con NaHCO₃ de 100 mM, pH 8,0, con el fin de retirar el medio residual y los ácidos grasos. Para la preparación del éster metílico de ácido graso (FAMES) se trataron los sedimentos de célula con H₂SO₄ metanólico de 1 M y dimetoxipropano al 2% (vol./vol.) por 1 hora a 80°C. El FAMES se extrajo dos veces con 2 ml de éter de petróleo, una vez con NaHCO₃ de 100 mM, pH 8,0, y una vez con agua destilada y se secó con Na₂SO₄. El solvente orgánico se evaporó bajo una corriente de argón y se disolvió el FAMES en 50 μ l de éter de petróleo. Las muestras se separaron en una columna capilar ZEBRONZB-Wax (30 m, 0,32 mm, 0,25 μ m; Phenomenex) en un cromatógrafo de gases Hewlett Packard-6850 con un detector de ionización de flama. La temperatura del horno se programó de 70°C (mantenida por 1 min) hasta 200°C con una velocidad de 20°C/min, luego a 250°C (mantenida por 5 min) con una velocidad de 5°C/min y finalmente a 260°C con una velocidad de 5°C/min. Se usó nitrógeno como gas portador (4,5 ml/min a 70°C). Los ácidos grasos se identificaron mediante comparación con tiempos de retención de los estándares FAMES (SIGMA).

60 *Análisis de expresión*

Los patrones de ácido graso de cinco cepas de levadura transgénica se muestran en la tabla 1 en % molar.

65 Las relaciones del ácido γ -linoléico que se han agregado y absorbido se resaltan con números impresos en negrilla, aquellos de los productos alargados mediante números en rojo y aquellos del ácido γ -linoléico mediante números impresos con negrilla (última línea). El análisis GC de FAMES que es de lípidos totales de las levaduras transformadas

ES 2 337 564 T3

con pYES2 (i/control) y pY2PSE1, se muestra en la figura 1. Para el análisis se cultivan levaduras transgénicas en presencia de γ -18:3.

5 Los resultados muestran que γ -18:3 se ha incorporado en todas las levaduras transgénicas en grandes cantidades. Todos los cuatro clones de levadura transgénica que se han transformado con pY2PSE1 muestran un pico adicional en el cromatograma de gases que se identificó como 20:3^{A8,11,14} mediante comparación de los tiempos de retención. Una cromatografía de gases/espectroscopía de masas puede proporcionar un apoyo adicional para confirmar esta identidad.

10 Los productos identificados mostraron que la secuencia de nucleótidos de PiPSE1 codifica una elongasa de ácido graso $\Delta 6$ -selectiva de la *Physcomitrella patens* del mohó, la cual conduce a la formación de nuevos ácidos grasos en levaduras transgénicas.

15 Pueden llevarse a cabo otros experimentos de alimentación con otros ácidos grasos diversos (por ejemplo, ácido linoleico (18:2^{A9,12?}), ácido esteárico (18:4^{A6,9,12,15})) para confirmar la selectividad del sustrato de esta elongasa con mayor detalle.

Ejemplo 14

20 *Purificación del producto deseado de organismos transformados en general*

La obtención del producto deseado puede efectuarse a partir de material vegetal u hongos, algas, ciliados, células animales o del sobrenadante de los cultivos antes descritos mediante diversos métodos conocidos en el campo técnico. Si el producto deseado no se secreta de las células, pueden cosecharse las células del cultivo mediante centrifugación 25 lenta y las células pueden lisar mediante técnicas estándar, tales como fuerza mecánica o tratamiento con ultrasonido. Pueden prepararse mecánicamente órganos de las plantas de otro tejido u otros órganos. Después de la homogeneización, el desecho celular se retira mediante centrifugación y la fracción del sobrenadante que contiene las proteínas solubles se retiene para aislar adicionalmente el compuesto deseado. Si el producto se secreta de células deseadas, las células se retiran a partir del cultivo mediante centrifugación lenta y la fracción del sobrenadante se retiene para el 30 aislamiento adicional.

La fracción del sobrenadante de cada paso de purificación se somete a cromatografía con una resina adecuada, se retiene la molécula deseada ya sea en la resina de cromatografía aunque no se retengan muchas impurezas en la muestra, o las impurezas permanecen en la resina aunque no en la muestra. Estos pasos de cromatografía pueden 35 repetirse cuando sea necesario y se usan las mismas resinas de cromatografía u otras. El experto en la materia está familiarizado con la selección de resinas de cromatografía adecuadas y con su uso más efectivo para una molécula en particular que va a purificarse. El producto purificado puede concentrarse mediante filtración o ultrafiltración y se almacena a una temperatura en la que la estabilidad del producto es máxima.

40 En el campo técnico se conoce un amplio espectro de métodos de purificación y el método de purificación previo no debe ser limitante. Estos métodos de purificación se describen, por ejemplo, en Bailey, J.E., & Ollis, D.F., *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986).

La identidad y pureza de los compuestos aislados puede determinarse por técnicas estándar del campo técnico. 45 Las mismas incluyen cromatografía líquida de alta resolución (HPLC), métodos espectroscópicos, métodos de tinción, cromatografía de capa delgada, NIRS, ensayos con enzimas o métodos microbiológicos. Véase una sinopsis de estos métodos de análisis en: Patek y colaboradores (1994) *Appl. Environ. Microbiol.* 60:133-140; Malakhova y colaboradores (1996) *Biotekhnologiya* 11: 27-32; y Schmidt y colaboradores (1998) *Bioprocess Engineer.* 19:67-70. *Ulmann's Encyclopedia of Industrial Chemistry* (1996) vol. A27, VCH: Weinheim, páginas 89-90, páginas 521-540, 50 páginas 540-547, páginas 559-566, 575-581 y páginas 581-587; Michal, G (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A., y colaboradores (1987) *Applications of HPLC in Biochemistry en: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.

55

60

65

REIVINDICACIONES

5 1. Método para la preparación de moléculas de ácidos grasos de C_{20} o C_{22} con al menos dos enlaces dobles en la molécula de ácido graso; el método comprende el cultivo de un microorganismo o de una planta, el/la cual comprende un ácido nucleico que codifica un polipéptido que alarga ácidos grasos de C_{16} o C_{18} que tienen al menos dos enlaces dobles en el ácido graso en al menos dos átomos de carbono; se prefieren los ácidos grasos $C_{18:3}^{\Delta 6,9,12}$, $C_{18:4}^{\Delta 6,9,12,15}$ y $C_{16:3}^{\Delta 7,10,13}$ al menos en el factor 1,5 frente a los ácidos insaturados $C_{18:2}^{\Delta 9,12}$, $C_{18:3}^{\Delta 4,7,10}$, $C_{18:3}^{\Delta 5,8,11}$, $C_{18:3}^{\Delta 7,10,13}$, $C_{18:3}^{\Delta 8,11,14}$, $C_{18:3}^{\Delta 9,12,15}$ o $C_{18:3}^{\Delta 5,e9,12}$; los ácidos grasos $C_{18:3}^{\Delta 5t,9,12}$, $C_{20:3}^{\Delta 8,11,14}$, $C_{20:4}^{\Delta 5,8,11,14}$ y $C_{20:5}^{\Delta 5,8,11,14,17}$ no se alargan y el ácido nucleico se selecciona del grupo compuesto de

- 10 a) una secuencia de ácido nucleico representada en SEQ ID NO:1,
- 15 b) una secuencia de ácido nucleico que se deriva según el código genético degenerado de la secuencia representada en SEQ ID NO: 2,
- c) Derivados de la secuencia representada en SEQ ID NO:1 que codifican polipéptidos que son idénticos al menos en 50% a la secuencia de aminoácidos representada en SEQ ID NO:2.

20 2. Método según la reivindicación 1, donde las moléculas de los ácidos grasos de C_{20} o C_{22} se aíslan del organismo en forma de un aceite, lípido o de un ácido graso libre.

25 3. Uso de un ácido nucleico tal como se describe en la reivindicación 1 para alargar ácidos grasos de C_{16} o C_{18} que tienen al menos dos enlaces dobles en el ácido graso en al menos dos átomos de carbono y se prefieren los ácidos grasos $C_{18:3}^{\Delta 6,9,12}$, $C_{18:4}^{\Delta 6,9,12,15}$ y $C_{16:3}^{\Delta 7,10,13}$ al menos en el factor 1,5 frente a los ácidos grasos insaturados $C_{18:2}^{\Delta 9,12}$, $C_{18:3}^{\Delta 4,7,10}$, $C_{18:3}^{\Delta 5,8,11}$, $C_{18:3}^{\Delta 7,10,13}$, $C_{18:3}^{\Delta 8,11,14}$, $C_{18:3}^{\Delta 9,12,15}$ o $C_{18:3}^{\Delta 5,e9,12}$ y no se alargan los ácidos grasos $C_{18:3}^{\Delta 5t,9,12}$, $C_{20:3}^{\Delta 8,11,14}$, $C_{20:4}^{\Delta 5,8,11,14}$ y $C_{20:5}^{\Delta 5,8,11,17}$.

30

35

40

45

50

55

60

65

ES 2 337 564 T3

LISTA DE SECUENCIAS

<110> BASF Plant Science GmbH

<120> Nuevo gen de elongasa y métodos para la producción de ácidos grasos poliinsaturados

<130> 925_2001

<140>

<141>

<160> 2

<170> PatentIn Vers. 2.0

<210> 1

<211> 1066

<212> DNA

<213> *Phytophthora infestans*

<220>

<221> CDS

<222> (52)..(888)

<400> 1

```

25      gaattcggca cgaggttcgc acgtccatcg tctactcacc aacaagaagt c atg tcg 57
                                         Met Ser
                                         1

30      act gag cta ctg cag agc tac tac gcg tgg gcc aac gcc acg gag gcc 105
      Thr Glu Leu Leu Gln Ser Tyr Tyr Ala Trp Ala Asn Ala Thr Glu Ala
          5          10          15

35      aag ctg ctg gac tgg gtc gac cct gag ggc ggc tgg aag gtg cat cct 153
      Lys Leu Leu Asp Trp Val Asp Pro Glu Gly Gly Trp Lys Val His Pro
          20          25          30

40      atg gca gac tac ccc cta gcc aac ttc tcc agc gtc tac gcc atc tgc 201
      Met Ala Asp Tyr Pro Leu Ala Asn Phe Ser Ser Val Tyr Ala Ile Cys
          35          40          45          50

45      gtc gga tac ttg ctc ttc gta atc ttc ggc acg gcc ctg atg aaa atg 249
      Val Gly Tyr Leu Leu Phe Val Ile Phe Gly Thr Ala Leu Met Lys Met
          55          60          65

50      gga gtc ccc gcc atc aag acc agt cca tta cag ttt gtg tac aac ccc 297
      Gly Val Pro Ala Ile Lys Thr Ser Pro Leu Gln Phe Val Tyr Asn Pro
          70          75          80

55      atc caa gtc att gcc tgc tct tat atg tgc gtg gag gcc gcc atc cag 345
      Ile Gln Val Ile Ala Cys Ser Tyr Met Cys Val Glu Ala Ala Ile Gln
          85          90          95

60      gcc tac cgc aac ggc tac acc gcc gcc ccg tgc aac gcc ttt aag tcc 393
      Ala Tyr Arg Asn Gly Tyr Thr Ala Ala Pro Cys Asn Ala Phe Lys Ser
          100

65      gac gac ccc gtc atg ggc aac gtt ctg tac ctc ttc tat ctc tcc aag 441
      Asp Asp Pro Val Met Gly Asn Val Leu Tyr Leu Phe Tyr Leu Ser Lys
          115          120          125          130

70      atg ctc gac ctg tgc gac aca gtc ttc att atc cta gga aag aag tgg 489
      Met Leu Asp Leu Cys Asp Thr Val Phe Ile Ile Leu Gly Lys Lys Trp
          135          140          145

75      aaa cag ctt tcc atc ttg cac gtg tac cac cac ctt acc gtg ctt ttc 537
      Lys Gln Leu Ser Ile Leu His Val Tyr His His Leu Thr Val Leu Phe
  
```

ES 2 337 564 T3

	150	155	160	
5	gtc tac tat Val Tyr Tyr 165	gtg acg ttc cgc gcc gct Val Thr Phe Arg Ala Ala 170	cag gac ggg gac tca Gln Asp Gly Asp Ser 175	tat gct Tyr Ala 585
	acc atc gtg ctc aac ggc ttc Thr Ile Val Leu Asn Gly Phe 180 185	gtg cac acc atc atg tac act Val His Thr Ile Met Tyr Thr 190	tac tac tac Tyr Tyr Tyr 633	
10	ttc gtc agc gcc cac acg cgc Phe Val Ser Ala His Thr Arg 195 200	aac att tgg tgg aag aag tac Asn Ile Trp Trp Lys Lys Tyr 205	ctc acg Leu Thr 210	681
15	cgc att cag ctt atc cag ttc Arg Ile Gln Leu Ile Gln Phe 215	gtg acc atg aac gtg cag ggc Val Thr Met Asn Val Gln Gly 220 225	tac ctg Tyr Leu 225	729
20	acc tac tct cga cag tgc cca Thr Tyr Ser Arg Gln Cys Pro 230	ggc atg cct cct aag gtg ccg Gly Met Pro Pro Lys Val Pro 235 240	ctc atg Leu Met 240	777
	tac ctt gtg tac gtg cag tca Tyr Leu Val Tyr Val Gln Ser 245 250	ctc ttc tgg ctc ttc atg aat Leu Phe Trp Leu Phe Met Asn 255	tac Tyr 825	
25	att cgc gcg tac gtg ttc ggc Ile Arg Ala Tyr Val Phe Gly 260 265	ccc aag aaa ccg gcc gtg gag Pro Lys Lys Pro Ala Val Glu 270	gaa tcg Glu Ser 873	
30	aag aag aag ttg taa Lys Lys Lys Leu	cggcgcttgt taaaaagtct aacctgctg 275	taacagctta 928	
	aaacacacac acacacaacg ctttgtagag gcagcatgcc catcaaatac atcccgtatg 35 aaaaaaaaaa aactcgag	gaggtaagta gtggcaactc gtgtagaaat aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	988 1048 1066	

<210> 2

<211> 278

40 <212> PRT

<213> *Phytophthora infestans*

<400> 2

45	Met Ser Thr Glu Leu Leu Gln Ser Tyr Tyr Ala Trp Ala Asn Ala Thr 1 5 10 15
	Glu Ala Lys Leu Leu Asp Trp Val Asp Pro Glu Gly Gly Trp Lys Val 20 25 30
50	His Pro Met Ala Asp Tyr Pro Leu Ala Asn Phe Ser Ser Val Tyr Ala 35 40 45
	Ile Cys Val Gly Tyr Leu Leu Phe Val Ile Phe Gly Thr Ala Leu Met 50 55 60
55	Lys Met Gly Val Pro Ala Ile Lys Thr Ser Pro Leu Gln Phe Val Tyr 65 70 75 80
60	Asn Pro Ile Gln Val Ile Ala Cys Ser Tyr Met Cys Val Glu Ala Ala 85 90 95
	Ile Gln Ala Tyr Arg Asn Gly Tyr Thr Ala Ala Pro Cys Asn Ala Phe 100 105 110

65

ES 2 337 564 T3

Lys Ser Asp Asp Pro Val Met Gly Asn Val Leu Tyr Leu Phe Tyr Leu
 115 120 125
 5 Ser Lys Met Leu Asp Leu Cys Asp Thr Val Phe Ile Ile Leu Gly Lys
 130 135 140
 Lys Trp Lys Gln Leu Ser Ile Leu His Val Tyr His His Leu Thr Val
 145 150 155 160
 10 Leu Phe Val Tyr Tyr Val Thr Phe Arg Ala Ala Gln Asp Gly Asp Ser
 165 170 175
 Tyr Ala Thr Ile Val Leu Asn Gly Phe Val His Thr Ile Met Tyr Thr
 180 185 190
 15 Tyr Tyr Phe Val Ser Ala His Thr Arg Asn Ile Trp Trp Lys Lys Tyr
 195 200 205
 20 Leu Thr Arg Ile Gln Leu Ile Gln Phe Val Thr Met Asn Val Gln Gly
 210 215 220
 Tyr Leu Thr Tyr Ser Arg Gln Cys Pro Gly Met Pro Pro Lys Val Pro
 225 230 235 240
 25 Leu Met Tyr Leu Val Tyr Val Gln Ser Leu Phe Trp Leu Phe Met Asn
 245 250 255
 30 Phe Tyr Ile Arg Ala Tyr Val Phe Gly Pro Lys Lys Pro Ala Val Glu
 260 265 270
 Glu Ser Lys Lys Lys Leu
 275

35

40

45

50

55

60

65