



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 367 182**

(51) Int. Cl.:

C12N 9/02 (2006.01)

C12N 15/82 (2006.01)

C12N 15/53 (2006.01)

C12N 5/10 (2006.01)

C12N 5/14 (2006.01)

A61K 38/44 (2006.01)

A61P 3/10 (2006.01)

A61P 35/00 (2006.01)

A61P 9/00 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **07019867 .6**

(96) Fecha de presentación : **28.09.2001**

(97) Número de publicación de la solicitud: **1911837**

(97) Fecha de publicación de la solicitud: **16.04.2008**

(54)

Título: **FAD5-2, miembro de la familia de la desaturasa de ácido graso y usos del mismo.**

(30)

Prioridad: **28.09.2000 US 236303 P**
12.06.2001 US 297562 P

(45)

Fecha de publicación de la mención BOPI:
31.10.2011

(45)

Fecha de la publicación del folleto de la patente:
31.10.2011

(73)

Titular/es: **BIORIGINAL FOOD & SCIENCE Corp.**
102 Melville Street
Saskatoon, Saskatchewan S7J 0R1, CA

(72)

Inventor/es: **Qui, Xiao y**
Hong, Haiping

(74)

Agente: **Carvajal y Urquijo, Isabel**

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

FAD5-2, miembro de la familia de la desaturasa de ácido graso y usos del mismo.

Antecedentes de la invención

5 Los ácidos grasos son ácidos carboxílicos con grupos laterales hidrocarbonados de cadena larga y juegan un papel fundamental en muchos procesos biológicos. Los ácidos grasos raramente se encuentran libres en la naturaleza pero, en vez de eso, se presentan en forma esterificada como el componente principal de los lípidos. Los lípidos/ácidos grasos son fuentes de energía (por ejemplo, β -oxidación) y son una parte integral de las membranas de la célula que son indispensables para el procesamiento de información biológica o bioquímica.

10 Los ácidos grasos pueden dividirse en dos grupos: los ácidos grasos saturados y los ácidos grasos insaturados que contienen uno o más dobles enlaces de carbono en configuración *cis*. Los ácidos grasos insaturados son producidos por desaturasas terminales que pertenecen a la clase de enzimas con hierro no hemo. Cada una de estas enzimas hacen parte del sistema de transporte de electrones que contiene otras dos proteínas, a saber citocromo b_5 y NADH-citocromo b_5 reductasa. Específicamente, tales enzimas catalizan la formación de enlaces dobles entre los átomos de carbono de una molécula de ácido graso. Los humanos y otros mamíferos tienen un espectro limitado de estas desaturasas que son requeridas para la formación de enlaces dobles particulares en ácidos grasos insaturados. De este modo, los humanos tienen que tomar algunos ácidos grasos a través de su dieta. Tales ácidos grasos esenciales son, por ejemplo, ácido linoleico (C18:2); ácido linolénico (C18:3), ácido araquidónico (C20:4). En contraste, los insectos y plantas son capaces de sintetizar una variedad mucho mayor de ácidos grasos insaturados y sus derivados.

20 Los ácidos grasos poliinsaturados de cadena larga (LCPUFA) tales como el ácido docosahexaenóico (DHA, 22:6 (4, 7, 10, 13, 16, 19)) son componentes esenciales de las membranas celulares de diferentes organelos y tejidos en mamíferos (nervios, retina, cerebro y células inmunes). Por ejemplo, más del 30% de los ácidos grasos en fosfolípidos del cerebro son 22:6 (n-3) y 20:4 (n-6). (Crawford, M. A., et al., (1997) Am. J. Clin. Nutr. 66: 1032S-1041S). En la retina, el DHA representa más del 60% de los ácidos grasos totales en el segmento exterior del bastón, la parte fotosensible de la célula fotorreceptora. (Giusto, N. M., et al. (2000) Prog. Lipid Res. 39: 315 - 391). Estudios clínicos han mostrado que el DHA es esencial para el crecimiento y desarrollo del cerebro en los infantes, y para el mantenimiento de la función normal del cerebro en adultos (Martinetz, M. (1992) J. Pediatr. 120: S129 - S138). El DHA también tiene efectos significativos sobre la función fotorreceptora involucrada en el proceso de transducción de la señal, activación de la rodopsina, y el desarrollo de conos y batones (Giusto, N. M., et al. (2000) Prog. Lipid Res. 39: 315 - 391). Además, se encontraron también algunos efectos positivos del DHA sobre enfermedades tales como hipertensión, artritis, aterosclerosis, depresión, trombosis y cánceres (Horrocks, L. A. y Yeo, Y. K. (1999) Pharmacol. Res. 40: 211 - 215). Por lo tanto, el suministro apropiado en la dieta del ácido graso es importante para que los humanos permanezcan sanos. Es particularmente importante para infantes, niños pequeños y adultos mayores ingerir adecuadamente estos ácidos grasos a partir de la dieta ya que ellos no pueden ser eficientemente sintetizados en sus organismos y deben ser suministrados por los alimentos (Spector, A. A. (1999) Lipids 34: S1 - S3).

40 El DHA es un ácido graso de la serie n-3 de acuerdo con la ubicación del último doble enlace en el metilo final. Se sintetiza a través de etapas alternantes de desaturación y alargamiento. Comenzando con 18:3 (9, 12, 15), la biosíntesis del DHA involucra desaturación $\Delta 6$ hasta 18:4 (6, 9, 12, 15), seguido por alargamiento hasta 20:4 (8, 11, 14, 17) y desaturación $\Delta 5$ hasta 20:5 (5, 8, 11, 14, 17). Más allá de este punto, existen algunas controversias acerca de la biosíntesis. El punto de vista convencional es que 20:5 (5, 8, 11, 14, 17) se larga hasta 22:5 (7, 10, 13, 16, 19) y luego convertido hasta 22:6 (4, 7, 10, 13, 16, 19) por medio de la desaturación $\Delta 4$ final (Horrobin, D. F. (1992) Prog. Lipid Res. 31: 163 - 194). Sin embargo, Sprecher et al. recientemente sugirieron una ruta alternativa para la biosíntesis del DHA, que es independiente de la desaturasa $\Delta 4$, que involucra dos alargamientos consecutivos, una desaturación $\Delta 6$ y un acortamiento de dos carbonos a través de una β -oxidación limitada en los peroxisomas (Sprecher, H., et al. (1995) J. Lipid Res. 36: 2471 - 2477; Sprecher, H., et al. (1999) Lipids 34: S153 - S156).

50 La producción del DHA es importante debido a su efecto benéfico sobre la salud humana. Actualmente las fuentes principales del DHA son aceites de pescado y de algas. El aceite de pescado es una fuente principal y tradicional para este ácido graso, sin embargo, usualmente se oxida durante el tiempo que dura su venta. Además, el suministro del aceite es altamente variable y su fuente está en peligro con la disminución de las poblaciones de peces mientras que la fuente de algas es costosa debido al bajo rendimiento y alto costo de extracción.

55 EPA y AA son ambos ácidos grasos esenciales $\Delta 5$. Ellos forman una clase única de constituyentes de los alimentos y los piensos para humanos y animales. EPA pertenece a la serie n-3 con cinco enlaces dobles en la cadena de acilo, se encuentra en alimentos marinos, y es abundante en pescado azul del Atlántico Norte. AA pertenece a la serie n-6 con cuatro enlaces dobles. La falta de un enlace doble en la posición ω -3 confiere sobre AA diferentes propiedades a aquellas encontradas en EPA. Los eicosanoides producidos a partir de AA tienen propiedades inflamatorias y de agregación plaquetaria fuertes, mientras que aquellos derivados de EPA tienen propiedades antiinflamatorias y de agregación antiplaquetaria. AA pueden ser obtenidos a partir de algunos alimentos tales como carne, pescado y huevos, pero la concentración es baja.

El ácido gamma-linolénico (GLA) es otro ácido graso esencial encontrado en mamíferos. GLA es el intermediario metabólico para ácidos grasos n-6 de cadena muy larga y para diferentes moléculas activas. En mamíferos, la formación de ácidos grasos poliinsaturados de cadena larga está limitada en velocidad por desaturación $\Delta 6$. Muchas condiciones fisiológicas y patológicas tales como envejecimiento, estrés, diabetes, eczema, y algunas infecciones han demostrado que reducen la etapa de desaturación $\Delta 6$. Además, GLA es fácilmente catabolizado a partir de la oxidación y la rápida división celular asociada con ciertos trastornos, por ejemplo, cáncer o inflamación. Por lo tanto, la suplementación de la dieta con GLA puede reducir los riesgos de estos trastornos. Estudios clínicos han demostrado que la suplementación de la dieta con GLA es efectiva en el tratamiento de algunas condiciones patológicas tales como eczema atópico, síndrome premenstrual, diabetes, hipercolesterolemia, y trastornos inflamatorios y cardiovasculares.

Las fuentes predominantes de GLA son aceites de plantas tales como onagra (*Oenothera biennis*), borraja (*Borago officinalis* L.), grosella negra (*Ribes nigrum*), y a partir de microorganismos tales como *Mortierella* sp., *Mucor* sp., y *Cyanobacteria*. Sin embargo, esas fuentes de GLA no son ideales para suplementar la dieta debido a grandes fluctuaciones en la disponibilidad y los costos asociados con los procesos de extracción.

Los genes que codifican desaturasas, incluidas las delta 5 desaturasas, han sido clonados a partir de *Thraustochytrium* sp. (Qiu X. et al., J. Biol. Chem. 276 (2001), N°. 34, páginas 31561 - 31566) y otros organismos (ver por ejemplo WO 98/46765, WO 98/46764, WO 00/20602, WO 00/20603 y WO 99/33958).

Resumen de la invención

La biosíntesis de ácidos grasos es una actividad importante de plantas y microorganismos. Sin embargo, los humanos tienen una capacidad limitada para sintetizar ácidos grasos esenciales, por ejemplo, ácidos grasos poliinsaturados de cadena larga (LCPUFA). La biotecnología ha sido considerada durante mucho tiempo una forma eficiente para manipular el proceso de producción de ácidos grasos en plantas y microorganismos. Es rentable y renovable con pocos efectos secundarios. Por lo tanto, se han hecho tremendos esfuerzos industriales encaminados a la producción de diferentes compuestos incluidos especialmente ácidos grasos y polipéptidos farmacéuticos a través de la manipulación de células de plantas, animales, y microorganismos. Por lo tanto, la biotecnología es una vía atractiva para la producción de ácidos grasos insaturados, especialmente LCPUFA, en una forma rentable y segura con el fin de recoger el máximo valor terapéutico de estos ácidos grasos.

La presente invención se basa, a menos en parte, en el descubrimiento de una familia de moléculas de ácido nucleico que codifican nuevas desaturasas. En particular, los presentes inventores han identificado la Fad5-2 ($\Delta 5$ desaturasa), que está involucrada en la biosíntesis de ácidos grasos poliinsaturados de cadena larga DHA (ácido docosahexaenóico, 22:6, n-3) y DPA (ácido docosapentanoico, 22:5, n-6); más específicamente, Fad5-2 desaturado 20:4 (n-3) y 20:3 (n-6) que resulta en EPA y AA.

En un primer aspecto, la presente invención está dirigida a una molécula aislada de ácido nucleico seleccionada del grupo que consiste de:

(a) una molécula aislada de ácido nucleico que comprende la secuencia de nucleótidos expuesta en la SEQ ID NO: 5, o un complemento de la misma;

(b) una molécula aislada de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos expuesta en la SEQ ID NO: 6, o un complemento de la misma; y

(c) una molécula aislada de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que es al menos 98% idéntica a la secuencia completa de aminoácidos de la SEQ ID NO: 6, o un complemento de la misma, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa.

En un segundo aspecto, la presente invención está dirigida a un polipéptido aislado seleccionado del grupo que consiste de:

(a) un polipéptido aislado que comprende la secuencia de aminoácidos expuesta en la SEQ ID NO: 6;

(b) un polipéptido aislado codificado por una molécula de ácido nucleico que consiste de la secuencia de nucleótidos expuesta en la SEQ ID NO: 5, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa; y

(c) un polipéptido aislado que comprende una secuencia de aminoácidos que es al menos 98% idéntica a la secuencia completa de aminoácidos de la SEQ ID NO: 6, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa.

En una modalidad, la invención se caracteriza por una molécula aislada de ácido nucleico que incluye la secuencia de nucleótidos expuesta en la SEQ ID NO: 5. En otra modalidad, la invención se caracteriza por una molécula aislada de ácido nucleico que codifica un polipéptido que incluye la secuencia de aminoácidos expuesta en la SEQ ID NO: 6.

En el presente documento se describen moléculas aisladas de ácido nucleico incluidas las secuencias de nucleótidos que son sustancialmente idénticas (por ejemplo, 70% idénticas) a la secuencia de nucleótidos expuesta como la SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, o SEQ ID NO: 7. En el presente documento se describen moléculas aisladas de ácido nucleico que incluyen al menos 30 nucleótidos contiguos de la secuencia de nucleótidos expuesta como la SEQ ID NO: 5. En otra modalidad, la invención se caracteriza por moléculas aisladas de ácido nucleico que codifican un polipéptido incluida una secuencia de aminoácidos que es al menos 98% idéntica a la secuencia de aminoácidos expuesta como la SEQ ID NO: 6. También se describen moléculas de ácido nucleico que codifican variantes alélicas del polipéptido que tiene la secuencia de aminoácidos expuesta como la SEQ ID NO: 6. También se describen moléculas de ácido nucleico que codifican fragmentos, por ejemplo, fragmentos biológicamente activos, de los polipéptidos de longitud completa de la presente invención (por ejemplo, fragmentos que incluyen al menos 10 residuos aminoácidos contiguos de la secuencia de aminoácidos de la SEQ ID NO: 6. En aún otras modalidades, la invención se caracteriza por moléculas de ácido nucleico que son complementarias a, o que hibridan bajo condiciones estrictas a las moléculas aisladas de ácido nucleico descritas aquí.

En un aspecto relacionado, la invención provee vectores incluidas las moléculas aisladas de ácido nucleico descritas aquí (por ejemplo, moléculas de ácido nucleico que codifican desaturasa). También se presentan células huésped incluidos tales vectores (por ejemplo, células huésped incluidos vectores adecuados para la producción de moléculas de ácido nucleico para desaturasa y polipéptidos).

La invención se caracteriza por polipéptidos aislados de desaturasa. Las modalidades de ejemplo cuentan con un polipéptido que incluye la secuencia de aminoácidos expuesta como la SEQ ID NO: 6, un polipéptido que incluye una secuencia de aminoácidos al menos 98% idéntica a la secuencia de aminoácidos expuesta como la SEQ ID NO: 6, y a polipéptido codificado por una molécula de ácido nucleico que consiste de una secuencia de nucleótidos expuesta como la SEQ ID NO: 5. También se describen fragmentos de los polipéptidos de longitud completa descritos aquí (por ejemplo, fragmentos que incluyen al menos 10 residuos contiguos de aminoácidos de la secuencia expuesta como la SEQ ID NO: 6) así como variantes alélicas del polipéptido que tiene la secuencia de aminoácidos expuesta como la SEQ ID NO: 6.

En una modalidad, un polipéptido desaturasa tiene actividad de desaturasa. En otra modalidad, un polipéptido desaturasa tiene un motivo de enlazamiento del hemo del terminal N, por ejemplo, un dominio del tipo del citocromo b5 encontrado en el extremo frontal de las desaturasas. En otra modalidad, una polipéptido desaturasa tiene al menos dos, preferiblemente aproximadamente tres, motivos conservados de histidina encontrados en todas las desaturasas microsomaes y, opcionalmente, tiene actividad de desaturasa. En una modalidad preferida, la polipéptido desaturasa tiene aproximadamente tres motivos de histidina.

Las construcciones que contienen los genes para la desaturasa pueden ser utilizadas en cualquier sistema de expresión incluidas plantas, animales, y microorganismos para la producción de células capaces de producir los LCPUFA por ejemplo DHA, EPA, AA, SDA, y GLA. Los ejemplos de plantas usadas para expresar las desaturasas de la presente invención incluyen, entre otras, plantas y semillas de plantas de cultivos de semillas oleaginosas, por ejemplo, linaza (*Linum sp.*), colza (*Brassica sp.*), soja (*Glycine* y *Soja sp.*), girasol (*Helianthus sp.*), algodón (*Gossypium sp.*), maíz (*Zea mays*), olivo (*Olea sp.*), cártamo (*Carthamus sp.*), cacao (*Theobroma cacao*), y cacahuate (*Arachis sp.*).

En un aspecto relacionado, la presente invención proporciona métodos nuevos y mejorados para producir ácidos grasos insaturados, por ejemplo, los LCPUFA, y otros compuestos clave de la ruta biosintética del ácido graso insaturado utilizando células, por ejemplo, células vegetales; células animales, y/o células microbianas en las cuales se ha manipulado la ruta biosintética del ácido graso insaturado de tal manera que se produzcan los LCPUFA u otros compuestos deseados de ácido graso insaturado.

Las metodologías nuevas y mejoradas de la presente invención incluyen métodos para la producción de ácidos grasos insaturados (por ejemplo, DHA) en células que tienen al menos una desaturasa de ácido graso de la ruta biosintética del ácido graso insaturado manipulada de tal manera que se produzcan ácidos grasos insaturados (por ejemplo, producidos en mayor cantidad). Por ejemplo, la invención se caracteriza por métodos para la producción de un ácido graso insaturado (por ejemplo, DHA) en células que comprenden al menos molécula aislada de ácido nucleico que codifica para desaturasa, por ejemplo, Fad5-2, como se describió anteriormente, de tal manera que se produzca un ácido graso insaturado, por ejemplo, LCPUFA, por ejemplo, DHA. Tales métodos pueden incluir además la etapa de recuperación del LCPUFA.

En otra modalidad, la presente invención proporciona métodos para la producción ácidos grasos insaturados, por ejemplo, los LCPUFA, por ejemplo, DHA, que comprende poner en contacto una composición que comprende al menos una molécula objetivo de desaturasa, como se define en este documento, con al menos una polipéptido desaturasa aislada, por ejemplo, Fad5-2, como se describió anteriormente, bajo unas condiciones de tal manera que se produzca un ácido graso insaturado, por ejemplo, LCPUFA, por ejemplo, DHA. Tales métodos pueden incluir además la etapa de recuperar el LCPUFA.

El ácido nucleico, las proteínas, y los vectores descritos anteriormente son particularmente útiles en las metodologías de la presente invención. En particular, la invención se caracteriza por métodos para mejorar la

producción de ácido graso insaturado (por ejemplo, la producción del DHA) que incluye el cultivo de una planta, animal, y/o microorganismo recombinante que contiene un ácido nucleico que codifica para desaturasa, por ejemplo, Fad5-2, bajo condiciones de tal manera que se mejore la producción de ácido graso.

5 En otra modalidad, la presente invención presenta métodos para la producción una célula capaz de producir ácidos grasos insaturados. Tales métodos incluyen la introducción en una célula, por ejemplo, una célula vegetal, de una molécula aislada de ácido nucleico como se describió anteriormente que codifica una proteína que tiene la capacidad de catalizar la formación de un enlace doble en una cadena de acilo graso.

10 En otra modalidad, la presente invención presenta métodos para modular la producción de ácidos grasos que comprende el cultivo de una célula que contiene una molécula aislada de ácido nucleico como se describió anteriormente que codifica un polipéptido que tiene la capacidad de catalizar la formación de un enlace doble, de tal manera que se produzca la modulación de la producción de ácido graso.

15 En otra modalidad, la presente invención incluye composiciones que contienen los ácidos grasos insaturados, ácidos nucleicos o polipéptidos descritos aquí. Las composiciones de la presente invención también pueden incluir las células capaces de producir tales ácidos grasos, como se describió anteriormente, y, opcionalmente, un portador farmacéuticamente aceptable.

20 En otra modalidad, se usan las composiciones de la presente invención como un suplemento para la dieta, por ejemplo, en alimentación animal o como un nutracéutico. Las composiciones de la presente invención pueden permitir un método para tratamiento de un paciente que tiene un trastorno, que comprende la administración de la composición de tal manera que se trate al paciente. Los trastornos abarcados por tales métodos incluyen, por ejemplo, estrés, diabetes, cáncer, trastornos inflamatorios, y trastornos cardiovasculares.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada y de las reivindicaciones.

Breve descripción de los dibujos

25 La Figura 1 muestra la secuencia de ADN y de proteína de Fad5-2 de *Pythium irregulare*; (A) la secuencia de ADNc del marco de lectura abierto (SEQ ID NO: 5); y (B) la secuencia traducida de proteína (SEQ ID NO: 6).

La Figura 2 muestra la secuencia de ADN y de proteína de Fad6 de *Pythium irregulare*; (A) la secuencia de ADNc del marco de lectura abierto (SEQ ID NO: 7); y (B) la secuencia traducida de proteína (SEQ ID NO: 8).

30 La Figura 3 muestra una comparación de las secuencias de proteína de Fad5-2 y Fad6 de *Pythium irregulare* (SEQ ID NO: 6 y 8, respectivamente). La barra vertical indica la identidad del aminoácido. Los motivos conservados tales como los motivos ricos en histidina y el enlace hemo del citocromo b5 están resaltados. Las dos flechas indican las posiciones de los enlaces de los dos iniciadores degenerados.

La Figura 4 es un análisis por GC de los FAME de la cepa de levadura AMY2 α que expresa Fad5-2 con sustrato exógeno 18:1 - 9 (el panel superior) y 18:1 - 11 (el panel inferior), respectivamente.

35 La Figura 5 muestra la composición del ácido graso de tejidos vegetativos (hojas, tallos, y raíces) de una línea transgénica T1 con Fad5-2 bajo el control del promotor 35S. Los niveles de ácido graso se muestran como el porcentaje en peso de ácidos grasos totales en *B. juncea*.

La Figura 6 es un análisis por GC de los FAME de la raíz de *B. juncea* que expresan Fad5-2 con sustrato exógeno del ácido homo- γ -linolénico (HGLA, 20:3 - 8, 11, 14).

40 La Figura 7 es un análisis por GC de los FAME preparados a partir de semillas de *B. juncea* que expresan Fad5-2 bajo el control del promotor de napina.

La Figura 8 es una tabla que muestra el perfil de ácido graso de *Pythium irregulare*. La Figura 9 es una tabla que muestra la conversión de ácidos grasos exógenos en levadura AMY-2 α /pFad5-2.

45 La Figura 10 es una tabla que muestra la acumulación de ácidos grasos interrumpidos por polimetileno Δ 5 insaturados (Δ 5-UIFAs) en semillas de linaza transgénico que expresan Fad5-2 bajo el control de napina (Napina) y promotores específicos de la semilla de linaza (Cln). Los niveles de ácido graso se muestran como el porcentaje en peso de los ácidos grasos totales.

Descripción detallada de la invención

50 La presente invención se basa, al menos en parte, en el descubrimiento de nuevos miembros de la familia de desaturasas de ácido graso, denominadas aquí indistintamente como "desaturasas" o ácido nucleico que codifica para "desaturasa" y moléculas de proteína (por ejemplo, Fad5-2). Estas nuevas moléculas son miembros de la

familia de desaturasas de ácido graso y se expresan en organismos que producen los LCPUFA, por ejemplo *Thraustochytrium*, *Pythium irregulare*, *Schizichytrium*, y *Cryptocodinium*.

Como se lo utiliza aquí, el término "ácidos grasos" es reconocido en el arte e incluye ácido carboxílico con base en hidrocarburos de cadena larga. Los ácidos grasos son componentes de muchos lípidos incluidos glicéridos. Los ácidos grasos más comunes de ocurrencia natural son ácidos monocarboxílicos que tienen un número par de átomos de carbono (16 ó 18) y que pueden ser saturados o insaturados. Los ácidos grasos "insaturados" contienen enlaces dobles cis entre los átomos de carbono. Los ácidos grasos insaturados abarcados por la presente invención incluyen, por ejemplo, DHA, GLA, y SDA. Los ácidos grasos "poliinsaturados" contienen más de un doble enlace y los dobles enlaces están dispuestos en un sistema interrumpido de metileno ($-\text{CH}=\text{CH}-\text{CH}_2-\text{CH}=\text{CH}-$).

Los ácidos grasos se describen aquí por medio de un sistema de numeración en el cual el número antes de los dos puntos indica el número de átomos de carbono en el ácido graso, mientras que el número después de los dos puntos es el número de dobles enlaces que están presentes. En el caso de ácidos grasos insaturados, este es seguido por un número entre paréntesis que indica la posición de los dobles enlaces. Cada número entre paréntesis es el átomo de carbono con el número más bajo de los dos conectados por el doble enlace. Por ejemplo, se puede describir el ácido oleico como 18:1 (9) y se puede describir el ácido linoléico como 18:2 (9, 12) indicando 18 carbonos, un doble enlace en el carbono 9, dos dobles enlaces en los carbonos 9 y 12, respectivamente.

Las etapas de control en la producción de ácidos grasos insaturados, es decir, la ruta biosintética del ácido graso insaturado, son catalizadas por desaturasas de ácido graso asociadas a la membrana, por ejemplo, Fad5-2. Específicamente, tales enzimas catalizan la formación de dobles enlaces entre los átomos de carbono de una molécula de ácido graso. Como se lo utiliza aquí, el término "ruta biosintética del ácido graso insaturado" se refiere a una serie de reacciones químicas que conducen a la síntesis de un ácido graso insaturado ya sea *in vivo* o *in vitro*. Tal ruta incluye una serie de etapas de desaturación y alargamiento que generan ácidos grasos insaturados y por último, ácidos grasos poliinsaturados de cadena larga. Tales ácidos grasos insaturados pueden incluir, GLA 18:3 (6, 9, 12), SDA 18:4 (6, 9, 12, 15), AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), y DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

Las desaturasas pueden contener un motivo de enlazamiento de hemo y/o aproximadamente tres motivos conservados de histidina, aunque pueden estar presentes dominios adicionales. Los miembros de la familia de desaturasas de ácido graso convierten ácidos grasos saturados en ácidos grasos insaturados, por ejemplo, ácidos grasos poliinsaturados de cadena larga (los LCPUFA), que son componentes de membranas celulares de diferentes tejidos y organelos en mamíferos (células nerviosas, de retina, de cerebro e inmunes). Los ejemplos de LCPUFA incluyen, entre otros, ácido docosahexaenóico (DHA, 22:6 (4, 7, 10, 13, 16, 19)). Estudios clínicos han mostrado que DHA es esencial para el crecimiento y desarrollo del cerebro en bebés, y para el mantenimiento de la función normal del cerebro en adultos (Martinetz, M. (1992) J. Pediatr. 120: S129 - S138). DHA también tiene efectos sobre la función fotorreceptora involucrada en el proceso de transducción de la señal, activación de la rodopsina, y el desarrollo de conos y bastones (Giusto, N. M., y colaboradores (2000) Prog. Lipid Res. 39: 315 - 391). Además, se encontraron también efectos positivos del DHA en el tratamiento de enfermedades tales como hipertensión, artritis, aterosclerosis, depresión, trombosis y cánceres (Horrocks, L. A. y Yeo, Y. K. (1999) Pharmacol. Res. 40: 211 - 215). En consecuencia, se pueden utilizar las moléculas de desaturasa para producir los LCPUFA útiles en el tratamiento de trastornos caracterizados por crecimiento, proliferación, o diferenciación regulados en forma aberrante. Tales trastornos incluyen cáncer, por ejemplo, carcinoma, sarcoma, o leucemia; angiogénesis tumoral y metástasis; displasia esquelética; trastornos hepáticos; síndromes mielodisplásicos; y trastornos hematopoyéticos y/o mieloproliferativos. Otros trastornos relacionados con angiogénesis y que son, por lo tanto, trastornos asociados con desaturasa incluyen telangiectasia hemorrágica hereditaria tipo 1, fibrodisplasia osificante progresiva, fibrosis pulmonar idiopática, y síndrome de Klippel-Trenaunay-Weber.

El término "familia" cuando se refiere a las moléculas de ácido nucleico y proteína de la presente invención debe entenderse como dos o más moléculas de proteína o ácido nucleico que tienen un dominio estructural común o motivo y tienen suficiente homología de secuencia de aminoácidos o nucleótidos como se define aquí. Tales miembros de la familia pueden ser de ocurrencia natural o no natural y pueden ser ya sea de la misma o de una especie diferente. Por ejemplo, una familia puede contener una primera proteína de origen humano así como otras proteínas distintas de origen humano o alternativamente, puede contener homólogos de origen no humano, por ejemplo, proteínas de rata o de ratón. Los miembros de una familia pueden también tener características funcionales comunes.

Por ejemplo, la familia de proteínas desaturasa de la presente invención incluye un motivo de enlazamiento de hemo del citocromo b5. Como se utiliza aquí, el término "motivo de enlazamiento de hemo" es una extensión del terminal N del dominio del tipo del citocromo b5 encontrado en el extremo frontal de las desaturasas.

En otra modalidad, los miembros de la familia de las desaturasas de proteínas incluyen "motivos de histidina" en la proteína, preferiblemente, aproximadamente tres o cuatro motivos de histidina. Como se lo utiliza aquí, el término "motivo de histidina" incluye un dominio de proteína que tiene aproximadamente al menos dos residuos del aminoácido histidina, preferiblemente aproximadamente tres o cuatro residuos del aminoácido histidina, y típicamente se lo encuentra en todas las desaturasas microsomales como el tercer motivo conservador de histidina.

Los ejemplos de motivos de enlazamiento de hemo del citocromo b5 y los motivos de histidina incluyen los residuos aminoácidos 40 - 45, 171 - 176, 208 - 213, y 395 - 400 de la SEQ ID NO: 6, como se muestra en la Figura 3.

Las proteínas desaturasas aisladas de la presente invención tienen una secuencia de aminoácidos suficientemente homóloga a la secuencia de aminoácidos de la SEQ ID NO: 6 o son codificadas por una secuencia de nucleótidos suficientemente homóloga a la SEQ ID NO: 5. Las proteínas de la invención son aquellas descritas en la reivindicación 11. Como se lo utiliza aquí, el término "suficientemente homóloga" se refiere a una primera secuencia de nucleótidos o aminoácidos que contiene un número mínimo o suficiente de residuos aminoácidos o nucleótidos idénticos o equivalentes (por ejemplo, un residuo aminoácido que tiene una cadena lateral similar) a una segunda secuencia de nucleótidos o aminoácidos de tal manera que la primera y la segunda secuencias de aminoácidos o nucleótidos comparten dominios estructurales comunes o motivos y/o una actividad funcional común. Por ejemplo, secuencias de aminoácidos o nucleótidos que comparten dominios estructurales comunes que tienen al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más homología o identidad a través de las secuencias de aminoácidos de los dominios y contienen al menos uno y preferiblemente dos dominios estructurales o motivos, son definidos aquí como suficientemente homólogos. Además, secuencias de aminoácidos o nucleótidos que comparten al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más homología o identidad y comparten una actividad funcional común son definidas aquí como suficientemente homólogas.

En una modalidad preferida, una proteína desaturasa incluye al menos uno o más de los siguientes dominios o motivos: un motivo de enlazamiento de hemo y/o un motivo de histidina y tiene una secuencia de aminoácidos de aproximadamente al menos 98%, 99% o más homóloga o idéntica a la secuencia de aminoácidos de la SEQ ID NO: 6. En aún otra modalidad preferida, una proteína desaturasa incluye al menos uno o más de los siguientes dominios: un motivo de enlazamiento de hemo y/o un motivo de histidina, y es codificado por una molécula de ácido nucleico que tiene una secuencia de nucleótidos que hibrida bajo condiciones de hibridación rigurosas hasta un complemento de una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEQ ID NO: 5. En otra modalidad preferida, una proteína desaturasa incluye al menos un motivo de enlazamiento de hemo y/o al menos aproximadamente tres motivos de histidina, y tiene una actividad de desaturasa.

Como se lo utiliza indistintamente aquí, una "actividad de desaturasa", "actividad biológica de una desaturasa" o "actividad funcional de una desaturasa" incluye una actividad ejercida o mediada por una proteína desaturasa, molécula de polipéptido o ácido nucleico sobre una célula sensible a la desaturasa o sobre un sustrato de desaturasa, como se determina *in vivo* o *in vitro*, de acuerdo a técnicas estándar. En una modalidad, una actividad de desaturasa es una actividad directa tal como una asociación con una molécula objetivo de desaturasa. Como se lo utiliza aquí, una "molécula objetivo" o "pareja de enlazamiento" es una molécula por ejemplo, una molécula involucrada en la síntesis de ácidos grasos insaturados, por ejemplo, un ácido graso intermedio, con el cual se enlaza o interactúa una proteína desaturasa de una manera tal que se logra una función mediada por desaturasa. Una actividad desaturasa directa también incluye la formación de un doble enlace entre los átomos de carbono de una molécula de ácido graso para formar una molécula insaturada de ácido graso.

La secuencia de nucleótidos de la desaturasa $\Delta 5$ de *Pythium irregulare*, Fad5-2, el ADNc y la secuencia predicha de aminoácidos codificada por el ADNc de Fad4 son mostrados en la Figura 1 y en las SEQ ID NOs: 1 y 2, respectivamente. El gen para Fad5-2 de *Pythium irregulare* que tiene una longitud aproximada de 1371 nucleótidos, que tiene aproximadamente 456 residuos aminoácidos de longitud.

La secuencia de nucleótidos para la desaturasa $\Delta 6$ de *Pythium irregulare*, Fad6, el ADNc y la secuencia predicha de aminoácidos codificada por el ADNc de Fad6 son mostrados en la Figura 2 y en las SEQ ID NOs: 7 y 8, respectivamente. El gen para Fad6 de *Pythium irregulare*, que tiene una longitud aproximada de 1383 nucleótidos, codifica una proteína que tiene aproximadamente 460 residuos aminoácidos de longitud.

Se describen diferentes aspectos de la invención en forma más detallada en las siguientes subsecciones:

I. Moléculas Aisladas de Ácido Nucleico

Un aspecto de la invención se relaciona con moléculas aisladas de ácido nucleico que codifican proteínas desaturasa. Se describen aquí fragmentos de ácido nucleico suficientes para uso como sondas de hibridación para

identificar moléculas de ácido nucleico que codifican desaturasa (por ejemplo, ARNm desaturasa) y fragmentos para uso como iniciadores PCR para la amplificación o mutación de moléculas de ácido nucleico para desaturasa. Como se lo utiliza aquí, el término "molécula de ácido nucleico" está destinado a incluir moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo, ARNm) y análogos del ADN o ARN generado utilizando análogos de nucleótido. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferiblemente es ADN bicatenario.

El término "molécula aislada de ácido nucleico" incluye moléculas de ácido nucleico que están separadas de las otras moléculas de ácido nucleico que están presentes en la fuente natural del ácido nucleico. Por ejemplo, con relación al ADN genómico, el término "aislado" incluye moléculas de ácido nucleico que están separadas del cromosoma con el cual está naturalmente asociado el ADN genómico. Preferiblemente, un ácido nucleico "aislado" está libre de secuencias que naturalmente flanquean al ácido nucleico (es decir, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del organismo a partir del cual se deriva el ácido nucleico. Por ejemplo, en diferentes modalidades, la molécula aislada de ácido nucleico de la desaturasa puede contener aproximadamente menos de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ó 0,1 kb de secuencias de nucleótidos que naturalmente flanquean la molécula de ácido nucleico en ADN genómico de la célula a partir de la cual se deriva el ácido nucleico. Además, una molécula "aislada" de ácido nucleico, tal como una molécula de ADNc, puede estar sustancialmente libre de otro material celular, o medio de cultivo cuando es producido por medio de técnicas recombinantes, o sustancialmente libre de precursores químicos o de otros compuestos químicos cuando se la sintetiza químicamente.

Una molécula de ácido nucleico de la presente invención, por ejemplo, una molécula de ácido nucleico que tiene la secuencia de nucleótidos de la SEQ ID NO: 5, puede ser aislada utilizando técnicas estándar de biología molecular y la información de la secuencia suministrada aquí. Utilizando toda o una porción de la secuencia de ácido nucleico de la SEQ ID NO: 5, como sondas de hibridación, se pueden aislar moléculas de ácido nucleico de desaturasa utilizando técnicas estándar de hibridación y clonación (por ejemplo, como se describe en Sambrook, J. y colaboradores. *Molecular Cloning: A Laboratory Manual* 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Además, una molécula de ácido nucleico que abarca todo o una porción de la SEQ ID NO: 5, puede ser aislada por medio de la reacción en cadena de la polimerasa (PCR) utilizando iniciadores oligonucleótidos sintéticos diseñados con base en la secuencia de la SEQ ID NO: 5.

Se puede amplificar un ácido nucleico de la invención utilizando ADNc, ARNm o alternativamente, ADN genómico, como molde e iniciadores oligonucleótidos apropiados de acuerdo a técnicas estándar de amplificación por medio de PCR. El ácido nucleico así amplificado puede ser clonado en un vector apropiado y caracterizado por medio de análisis de secuencias de ADN. Además, se pueden preparar oligonucleótidos correspondientes a secuencias de nucleótido de desaturasa por medio de técnicas estándar de síntesis, por ejemplo, utilizando un sintetizador automático de DNA.

En otra modalidad, la molécula de ácido nucleico consiste de la secuencia de nucleótidos expuesta como la SEQ ID NO: 5.

En aún otra modalidad, una molécula aislada de ácido nucleico de la invención incluye una molécula de ácido nucleico que es un complemento de la secuencia de nucleótidos mostrada en la SEQ ID NO: 5. Una molécula de ácido nucleico que sea complementaria a la secuencia de nucleótidos mostrada en la SEQ ID NO: 5 es una que sea suficientemente complementaria a la secuencia de nucleótidos mostrada en la SEQ ID NO: 5, de tal manera que puede hibridar a la secuencia de nucleótidos mostrada en la SEQ ID NO: 5, formando así un dúplex estable.

Las moléculas de ácido nucleico que incluyen únicamente una porción de la secuencia de ácido nucleico de la SEQ ID NO: 5, por ejemplo, un fragmento, pueden ser utilizadas como sonda o iniciador. La secuencia de nucleótidos determinada a partir de la clonación del gen para la desaturasa permite la generación de sondas e iniciadores diseñados para ser utilizados en la identificación y/o clonación de otros miembros de la familia de la desaturasa, así como homólogos de desaturasa de otras especies. La sonda/iniciador (por ejemplo, oligonucleótido) incluye típicamente un oligonucleótido sustancialmente purificado. El oligonucleótido típicamente comprende una región de secuencia de nucleótidos que hibrida bajo condiciones rigurosas al menos aproximadamente hasta 12 ó 15, preferiblemente aproximadamente 20 ó 25, más preferiblemente aproximadamente 30, 35, 40, 45, 50, 55, 60, 65 ó 75 nucleótidos consecutivos de una secuencia sentido de la SEQ ID NO: 5, de una secuencia antisentido de la SEQ ID NO: 5, o de una variante alélica de ocurrencia natural o mutante de la SEQ ID NO: 5.

Los ejemplos de sondas o iniciadores son de al menos (o no más de) 12 ó 15, 20 ó 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 o más nucleótidos de longitud y/o incluyen nucleótidos consecutivos de una molécula aislada de ácido

nucleico descrita aquí. También se describen sondas o iniciadores que comprenden nucleótidos contiguos o consecutivos de una molécula aislada de ácido nucleico descrita aquí, pero para la diferencia de 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 bases dentro de la secuencia de la sonda o el iniciador. Se pueden utilizar sondas con base en las secuencias de nucleótidos de la desaturasa para detectar (por ejemplo, detectar específicamente) transcritos o secuencias genómicas que codifican las mismas proteínas o proteínas homologas. La sonda puede incluir además un grupo marcador unido a la misma, por ejemplo, el grupo marcador puede ser radioisótopo, un compuesto fluorescente, una enzima, o un cofactor de una enzima. También se describe un conjunto de iniciadores, por ejemplo, de iniciadores adecuados para uso en una PCR, que pueden ser utilizados para amplificar una región seleccionada de una secuencia de desaturasa, por ejemplo, un dominio, región, sitio u otra secuencia descrita aquí. Los iniciadores deben tener al menos 5, 10 ó 50 pares de bases de longitud y menos de 100, o menos de 200, pares de bases de longitud. Los iniciadores deben ser idénticos, o diferentes pero no superiores a 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 bases cuando se los compara con una secuencia descrita aquí o con la secuencia de una variante de ocurrencia natural. Tales sondas pueden ser utilizadas como parte de un kit para análisis de diagnostico para identificar células o tejido que expresa en forma incorrecta una proteína desaturasa, por ejemplo por medio de la medición del nivel de ácido nucleico que codifica una desaturasa en una muestra de células de un individuo, por ejemplo, detectando los niveles ARNm para la desaturasa o determinando si un gen genómico para la desaturasa ha sido mutado o suprimido.

Se puede preparar un fragmento de ácido nucleico que codifica una "porción biológicamente activa de una proteína desaturasa" por medio del aislamiento de una porción de la secuencia de nucleótidos de la SEQ ID NO: 1 ó 3, que codifica un polipéptido que tiene actividad biológica de desaturasa (las actividades biológicas de las proteínas desaturasa están descritas aquí), que expresa la porción codificada de la proteína desaturasa (por ejemplo, por medio de expresión recombinante *in vitro*) y evaluando la actividad de la porción codificada de la proteína desaturasa. En una modalidad de ejemplo, la molécula de ácido nucleico tiene al menos 50-100, 100-250, 250-500, 500-700, 750-1000, 1000-1250, 1250-1500, 1500-1750, 1750-2000, 2000-2250, 2250-2500, 2500-2750, 2750-3000, 3250-3500, 3500-3750 o más nucleótidos de longitud y codifica una proteína que tiene actividad de desaturasa, (como se describe aquí).

La invención abarca además moléculas de ácido nucleico que difieren de la secuencia de nucleótidos mostrada en la SEQ ID NO: 5, debido a degeneración del código genético y codifican por lo tanto las mismas proteínas desaturasa que aquellas codificadas por la secuencia de nucleótidos mostrada en la SEQ ID NO: 5. En otra modalidad, una molécula aislada de ácido nucleico de la invención tiene una secuencia de nucleótidos que codifica una proteína que tiene una secuencia de aminoácidos que difiere en al menos 1, pero no más de 5, 10, 20, 50 ó 100 residuos aminoácidos de la secuencia de aminoácidos mostrada en la SEQ ID NO: 6. En aún otra modalidad, la molécula de ácido nucleico codifica la secuencia de aminoácidos de desaturasa humana. Si se requiere una alineación para esta comparación, las secuencias deben ser alineadas para homología máxima.

Las variantes de ácido nucleico pueden ser de ocurrencia natural, por ejemplo variantes alélicas (mismo locus), homólogos (diferente locus), y ortólogos (diferente organismo) o pueden ser de ocurrencia no natural. Las variantes de ocurrencia no natural se pueden elaborar por medio de técnicas de mutagénesis, incluidas aquellas aplicadas a polinucleótidos, células, u organismos. Las variantes pueden contener sustituciones, supresiones, inversiones e inserciones de nucleótidos. La variación puede presentarse ya sea en una o en ambas regiones de codificación o no codificadora. Las variaciones pueden producir tanto sustituciones conservadoras como no conservadoras de aminoácidos (en comparación con el producto codificado).

Las variantes alélicas resultan, por ejemplo, de polimorfismos de la secuencia de ADN dentro de una población, (por ejemplo, la población humana) que conduce a cambios en las secuencias de aminoácidos de las proteínas desaturasa. Tal polimorfismo genético en los genes para la desaturasa puede existir entre individuos dentro de una población debido a una variación alélica natural.

Como se los utiliza aquí los términos "genes" y "gen recombinante" se refieren a moléculas de ácido nucleico que incluyen un marco de lectura abierto que codifica una proteína desaturasa, por ejemplo, proteína desaturasa de semilla oleaginosa, y pueden incluir además secuencias reguladoras no codificadoras, e intrones.

Se describen aquí moléculas aisladas de ácido nucleico que codifican una variante alélica de ocurrencia natural de un polipéptido que comprende la secuencia de aminoácidos de la SEQ ID NO: 6, en donde la molécula de ácido nucleico hibrida hasta un complemento de una molécula de ácido nucleico que comprende la SEQ ID NO: 5, por ejemplo, bajo condiciones rigurosas de hibridación.

Las variantes alélicas de la desaturasa, por ejemplo, Fad5-2, incluyen tanto proteínas desaturasa funcionales como no funcionales. Las variantes alélicas funcionales son variantes de la secuencia de aminoácidos de ocurrencia natural de la proteína desaturasa que mantienen la habilidad, por ejemplo, para (i) interactuar con un sustrato de desaturasa o molécula objetivo (por ejemplo, un ácido graso, por ejemplo, DRA); y/o (ii) formar un doble enlace entre

átomos de carbono en un sustrato de desaturasa o molécula objetivo. Los ácidos grasos producidos por las moléculas de ácido nucleico y de proteína de la presente invención son también útiles para el tratamiento de trastornos tales como el envejecimiento, el estrés, la diabetes, el cáncer, trastornos inflamatorios (por ejemplo, artritis, eczema), y trastornos cardiovasculares. Las variantes alélicas funcionales contendrán típicamente únicamente un sustitución conservadora de uno o más aminoácidos de la SEQ ID NO: 6, o una sustitución, supresión o inserción de residuos no críticos en regiones no críticas de la proteína.

Las variantes alélicas no funcionales son variantes de la secuencia de aminoácidos de ocurrencia natural de la proteína desaturasa, por ejemplo, Fad5-2, que no tienen la habilidad, por ejemplo, (i) interactuar con un sustrato de desaturasa o molécula objetivo (por ejemplo, un ácido graso intermediario, tal como 18:4 (6, 9, 12, 15)); y/o (ii) formar un doble enlace entre átomos de carbono en un sustrato de desaturasa o molécula objetivo. Las variantes alélicas no funcionales contendrán típicamente una sustitución no conservadora, una supresión, o inserción, o truncamiento prematuro de la secuencia de aminoácidos de la SEQ ID NO: 6, o una sustitución, inserción, o supresión en residuos críticos o en regiones críticas de la proteína.

La presente invención provee además ortólogos (por ejemplo, ortólogos humanos de las proteínas desaturasa). Los ortólogos de las proteínas desaturasa de *Pythium irregulare* son proteínas que se aíslan de otros organismos y poseen los mismos mecanismos de enlazamiento del sustrato de desaturasa o molécula objetivo, mecanismos de formación de dobles enlaces, mecanismos moduladores del crecimiento y desarrollo del cerebro en bebés, mecanismos de mantenimiento de la función normal del cerebro en adultos, la habilidad para afectar la función fotorreceptora involucrada en el proceso de transducción de la señal, la habilidad para afectar la activación de la rodopsina, mecanismos de desarrollo de conos y/o bastones, y/o mecanismos de modulación de crecimiento y/o proliferación celular de las proteínas desaturasa no humanas. Los ortólogos de las proteínas desaturasa de *Pythium irregulare* pueden ser fácilmente identificados por contener una secuencia de aminoácidos que es sustancialmente homóloga a la SEQ ID NO: 6.

Se describen aquí moléculas de ácido nucleico que codifican otros miembros de la familia de las desaturasas que tienen una secuencia de nucleótidos que difiere de las secuencias de la desaturasa de la SEQ ID NO: 5. Por ejemplo, se pueden identificar otro ADNc para desaturasa con base en la secuencia de nucleótidos de Fad5-2. Se describen aquí moléculas de ácido nucleico que codifican proteínas desaturasa de diferentes especies que tienen una secuencia de nucleótidos que difiere de las secuencias de la desaturasa de la SEQ ID NO: 5. Por ejemplo, se puede identificar el ADNc para la desaturasa de *Schizochytrium* o *Crythecodinium* con base en la secuencia de nucleótidos de una Fad5-2.

Las moléculas de ácido nucleico correspondientes a las variantes alélicas naturales y homólogos de los ADNc de desaturasa de la invención se pueden aislar con base en su homología con los ácidos nucleicos de la desaturasa descritos aquí utilizando los ADNc divulgados aquí, o una porción de los mismos, como sonda de hibridación de acuerdo con técnicas estándar de hibridación bajo condiciones estrictas de hibridación.

Se pueden identificar los ortólogos, homólogos y las variantes alélicas utilizando métodos conocidos en el arte (por ejemplo, por medio de hibridación con una molécula aislada de ácido nucleico de la presente invención, por ejemplo, bajo condiciones estrictas de hibridación). Se describe aquí una molécula aislada de ácido nucleico que tiene al menos 15, 20, 25, 30 o más nucleótidos de longitud y que hibrida bajo condiciones estrictas con la molécula de ácido nucleico que contiene la secuencia de nucleótidos de la SEQ ID NO: 1 ó 3. También se describe un ácido nucleico que tiene al menos 50-100, 100-250, 250-500, 500-700, 750-1000, 1000-1250, 1250-1500, 1500-1750, 1750-2000, 2000-2250, 2250-2500, 2500-2750, 2750-3000, 3250-3500, 3500-3750 o más nucleótidos de longitud.

Como se lo utiliza aquí, el término "hibrida bajo condiciones estrictas" pretende describir condiciones de hibridación y lavado bajo las cuales las secuencias de nucleótido que son significativamente idénticas u homólogas entre sí permanecen hibridadas entre sí. Preferiblemente, las condiciones son tales que las secuencias que son aproximadamente al menos 70%, más preferiblemente aproximadamente 80%, incluso más preferiblemente aproximadamente a menos 85% ó 90% idénticas entre sí permanecen hibridadas entre sí. Tales condiciones estrictas son conocidas por aquellos capacitados en el arte y pueden encontrarse en Current Protocols in Molecular Biology, Ausubel y colaboradores, eds., John Wiley & Sons, Inc. (1995), secciones 2, 4, y 6. Se pueden encontrar condiciones estrictas adicionales en Molecular Cloning: A Laboratory Manual, Sambrook y colaboradores, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), capítulos 7, 9, y 11. Un ejemplo preferido no limitante de condiciones estrictas de hibridación incluye hibridación en 4X cloruro de sodio/citrato de sodio (SSC), aproximadamente a 65 - 70°C (o alternativamente hibridación en 4X SSC más 50% de formamida aproximadamente a 42 - 50°C) seguido por uno o más lavados en 1X SSC, aproximadamente a 65 - 70°C. Un ejemplo preferido, no limitante de condiciones de hibridación altamente estrictas incluye hibridación en 1X SSC, aproximadamente a 65 - 70°C (o alternativamente hibridación en 1X SSC más 50% de formamida aproximadamente a 42 - 50°C) seguido por uno o más lavados en 0,3X SSC, aproximadamente a 65 - 70°C. Un ejemplo preferido no limitante de condiciones de hibridación poco estrictas incluyen hibridación en 4X SSC, aproximadamente a 50 - 60°C (o alternativamente

hibridación en 6X SSC más 50% de formamida aproximadamente a 40 - 45°C) seguido por uno o más lavados en 2X SSC, aproximadamente a 50 - 60°C. Los rangos intermedios de los valores anteriormente mencionados, por ejemplo, a 65-70°C o a 42 - 50°C también son abarcados por la presente invención. Se puede sustituir SSPE (1X SSPE es NaCl 0,15 M, NaH₂PO₄ 10 mM, y EDTA 1,25 mM, pH 7,4) por SSC (1X SSC es NaCl 0,15M y citrato de sodio 15 mM) en los reguladores de hibridación y lavado; los lavados se llevan a cabo durante 15 minutos cada vez después de completar la hibridación. La temperatura de hibridación para los híbridos que se anticipó eran menores a 50 pares de bases de longitud debe ser 5 - 10°C menor que la temperatura de fusión (T_m) del híbrido, donde T_m de acuerdo a las ecuaciones siguientes. Para híbridos de menos de 18 pares de bases de longitud, T_m (°C) = 2(# de A + T bases) + 4(# de G + C bases). Para híbridos entre 18 y 49 pares de bases de longitud, T_m (°C) = 81,5 + 16,6 (log₁₀[Na⁺]) + 0,41 (% de G + C) - (600/N), donde N es el número de bases en el híbrido, y [Na⁺] es la concentración de iones sodio en el regulador de hibridación ([Na⁺] para 1X SSC = 0,165 M). La persona capacitada en el arte se dará cuenta también que pueden añadirse reactivos adicionales para hibridación y/o reguladores de lavado para disminuir la hibridación no específica de moléculas de ácido nucleico con membranas, por ejemplo, membranas de nitrocelulosa o nylon, incluyendo pero sin limitarse a agentes de bloqueo (por ejemplo, BSA o esperma de salmón o de arenque portador de ADN), detergentes (por ejemplo, SDS), agentes de quelación (por ejemplo, EDTA), Ficoll, PVP y similares. Cuando se utilizan membranas de nylon en particular, un ejemplo preferido adicional no limitante de condiciones estrictas de hibridación es la hibridación en NaH₂PO₄ 0,25 - 0,5 M, 7% de SDS aproximadamente a 65°C, seguido por uno o más lavados con NaH₂PO₄ 0,02 M, 1% de SDS a 65°C (ver por ejemplo, Church y Gilbert (1984) Proc. Natl. Acad. Sci. USA 81: 1991 - 1995), o alternativamente 0,2X SSC, 1% de SDS.

Preferiblemente, una molécula aislada de ácido nucleico de la invención que hibrida bajo condiciones estrictas con la secuencia de la SEQ ID NO: 5, corresponde a una molécula de ácido nucleico de ocurrencia natural. Como se lo utiliza aquí, una molécula de ácido nucleico "de ocurrencia natural" se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que se presenta en la naturaleza (por ejemplo, que codifica una proteína natural).

Además de las variantes alélicas de ocurrencia natural de las secuencias de desaturasa que pueden existir en la población, la persona capacitada se dará cuenta además que pueden introducirse cambios por medio de mutación en las secuencias de nucleótidos de las SEQ ID NO: 5, conduciendo por lo tanto a cambios en la secuencia de aminoácidos de las proteínas desaturasa codificadas, sin alterar la habilidad funcional de las proteínas desaturasa. Por ejemplo, pueden hacerse sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en residuos aminoácidos "no esenciales" en la secuencia de la SEQ ID NO: 5. Un residuo aminoácido "no esencial" es un residuo que puede ser alterado a partir de la secuencia de tipo silvestre de Fad5-2 (por ejemplo, la secuencia de la SEQ ID NO: 6) sin alterar la actividad biológica, mientras que se requiere un residuo aminoácido "esencial" para la actividad biológica. Por ejemplo, los residuos aminoácidos que se conservan entre las proteínas desaturasa de la presente invención, por ejemplo, aquellos presentes en un motivo de enlazamiento hemo o un motivo de histidina, se predice que son particularmente no sensibles a alteración. Además, los residuos aminoácidos adicionales que se conservan entre las proteínas desaturasa de la presente invención y otros miembros de la familia de desaturasas de ácido graso probablemente no son sensibles a alteración.

Por lo tanto, otro aspecto de la invención se relaciona con moléculas de ácido nucleico que codifican proteínas desaturasa que contienen cambios en los residuos aminoácidos que no son esenciales para la actividad. Tales proteínas desaturasa difieren en la secuencia de aminoácidos de la SEQ ID NO: 6 u 8, reteniendo aún la actividad biológica. En una modalidad, la molécula aislada de ácido nucleico incluye una secuencia de nucleótidos que codifica una proteína, en donde la proteína incluye una secuencia de aminoácidos que es al menos 98% 99% o más homóloga con la SEQ ID NO: 6, por ejemplo, con la longitud completa de la SEQ ID NO: 6.

Se puede crear una molécula aislada de ácido nucleico que codifica una proteína desaturasa homóloga a la proteína de la SEQ ID NO: 6, por medio de la introducción de una o más sustituciones, adiciones o supresiones de nucleótidos dentro de la secuencia de nucleótidos de la SEQ ID NO: 5, de tal manera que se introduzcan una o más sustituciones, adiciones o supresiones de aminoácidos dentro de la proteína codificada. Se pueden introducir mutaciones en la SEQ ID NO: 5, por medio de técnicas estándar, tales como mutagénesis dirigida al sitio y mutagénesis mediada por PCR. Preferiblemente se hacen sustituciones conservadoras de aminoácidos en uno o más residuos predichos de aminoácidos no esenciales. Una "sustitución conservadora de aminoácidos" es una en la cual el residuo aminoácido es remplazado con un residuo aminoácido que tiene una cadena lateral similar. Las familias de residuos aminoácidos que tienen cadenas laterales similares han sido definidas en el arte. Estas familias incluyen 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, triptófano), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina), cadenas laterales beta-ramificadas (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). De este modo, un residuo predicho de aminoácido no esencial en una proteína desaturasa es preferiblemente remplazado con otro residuo aminoácido de la misma familia de cadena lateral. Alternativamente, se puede introducir mutaciones aleatorias a lo largo de toda o parte de una secuencia de codificación de desaturasa, tal como

por medio de mutagénesis por saturación, y se pueden seleccionar los mutantes resultantes por la actividad biológica de la desaturasa para identificar mutantes que retengan actividad. Después de la mutagénesis de la SEQ ID NO: 5, se puede expresar la proteína codificada en forma recombinante y se puede determinar la actividad de la proteína.

- 5 Se puede analizar una proteína desaturasa mutante por la habilidad para (i) interactuar con un sustrato de desaturasa o una molécula objetivo (por ejemplo, un ácido graso intermedio); y/o (ii) formar un doble enlace entre átomos de carbono en un sustrato de desaturasa o una molécula objetivo.

II. Proteínas Desaturasa Aisladas

- 10 Un aspecto de la invención se relaciona con proteínas desaturasa recombinantes o aisladas y polipéptidos, y con porciones biológicamente activas de los mismos. En una modalidad, se pueden aislar proteínas desaturasa nativas de fuentes de células o tejidos por medio de un esquema apropiado de purificación utilizando técnicas estándar para purificación de proteínas. En otra modalidad, se producen proteínas desaturasa por medio de técnicas de ADN recombinante. En forma alternativa a la expresión recombinante, se puede sintetizar químicamente una proteína desaturasa o polipéptido utilizando técnicas estándar para síntesis de péptidos.

- 15 Una proteína "aislada" o "purificada" o una porción biológicamente activa de la misma está sustancialmente libre de material celular o de otras proteínas recombinantes de la fuente de células o tejidos a partir de la cual se deriva la proteína desaturasa, o sustancialmente libre de precursores químicos u otros compuestos químicos cuando se la sintetiza químicamente. La expresión "sustancialmente libre de material celular" incluye preparaciones de proteína desaturasa en las cuales se separa la proteína de componentes celulares de las células a partir de las cuales se
20 aísla o se la produce en forma recombinante. En una modalidad, la expresión "sustancialmente libre de material celular" incluye preparaciones de proteína desaturasa que tienen aproximadamente menos de 80%, 70%, 60%, 50%, 40%, o 30% (en peso seco) de proteína no desaturasa (también denominada aquí como una "proteína contaminante"), más preferiblemente aproximadamente menos del 20% de proteína no desaturasa, aún más preferiblemente menos del 10% de proteína no desaturasa, y lo más preferible aproximadamente menos del 5% de proteína no desaturasa. Cuando la proteína desaturasa o una porción biológicamente activa de la misma es
25 producida en forma recombinante, está también preferiblemente sustancialmente libre de medio de cultivo, es decir, que el medio de cultivo representa aproximadamente menos del 20%, más preferiblemente aproximadamente menos del 10%, y lo más preferible aproximadamente menos del 5% del volumen de la preparación de proteína.

- 30 La expresión "sustancialmente libre de precursores químicos o de otros compuestos químicos" incluye preparaciones de proteína desaturasa en las cuales la proteína está separada de los precursores químicos o de otros compuestos químicos que están involucrados en la síntesis de la proteína. En una modalidad, la expresión "sustancialmente libre de precursores químicos o de otros compuestos químicos" incluye preparaciones de proteína desaturasa que tiene aproximadamente menos del 30% (en peso seco) de precursores químicos o de productos
35 químicos que no son de desaturasa, más preferiblemente aproximadamente menos del 20% de precursores químicos o de productos químicos que no son de desaturasa, aún más preferiblemente aproximadamente menos del 10% de precursores químicos o de productos químicos que no son de desaturasa, y lo más preferible aproximadamente menos del 5% de precursores químicos o de productos químicos que no son de desaturasa. Debe entenderse que las proteínas o esta invención pueden estar también en una forma que sea diferente que sus correspondientes proteínas de ocurrencia natural y/o que estén asociadas con al menos algunos componentes
40 celulares. Por ejemplo, la proteína puede estar asociada con una membrana celular.

- Como se lo utiliza, una "porción biológicamente activa" de una proteína desaturasa incluye un fragmento de una proteína desaturasa que participa en una interacción entre una molécula de desaturasa y una molécula que no es de desaturasa (por ejemplo, un sustrato de desaturasa tal como ácido graso). Las porciones biológicamente activas de una proteína desaturasa incluyen péptidos que contienen secuencias de aminoácidos suficientemente homólogas a
45 o derivadas de las secuencias de aminoácidos de desaturasa, por ejemplo, las secuencias de aminoácidos mostradas en la SEQ ID NO: 6 que incluyen suficientes residuos aminoácidos para exhibir al menos una actividad de una proteína desaturasa. Típicamente, las porciones biológicamente activas contienen un dominio o motivo con al menos una actividad de la proteína desaturasa; la habilidad para (i) interactuar con un sustrato de desaturasa o molécula objetivo (por ejemplo, un ácido graso intermedio); y/o (ii) formar un enlace doble entre átomos de carbono
50 en un sustrato de desaturasa o molécula objetivo. Una porción biológicamente activa de una proteína desaturasa puede ser un polipéptido que sea, por ejemplo, de 10, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200 o más aminoácidos de longitud.

- Una porción biológicamente activa de una proteína desaturasa contiene un motivo de enlazamiento de hemo y/o al menos motivos de histidina, preferiblemente aproximadamente tres motivos de histidina. Además, se pueden
55 preparar otras porciones biológicamente activas, en las cuales se suprimen otras regiones de la proteína, por medio de técnicas recombinantes y evaluar una o más de las actividades funcionales de una proteína desaturasa nativa.

En una modalidad preferida, una proteína desaturasa tiene una secuencia de aminoácidos mostrada en la SEQ ID NO: 6. En otras modalidades, la proteína desaturasa es sustancialmente idéntica a la SEQ ID NO: 6 y retiene la actividad funcional de la proteína de la SEQ ID NO: 6, incluso difiere en la secuencia de aminoácidos debido a una variación alélica natural o mutagénesis, como se describe en detalle en la subsección I anterior. En otra modalidad, la proteína desaturasa es una proteína que contiene una secuencia de aminoácidos al menos 98%, 99% o más idéntica a la SEQ ID NO: 6.

En otra modalidad, la invención se caracteriza por una proteína desaturasa que es codificada por una molécula de ácido nucleico que consiste de la secuencia e nucleótidos de la SEQ ID NO: 5, o un complemento de la misma. También se describe aquí una proteína desaturasa que es codificada por una molécula de ácido nucleico que consiste de una secuencia de nucleótidos que hibrida bajo condiciones de hibridación estrictas a un complemento de una molécula de ácido nucleico que contiene la secuencia de nucleótidos de la SEQ ID NO: 5, o un complemento de la misma.

Para determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácido nucleico, se alinean las secuencias con el propósito de una comparación óptima (por ejemplo, se pueden introducir espacios en una o en ambas de una primera y una segunda secuencia de aminoácidos o de ácido nucleico para alineación óptima y se pueden ignorar secuencias no homólogas para propósitos de comparación). La longitud de una secuencia de referencia alineada para propósitos de comparación es al menos 30%, preferiblemente al menos 40%, más preferiblemente al menos 50%, incluso más preferiblemente al menos 60%, e incluso más preferiblemente al menos 70%, 80%, ó 90% de la longitud de la secuencia preferida (por ejemplo, cuando se alinea una segunda secuencia con la secuencia de aminoácidos de Fad5-2 de la SEQ ID NO: 6 que tiene 456 residuos aminoácidos, al menos 137, preferiblemente al menos 182, más preferiblemente al menos 228, incluso más preferiblemente al menos 273, e incluso más preferiblemente se alinean al menos 319, 365, ó 419 residuos aminoácidos. Se comparan luego los residuos aminoácidos o nucleótidos con las correspondientes posiciones de aminoácidos o posiciones de nucleótidos positions. Cuando una posición en la primera secuencia está ocupada por el mismo residuo aminoácido o nucleótido que la correspondiente posición en la segunda secuencia, entonces las moléculas son idénticas en esa posición (como se utiliza aquí, "identidad" de aminoácido o ácido nucleico es equivalente a "homología" de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos, y la longitud de cada hueco, que deben ser introducidos para una alineación óptima de las dos secuencias.

La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias pueden ser logradas utilizando un algoritmo matemático. En una modalidad preferida, el porcentaje de identidad entre dos secuencias de aminoácidos se determina utilizando el algoritmo de Needleman y Wunsch (J. Mol. Biol. (48): 444 - 453 (1970)) que ha sido incorporado en el programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), utilizando ya sea una matriz Blossum 62 o una matriz PAM250; y un peso de hueco de 16, 14, 12, 10, 8, 6, ó 4 y un peso de longitud de 1, 2, 3, 4, 5, ó 6. En aún otra modalidad preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina utilizando el programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), utilizando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70, ó 80 y un peso de longitud de 1, 2, 3, 4, 5, ó 6. Un ejemplo no limitante preferido de parámetros para ser utilizados junto con el programa GAP incluyen una matriz de puntuación Blosum 62 con una penalización por hueco de 12, una penalización extendida por hueco de 4, y una penalización por hueco por cambio de marco de 5.

En otra modalidad, el porcentaje de identidad entre dos aminoácidos o secuencias de nucleótidos se determina utilizando el algoritmo de Meyers y Miller (Comput. Appl. Biosci., 4: 11 - 17 (1988)) que ha sido incorporado en el programa ALIGN (versión 2.0 o versión 2.0U), utilizando una tabla de residuos de peso PAM120, una penalización por longitud de hueco de 12 y una penalización por hueco de 4.

Las secuencias de ácido nucleico y de proteína de la presente invención pueden ser utilizadas adicionalmente como una "secuencia de consulta" para llevar a cabo una búsqueda contra bases de datos públicas, por ejemplo, para identificar otros miembros de la familia o secuencias relacionadas. Tales búsquedas pueden ser llevadas a cabo utilizando los programas NBLAST y XBLAST (versión 2.0) de Altschul et al. (1990) J. Mol. Biol. 215: 403 - 10. Las búsquedas de nucleótidos por BLAST pueden ser llevadas a cabo con el programa NBLAST, puntuación = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas a moléculas de ácido nucleico para desaturasa de la invención. Las búsquedas de proteína por BLAST pueden ser llevadas a cabo con el programa XBLAST, puntuación = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas a moléculas de proteína desaturasa de la invención. Para obtener alineaciones lineales para propósitos de comparación, se puede utilizar Gapped BLAST como se describe en Altschul et al. (1997) Nucleic Acids Res. 25(17): 3389 - 3402. Cuando se utilizan los programas BLAST y Gapped BLAST, se pueden usar los parámetros predeterminados de los programas respectivos (por ejemplo, XBLAST y NBLAST). Ver <http://www.ncbi.nlm.nih.gov>.

III. Métodos para producir ácidos grasos insaturados

La presente invención provee métodos nuevos y mejorados para la producción de ácidos grasos insaturados, por ejemplo, los LCPUFA, por ejemplo, DHA (ácido docosahexaenóico, 22:6 (n-6)), DPA (ácido docosapentaenóico, 22:5 (n-6)), AA (ácido araquidónico, 20: 4 (n-6)) y EPA (ácido eicosapentaenóico, 20:5(n-3)).

5 A. Células recombinantes y métodos para cultivar células

La presente invención presenta adicionalmente vectores recombinantes que incluyen secuencias de ácido nucleico que codifican los productos génicos tal como se describe en este documento, preferiblemente los productos génicos Fad4, Fad5, Fad5-2, y Fad6. El término vector recombinante incluye un vector (por ejemplo, plásmido) que ha sido alterado, modificado o tratado por ingeniería genética de tal manera que contenga más, menos o diferentes secuencias de ácido nucleico que aquellas incluidas en el vector nativo o plásmido. En una modalidad, un vector recombinante incluye la secuencia de ácido nucleico que codifica al menos una enzima desaturasa de ácido graso operativamente enlazada a secuencias reguladoras. La frase "operativamente enlazada a una secuencia(s) reguladora(s)" significa que la secuencia de nucleótidos de interés está enlazada a la(s) secuencia(s) reguladora(s) en una forma que permita la expresión (por ejemplo, expresión mejorada, incrementada, constitutiva, basal, atenuada, disminuida o reprimida) de la secuencia de nucleótidos, preferiblemente expresión de un producto génico codificado por la secuencia de nucleótidos (por ejemplo, cuando se introduce el vector recombinante en una célula). Los ejemplos de vectores son descritos aquí con más detalle así como, por ejemplo, en Frascotti et al., Patente Estadounidense No. 5.721.137.

El término "secuenciador regulador" incluye secuencias de ácido nucleico que afectan (por ejemplo, modulan o regulan) la expresión de otras secuencias (no reguladoras) de ácido nucleico. En una modalidad, una secuencia reguladora está incluida en un vector recombinante en una posición similar o idéntica y/o orientación relativa a un gen particular de interés como se observa para la secuencia reguladora y el gen de interés tal y como se presenta en la naturaleza, por ejemplo, en una posición y/o orientación nativa. Por ejemplo, se puede incluir un gen de interés (por ejemplo, un gen para Fad5-2) en un vector recombinante operativamente enlazado a una secuencia reguladora que acompaña o es adyacente al gen en el organismo natural (por ejemplo, operativamente enlazado a una secuencia reguladora "nativa" de Fad5-2 (por ejemplo, al promotor "nativo" de Fad5-2). Alternativamente, un gen de interés (por ejemplo, un gen para Fad5-2) puede ser incluido en un vector recombinante operativamente enlazado a una secuencia reguladora que acompaña o es adyacente a otro gen (por ejemplo, uno diferente) en el organismo natural. Por ejemplo, un gen para Fad5-2 puede ser incluido en un vector operativamente enlazado a secuencias reguladoras que no son de Fad4, Fad5, Fad5-2, o Fad6. Alternativamente, se puede incluir un gen de interés (por ejemplo, un gen para Fad5-2) en un vector operativamente enlazado a una secuencia reguladora de otro organismo. Por ejemplo, se pueden enlazar operativamente secuencias reguladoras de otros microbios (por ejemplo, otras secuencias reguladoras bacterianas, secuencias reguladoras de bacteriófagos y similares) a un gen particular de interés.

Las secuencias reguladoras preferidas incluyen promotores, reforzadores, señales de terminación y otros elementos de control de la expresión (por ejemplo, sitios de enlazamiento para proteínas reguladoras transcripcionales y/o de traducción, por ejemplo, en el ARNm transcrito). Tales secuencias reguladoras están descritas, por ejemplo, en Sambrook, J., Fritsh, E. F., y Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. Las secuencias reguladoras incluyen a aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en una célula (por ejemplo, promotores constitutivos y promotores constitutivos fuertes), a aquellas que dirigen expresión inducible de una secuencia de nucleótidos en una célula (por ejemplo, promotores inducibles, por ejemplo, promotores inducibles de xilosa) y aquellas que atenúan o reprimen la expresión de una secuencia de nucleótidos en una célula (por ejemplo, señales de atenuación o secuencias represoras). También se encuentra dentro del alcance de la presente invención regular la expresión de un gen de interés por medio de la remoción o supresión de secuencias reguladoras. Por ejemplo, se pueden remover secuencias involucradas en la regulación negativa de transcripción de tal manera que se refuerza la expresión de un gen de interés.

En una modalidad, un vector recombinante de la presente invención incluye secuencias de ácido nucleico que codifican al menos un producto génico (por ejemplo, Fad5-2) operativamente enlazadas a un promotor o a una secuencia promotora.

En aún otra modalidad, un vector recombinante de la presente invención incluye una secuencia terminadora o secuencias terminadoras (por ejemplo, secuencias terminadoras de la transcripción). El término "secuencias terminadoras" incluye secuencias reguladoras que sirven para terminar la transcripción del ARNm. Las secuencias terminadoras (o los terminadores en tándem de la transcripción) pueden servir además para estabilizar el ARNm (por ejemplo, por medio de la adición de estructura al ARNm), por ejemplo, contra nucleasas.

En aún otra modalidad, un vector recombinante de la presente invención incluye secuencias de resistencia a los antibióticos. El término "secuencias de resistencia a los antibióticos" incluye secuencias que promueven o confieren resistencia a los antibióticos a los organismos huésped. En una modalidad, las secuencias de resistencia a los antibióticos se seleccionan del grupo que consiste de secuencias *cat* (resistencia al cloranfenicol), *tet* (resistencia a

la tetraciclina), secuencias *erm* (resistencia a la eritromicina), secuencias *neo* (resistencia a la neomicina) y secuencias *spec* (resistencia a la espectinomina). Los vectores recombinantes de la presente invención pueden incluir adicionalmente secuencias de recombinación homólogas (por ejemplo, secuencias diseñadas para permitir la recombinación del gen de interés dentro del cromosoma del organismo huésped). Por ejemplo, se pueden utilizar secuencias *amyE* como objetivos de homología para recombinación dentro del cromosoma huésped.

El término "célula manipulada" incluye una célula que ha sido modificada por ingeniería (por ejemplo, modificada por medio de ingeniería genética) o modificada de tal manera que la célula tiene al menos una desaturasa de ácido graso, por ejemplo, Fad5-2, de tal manera que se produce un ácido graso insaturado. La modificación o tratamiento por ingeniería genética de tales microorganismos puede ser de acuerdo con la metodología descrita aquí incluyendo, pero sin limitarse a, desregulación de una ruta biosintética y/o sobreexpresión de al menos una enzima biosintética. Una enzima "manipulada" (por ejemplo, una enzima biosintética "manipulada") incluye una enzima, cuya expresión o producción ha sido alterada o modificada de tal manera que se altera o modifica al menos un precursor, sustrato o producto secuencia arriba o secuencia debajo de la enzima, por ejemplo, comparado con la correspondiente enzima de origen natural o de tipo silvestre.

El término "sobreexpresado" o "sobreexpresión" incluye la expresión de un producto génico (por ejemplo, una desaturasa de ácido graso) a un nivel mayor que aquella expresada antes de la manipulación de la célula o en una célula comparable que no ha sido manipulada. En una modalidad, la célula puede ser genéticamente manipulada (por ejemplo, modificada por ingeniería genética) para sobreexpresar un nivel de producto génico superior a aquella expresada antes de la manipulación de la célula o en una célula comparable que no ha sido manipulada. La manipulación genética puede incluir, pero no se limita a, alterar o modificar secuencias reguladoras o sitios asociados con la expresión de un gen particular (por ejemplo, por medio de la adición de promotores fuertes, promotores inducibles o promotores múltiples o por medio de la remoción de secuencias reguladoras de tal manera que la expresión es constitutiva), modificando la ubicación cromosómica de un gen particular, alterando las secuencias de ácido nucleico adyacentes a un gen particular tal como un sitio de enlazamiento del ribosoma o terminador de la transcripción, incrementando el número de copias de un gen particular, modificando proteínas (por ejemplo, proteínas reguladoras, supresores, reforzadores, activadores transcripcionales y similares) involucradas en la transcripción de un gen particular y/o la traducción de un producto génico particular, o cualquier otro medio convencional de desregulación de la expresión de genes rutinario en el arte (incluyendo pero sin limitarse al uso de moléculas antisentido de ácido nucleico, por ejemplo, para bloquear la expresión de proteínas represoras).

En otra modalidad, la célula puede ser física o ambientalmente manipulada para sobreexpresar un nivel de producto génico superior a aquel expresado antes de la manipulación de la célula o en una célula comparable que no ha sido manipulada. Por ejemplo, una célula puede ser tratada con o cultivada en presencia de un agente conocido o que se sospecha que incrementa la transcripción de un gen particular y/o la traducción de un producto génico particular de tal manera que se refuerzan o se incrementan la transcripción y/o la traducción. Alternativamente, una célula puede ser cultivada a una temperatura seleccionada para incrementar la transcripción de un gen particular y/o la traducción de un producto génico particular de tal manera que se refuerzan o se incrementan la transcripción y/o la traducción.

El término "desregulado" o "desregulación" incluyen la alteración o la modificación de al menos un gen en una célula que codifica una enzima en una ruta biosintética, de tal manera que se altera o modifica el nivel o la actividad de la enzima biosintética en la célula. Preferiblemente, se altera o modifica al menos un gen que codifica una enzima en una ruta biosintética de tal manera que se refuerza o incrementa el producto génico. La frase "ruta desregulada" puede incluir también una ruta biosintética en la cual se altera o modifica más de un gen que codifica una enzima en una ruta biosintética de tal manera que se altera o modifica el nivel o actividad de más de una enzima biosintética. La habilidad para "desregular" una ruta (por ejemplo, para desregular simultáneamente más de un gen en una ruta biosintética dada) en una célula surge del fenómeno particular de células en las cuales más de una enzima (por ejemplo, dos o tres enzimas biosintéticas) son codificadas por genes adyacentes entre sí sobre una pieza contigua de material genético llamada un "operón".

El término "operón" incluye una unidad coordinada de expresión génica que contiene un promotor y posiblemente un elemento regulador asociado con uno o más, preferiblemente al menos dos, genes estructurales (por ejemplo, genes que codifican enzimas, por ejemplo, enzimas biosintéticas). La expresión de los genes estructurales puede ser regulada en forma coordinada, por ejemplo, por medio del enlazamiento de proteínas reguladoras al elemento regulador o por medio de anti-terminación de la transcripción. Los genes estructurales pueden ser transcritos para producir un solo ARNm que codifica todas las proteínas estructurales. Debido a la regulación coordinada de genes incluidos en un operón, a la alteración o modificación del promotor individual y/o del elemento regulador puede dar como resultado una alteración o modificación de cada producto génico codificado por el operón. La alteración o modificación del elemento regulador puede incluir, pero no se limita a la remoción del promotor endógeno y/o del(de los) elemento(s) regulador(es), añadiendo promotores fuertes, promotores inducibles o promotores múltiples o removiendo secuencias reguladoras de tal manera que se modifica la expresión de los productos génicos, modificando la ubicación cromosómica del operón, alterando las secuencias de ácido nucleico adyacentes al operón o dentro del operón tal como un sitio de enlazamiento del ribosoma, incrementando el número de copias del operón, modificando proteínas (por ejemplo, proteínas reguladoras, supresores, reforzadores, activadores transcripcionales y similares) involucradas en la transcripción del operón y/o la traducción de los productos génicos del operón, o

cualquier otro medio convencional de desregulación de la expresión de genes rutinario en el arte (incluyendo pero sin limitarse al uso de moléculas antisentido de ácido nucleico, por ejemplo, para bloquear la expresión de proteínas represoras). La desregulación también puede involucrar la alteración de la región de codificación de uno o más genes para producir, por ejemplo, una enzima que sea resistente a la retroalimentación o que tenga una actividad específica mayor o menor.

Se ha modificado por ingeniería genética una célula "recombinante" particularmente preferida de la presente invención para sobreexpresar un gen derivado de una planta o un producto génico o un gen derivado de un microorganismo o producto génico. El término "derivado de una planta", "derivado de un microorganismo", o "derivado de" por ejemplo, incluye un gen que se encuentra naturalmente en un microorganismo o una planta, por ejemplo, una planta oleaginosa, o un producto génico (por ejemplo, Fad5-2) o que es codificado por un gen de una planta o un gen de un microorganismo (por ejemplo, codificado por la SEQ ID NO: 5).

Las metodologías de la presente invención se relacionan con células recombinantes que sobreexpresan al menos una desaturasa de ácido graso.

En otra modalidad, se ha modificado por ingeniería genética una célula recombinante de la presente invención para sobreexpresar una $\Delta 5$ desaturasa de *Pythium irregulare* (el producto génico Fad5-2) (por ejemplo, una desaturasa de ácido graso que tiene la secuencia de aminoácidos de la SEQ ID NO: 6 o codificada por una molécula de ácido nucleico que tiene la secuencia de nucleótidos de la SEQ ID NO: 5).

En otra modalidad, la invención se caracteriza por una célula (por ejemplo, una célula microbiana) que ha sido transformada con un vector que comprende una secuencia de ácido nucleico para desaturasa de ácido graso como la expuesta en la SEQ ID NO: 5).

Otro aspecto de la presente invención se relaciona con un método para modulación de la producción de ácidos grasos que comprende el cultivo de células transformadas por las moléculas de ácido nucleico de la presente invención (por ejemplo, una desaturasa) de tal manera que se presenta la modulación de la producción de ácido graso (por ejemplo, se mejora la producción de ácidos grasos insaturados). El método para cultivar células transformadas por las moléculas de ácido nucleico de la presente invención (por ejemplo, Fad5-2) para modular la producción de ácidos grasos es denominada aquí como "biotransformación." Los procesos de biotransformación pueden utilizar células recombinantes y/o desaturasas descritas aquí. El término "proceso de biotransformación", también denominado aquí como "procesos de bioconversión" incluyen procesos biológicos que resultan en la producción (por ejemplo, transformación o conversión) de cualquier compuesto (por ejemplo, sustrato, compuesto intermedio o producto) que está corriente arriba de una desaturasa de ácido graso en un compuesto (por ejemplo, sustrato, compuesto intermedio, o producto) que está corriente debajo de una desaturasa de ácido graso, en particular, un ácido graso insaturado. En una modalidad, la invención se caracteriza por un proceso de biotransformación para la producción de un ácido graso insaturado que comprende poner en contacto una célula que sobreexpresa al menos una desaturasa de ácido graso con al menos un sustrato apropiado bajo condiciones de tal manera que se produce un ácido graso insaturado y, opcionalmente, recuperar el ácido graso. En una modalidad preferida, la invención se caracteriza por un proceso de biotransformación para la producción de ácidos grasos insaturados que comprende poner en contacto una célula que sobreexpresa Fad5-2 con un sustrato apropiado (por ejemplo, un ácido graso intermedio) bajo condiciones de tal manera que se produzca un ácido graso insaturado (por ejemplo, DHA, SDA, o GLA) y, opcionalmente, recuperar el ácido graso insaturado. Las condiciones bajo las cuales se produce un ácido graso insaturado pueden incluir cualquiera de las condiciones que resultan en la producción deseada de un ácido graso insaturado.

La(s) célula(s) y/o enzimas utilizada(s) en la(s) reacciones de biotransformación están en una forma que le(s) permite llevar a cabo su pretendida función (por ejemplo, la producción de ácidos grasos deseados). Las células pueden ser células enteras, o pueden ser únicamente aquellas porciones de las células necesarias para obtener el resultado final deseado. Las células pueden ser suspendidas (por ejemplo, en una solución apropiada tal como soluciones o medios amortiguados), lavadas (por ejemplo, lavadas hasta quedar libres del medio de cultivo celular), secadas con acetona, inmovilizadas (por ejemplo, con gel de poliácridamida o carragenina kappa o sobre soportes sintéticos, por ejemplo, perlas, matrices y similares), fijadas; entrecruzadas o permeabilizadas (por ejemplo, que tengan membranas y/o paredes permeabilizadas de tal manera que los compuestos, por ejemplo, sustratos, compuestos intermedios o productos puedan pasar más fácilmente a través de dicha membrana o pared). El tipo de célula puede ser cualquier célula que pueda ser utilizada en los métodos de la invención, por ejemplo, células vegetales, animales, o microbianas.

Un aspecto importante de la presente invención involucra el crecimiento de la planta recombinante o el cultivo de los microorganismos recombinantes descritos aquí, de tal manera que se produzca un compuesto deseado (por ejemplo, un ácido graso insaturado deseado). El término "cultivo" incluye el mantenimiento y/o el desarrollo de un microorganismo vivo de la presente invención (por ejemplo, el mantenimiento y/o el desarrollo de un cultivo o cepa). En una modalidad, se cultiva un microorganismo en medio líquido. En otra modalidad, se cultiva un microorganismo en un medio sólido o en un medio semisólido. En una modalidad preferida, se cultiva un microorganismo en un medio (por ejemplo, un medio líquido estéril) que contiene nutrientes esenciales o benéficos para el mantenimiento y/o el desarrollo del microorganismo (por ejemplo, fuentes de carbono o un sustrato de carbono, por ejemplo

carbohidratos complejos tales como alubias o harina de granos, almidones, azúcares, alcoholes de azúcar, hidrocarburos, aceites, grasas, ácidos grasos, ácidos orgánicos y alcoholes; fuentes de nitrógeno, por ejemplo, proteínas vegetales, peptonas, péptidos y aminoácidos derivados de granos, alubias y tubérculos, proteínas, péptidos y aminoácidos derivados de fuentes animales tales como carne, leche y subproductos animales tales como peptonas, extractos de carne e hidrolizados de caseína; fuentes de nitrógeno inorgánico tales como urea, sulfato de amonio, cloruro de amonio, nitrato de amonio y fosfato de amonio; fuentes de fósforo, por ejemplo, ácido fosfórico, sales de sodio y potasio de los mismos; elementos traza, por ejemplo, magnesio, hierro, manganeso, calcio, cobre, zinc; sales de boro, molibdeno, y/o cobalto; así como factores de crecimiento tales como aminoácidos; vitaminas, promotores de crecimiento y similares).

10 Preferiblemente, se cultivan los microorganismos bajo un pH controlado. El término "pH controlado" incluye cualquier pH que resulta en la producción del producto deseado (por ejemplo, un ácido graso insaturado). En una modalidad, se cultivan los microorganismos a un pH de aproximadamente 7. En otra modalidad, se cultivan los microorganismos a un pH entre 6,0 y 8,5. El pH deseado se puede mantener por medio de una cantidad de métodos conocidos por aquellos capacitados en el arte.

15 Preferiblemente también, se cultivan los microorganismos bajo una aireación controlada. El término "aireación controlada" incluye una aireación suficiente (por ejemplo, con oxígeno) que resulta en la producción del producto deseado (por ejemplo, un ácido graso insaturado). En una modalidad, se controla la aireación por medio de la regulación de los niveles de oxígeno en el cultivo, por ejemplo, regulando la cantidad de oxígeno disuelto en el medio de cultivo. Preferiblemente, se controla la aireación del cultivo por medio de la agitación del mismo. La agitación puede ser proporcionada por una hélice o por un equipo similar de agitación mecánica, revolviendo o agitando el recipiente de cultivo (por ejemplo, un fermentador) o por medio de diferentes equipos de bombeo. Adicionalmente se puede controlar la aireación por medio del paso de aire estéril o de oxígeno a través del medio (por ejemplo, a través de la mezcla de fermentación). Preferiblemente también, se cultivan los microorganismos de la presente invención sin excesiva formación de espuma (por ejemplo, a través de la adición de agentes antiespumantes).

Además, se pueden cultivar plantas o microorganismos bajo temperaturas controladas. El término "temperatura controlada" incluye cualquier temperatura que resulte en la producción del producto deseado (por ejemplo, un ácido graso insaturado). En una modalidad, las temperaturas controladas incluyen entre 15°C y 95°C. En otra modalidad, las temperaturas controladas incluyen temperaturas entre 15°C y 70°C. Las temperaturas preferidas están entre 20°C y 55°C, más preferiblemente entre 30°C y 45°C o entre 30°C y 50°C.

Los microorganismos pueden ser cultivados (por ejemplo, mantenidos y/o cultivados) en medio líquido y preferiblemente se cultivan, ya sea en forma continua o intermitente, por medio de métodos de cultivo convencionales tales como un cultivo en reposo, cultivo en tubo de ensayo, cultivo con agitación (por ejemplo, un cultivo con agitación rotatoria, cultivo en un recipiente con agitación, etc.), cultivo agitado aireado, o fermentación. En una modalidad preferida, se cultivan los microorganismos en matraces con agitación. En una modalidad más preferida, se cultivan los microorganismos en un fermentador (por ejemplo, en un proceso de fermentación). El proceso de fermentación de la presente invención incluye, pero no se limita a, métodos de fermentación por tandas, alimentados por tandas y continuos. La frase "proceso por tandas" o "fermentación por tandas" se refiere a un sistema cerrado en el cual la composición del medio, los nutrientes, aditivos suplementarios y similares se fijan al comienzo de la fermentación y no está sometido a alteración durante la fermentación, sin embargo, deben hacerse intentos para controlar factores tales como el pH y la concentración de oxígeno para evitar una acidulación excesiva del medio y/o la muerte de los microorganismos. La frase "proceso por tandas alimentado" o fermentación "por tandas alimentada" se refiere a una fermentación por tandas con la excepción de que se añaden uno o más sustratos o suplementos (por ejemplo, añadidos en incrementos o en forma continua) a medida que avanza la fermentación. La frase "proceso continuo" o "fermentación continua" se refiere a un sistema en el cual se añade un medio de fermentación definido en forma continua a un fermentador y se remueve simultáneamente una cantidad igual de medio usado o "acondicionado", preferiblemente para la recuperación del producto deseado (por ejemplo, un ácido graso insaturado). Se han desarrollado una variedad de tales procesos y son bien conocidos en el arte.

La frase "cultivo bajo condiciones de tal manera que se produzca un compuesto deseado (por ejemplo, un ácido graso insaturado, por ejemplo, DHA)" incluye el mantenimiento y/o el cultivo plantas o de microorganismos bajo condiciones (por ejemplo, temperatura, presión, pH, duración, etc.) apropiadas o suficientes para obtener la producción del compuesto deseado o para obtener rendimientos deseados del compuesto particular que está siendo producido. Por ejemplo, se continúa el cultivo durante un período de tiempo suficiente para producir la cantidad deseada de un ácido graso insaturado (por ejemplo, DHA). Preferiblemente, se continúa el cultivo durante un período de tiempo suficiente para alcanzar sustancialmente una producción máxima del ácido graso insaturado. En una modalidad, se continúa el cultivo aproximadamente durante 12 a 24 horas. En otra modalidad, se continúa el cultivo aproximadamente durante 24 a 36 horas, 36 a 48 horas, 48 a 72 horas, 72 a 96 horas, 96 a 120 horas, 120 a 144 horas, o más de 144 horas. En otra modalidad, se continúa el cultivo durante un período de tiempo suficiente para alcanzar rendimientos en la producción de ácidos grasos insaturados, por ejemplo, se cultivan células de tal manera que se produzcan al menos aproximadamente 15 a 20 g/L de ácidos grasos insaturados, se produzcan al menos aproximadamente 20 a 25 g/L de ácidos grasos insaturados, se produzcan al menos aproximadamente 25 a

30 g/L de ácidos grasos insaturados, se produzcan al menos aproximadamente 30 a 35 g/L de ácidos grasos insaturados, se produzcan al menos aproximadamente 35 a 40 g/L de ácidos grasos insaturados (por ejemplo, aproximadamente al menos 37 g/L de ácidos grasos insaturados) o se produzcan al menos aproximadamente 40 a 50 g/L de ácidos grasos insaturados. En aún otra modalidad, se cultivan microorganismos bajo condiciones de tal manera que se produzca un rendimiento preferido de ácidos grasos insaturados, por ejemplo, un rendimiento dentro de un rango expuesto más arriba, en aproximadamente 24 horas, en aproximadamente 36 horas, en aproximadamente 48 horas, en aproximadamente 72 horas, o en aproximadamente 96 horas.

En la producción de ácidos grasos insaturados, adicionalmente puede ser deseable cultivar células de la presente invención en presencia de sustratos biosintéticos de ácido graso como suplemento. El término "sustrato biosintético de ácido graso como suplemento" incluye un agente o compuesto que, cuando se lo pone en contacto con una célula o se lo incluye en el medio de cultivo de una célula, sirve para mejorar o incrementar la biosíntesis de ácido graso insaturado. Se pueden añadir sustratos biosintéticos de ácido graso como suplemento en la forma de una solución concentrada o suspensión (por ejemplo, en un solvente adecuado tal como agua o un amortiguador) o en la forma de un sólido (por ejemplo, en la forma de un polvo). Además, se pueden añadir sustratos biosintéticos de ácido graso como suplemento de la presente invención como una sola alícuota, en forma continua o intermitente durante un periodo dado de tiempo.

La metodología de la presente invención puede incluir adicionalmente una etapa de recuperación de un compuesto deseado (por ejemplo, un ácido graso insaturado). El término "recuperación" de un compuesto deseado incluye extracción, cosecha, aislamiento o purificación del compuesto a partir del medio de cultivo. La recuperación del compuesto se puede llevar a cabo de acuerdo con cualquier metodología convencional de aislamiento o purificación conocida en el arte incluyendo, pero sin limitarse al tratamiento con una resina convencional (por ejemplo, una resina de intercambio aniónico o catiónico, una resina de adsorción no iónica, etc.), el tratamiento con un adsorbente convencional (por ejemplo, carbón activado, ácido silícico, gel de sílice, celulosa, alúmina, etc.), alteración del pH, extracción con solvente (por ejemplo, con un solvente convencional tal como un alcohol, acetato de etilo, hexano y similares), diálisis, filtración, concentración, cristalización, recristalización, ajuste del pH, liofilización y similares. Por ejemplo, se puede recuperar un compuesto del medio de cultivo removiendo primero los microorganismos del cultivo. Se pasa luego el medio a través o sobre una resina de intercambio catiónico para remover cationes no deseados y luego a través o sobre una resina de intercambio aniónico para remover aniones inorgánicos no deseados y ácidos orgánicos que tienen una acidez mayor que la del ácido graso insaturado de interés (por ejemplo, DHA).

Preferiblemente, se "extrae", "aisla" o "purifica" un compuesto deseado de tal manera que la preparación resultante esté sustancialmente libre de otros componentes (por ejemplo, libre de componentes del medio y/o de subproductos de la fermentación). La expresión "sustancialmente libre de otros componentes" incluye preparaciones del compuesto deseado en las cuales se separa el compuesto (por ejemplo, purificado o parcialmente purificado) de los componentes del medio o de los subproductos de la fermentación del cultivo a partir del cual es producido. En una modalidad, la preparación tiene aproximadamente más del 80% (en peso seco) del compuesto deseado (por ejemplo, aproximadamente menos del 20% de los otros componentes del medio o subproductos de fermentación), más preferiblemente aproximadamente más del 90% del compuesto deseado (por ejemplo, aproximadamente menos del 10% de los otros componentes del medio o subproductos de la fermentación), aún más preferiblemente aproximadamente más del 95% del compuesto deseado (por ejemplo, aproximadamente menos del 5% de los otros componentes del medio o subproductos de fermentación), y lo más preferible aproximadamente más del 98 - 99% del compuesto deseado (por ejemplo, aproximadamente menos del 1 - 2% de los otros componentes del medio o subproductos de la fermentación). Cuando el compuesto deseado es un ácido graso insaturado que ha sido convertido en una sal, el compuesto está preferiblemente libre adicionalmente (por ejemplo, sustancialmente libre) de contaminantes químicos asociados con la formación de la sal. Cuando el compuesto deseado es un ácido graso insaturado que ha sido convertido en un alcohol, el compuesto está preferiblemente libre adicionalmente (por ejemplo, sustancialmente libre) de contaminantes químicos asociados con la formación del alcohol.

En una modalidad alternativa, el ácido graso insaturado deseado no es purificado de la planta o el microorganismo, por ejemplo, cuando la planta o el microorganismo no es biológicamente peligroso (por ejemplo, seguro). Por ejemplo, se puede usar la planta completa o el cultivo (o el sobrenadante del cultivo) como fuente del producto (por ejemplo, el producto crudo). En una modalidad, se usa la planta o el sobrenadante del cultivo (o el sobrenadante del cultivo) sin modificación. En otra modalidad, se concentra la planta o el cultivo (o el sobrenadante del cultivo). En aún otra modalidad, se pulveriza, seca, o liofiliza la planta o el cultivo (o el sobrenadante del cultivo).

B. Metodologías de producción de alto rendimiento

Una modalidad particularmente preferida de la presente invención es un método de producción de alto rendimiento para la producción de ácidos grasos insaturados, por ejemplo, DHA, que comprende el cultivo de una planta o de un microorganismo manipulados bajo condiciones tales que se produce el ácido graso insaturado con un rendimiento significativamente alto. La frase "método de producción de alto rendimiento", por ejemplo, un método de producción de alto rendimiento para la producción de un compuesto deseado (por ejemplo, para la producción de un ácido graso insaturado) incluye un método que resulta en la producción del compuesto deseado a un nivel que es elevado o está

por encima de lo usual para métodos de producción comparables. Preferiblemente, un método de producción de alto rendimiento resulta en la producción del compuesto deseado con un rendimiento significativamente alto. La frase "rendimiento significativamente alto" incluye un nivel de producción o rendimiento que es suficientemente elevado o está por encima de lo usual para métodos de producción comparables, por ejemplo, que está elevado hasta un nivel suficiente para una producción comercial del producto deseado (por ejemplo, la producción del producto a un costo comercialmente viable). En una modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido grasos insaturado con un nivel superior a 2 g/L. En otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 10 g/L. En otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 20 g/L. En aún otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 30 g/L. En aún otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 40 g/L.

La invención se caracteriza además por un método de producción de alto rendimiento para la producción de un compuesto deseado (por ejemplo, para la producción de un ácido graso insaturado) que involucra el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un nivel suficientemente elevado de compuesto dentro de un período de tiempo comercialmente deseado. En una modalidad de ejemplo, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 15 - 20 g/L en 36 horas. En otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 25 - 30 g/L en 48 horas. En otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 35 - 40 g/L en 72 horas, por ejemplo, superior a 37g/L en 72 horas. En otra modalidad, la invención se caracteriza por un método de producción de alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce un ácido graso insaturado con un nivel superior a 30 - 40 g/L en 60 horas, por ejemplo, superior a 30, 35 ó 40 g/L en 60 horas. Los valores y los rangos incluidos y/o intermedios dentro de los rangos expuestos aquí también se pretende que estén incluidos dentro del alcance de la presente invención. Por ejemplo, se pretende que la producción de ácido graso insaturado en niveles de al menos 31, 32, 33, 34, 35, 36, 37, 38 y 39 g/L en 60 horas esté incluida dentro del rango de 30 - 40 g/L en 60 horas. En otro ejemplo, se pretende que rangos de 30 - 35 g/L o 35 - 40 g/L estén incluidos dentro del rango de 30 - 40 g/L en 60 horas. Además, la persona experta se dará cuenta que el cultivo de un microorganismo manipulado para lograr un nivel de producción, por ejemplo, de "30 - 40 g/L en 60 horas" incluye el cultivo del microorganismo durante períodos de tiempo adicionales (por ejemplo, períodos de tiempo superiores a 60 horas), resultando opcionalmente incluso en rendimientos más altos del ácido graso insaturado que está siendo producido.

IV. Composiciones

Las moléculas de ácido nucleico para desaturasa y las proteínas de la invención pueden ser utilizadas para producir ácidos grasos insaturados que pueden ser incorporados en composiciones. Las composiciones de la presente invención incluyen, por ejemplo, composiciones para uso como alimento para animales, composiciones para uso como neutracéuticos (por ejemplo, suplementos dietéticos), y composiciones farmacéuticas adecuadas para administración.

Tales composiciones farmacéuticas típicamente contienen un ácido graso insaturado y un portador farmacéuticamente aceptable. Como se la utiliza aquí la expresión "portador farmacéuticamente aceptable" pretende incluir todos y cada uno de los solventes, medios de dispersión, recubrimientos, agentes antibacterianos y contra hongos, agentes isotónicos y que retrasan la absorción, y similares, compatibles con administración farmacéutica. El uso de tales medios y agentes para sustancias farmacéuticamente activas es bien conocido en el arte. Excepto en la medida en que cualquier medio convencional o agente sea incompatible con el compuesto activo, se contempla el uso de los mismos en las composiciones. También se pueden incorporar compuestos activos suplementarios en las composiciones.

Una composición farmacéutica de la invención se formula para que sea compatible con su pretendida ruta de administración. Los ejemplos de rutas de administración incluyen administración parenteral, por ejemplo, intravenosa, intradérmica, subcutánea; administración oral (por ejemplo, inhalación), transdérmica (tópica),

transmucosa, y rectal. Las soluciones o suspensiones utilizadas para aplicación parenteral, intradérmica, o subcutánea pueden incluir los siguientes componentes: un diluyente estearil tal como agua para inyección, solución salina, aceites fijos, polietilén glicoles, glicerina, propilén glicol u otros solventes sintéticos; agentes antibacteriales tales como alcohol bencílico o metil parabenos; antioxidantes tales como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como el ácido etilendiaminotetraacético; amortiguadores tales como acetatos, citratos o fosfatos y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa. Se puede ajustar el pH con ácidos o bases, tales como ácido clorhídrico o hidróxido de sodio. La preparación parenteral puede estar contenida en ampolletas, jeringas desechables o viales de dosis múltiples de vidrio o de plástico.

Las composiciones farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas estériles (solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones inyectables estériles o dispersiones. Para administración intravenosa, los portadores adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor EL™ (BASF, Parsippany, NJ) o solución salina amortiguada con fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida hasta un grado en que salga fácilmente por la jeringa. Debe ser estable bajo las condiciones de fabricación y almacenamiento y debe ser preservada contra la acción contaminante de microorganismos tales como bacterias y hongos. El portador puede ser un medio solvente o dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilén glicol, y polietilén glicol, y similares), y mezclas adecuadas de los mismos. La fluidez adecuada puede ser mantenida, por ejemplo, por medio del uso de un recubrimiento tal como lecitina, por medio del mantenimiento del tamaño de partícula requerido en el caso de una dispersión y por medio del uso de tensoactivos. La prevención de la acción de los microorganismos puede ser lograda por medio de diferentes agentes antibacteriales y contra hongos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal, y similares. En muchos casos, sería preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol, cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables puede ser lograda por medio de la inclusión en la composición de un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las soluciones inyectables estériles pueden ser preparadas por medio de la incorporación del compuesto activo (por ejemplo, un LCPUFA, o un fragmento del mismo, producido por medio de moléculas de ácido nucleico y de proteína de la presente invención) en la cantidad requerida en un solvente apropiado con uno o con una combinación de ingredientes enumerada anteriormente, según se requiera, seguido por esterilización por filtración. Generalmente, se preparan las dispersiones por medio de la incorporación del compuesto activo en un vehículo estéril que contiene un medio básico de dispersión y los otros ingredientes requeridos a partir de aquellos enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables, los métodos preferidos de preparación son secado al vacío y liofilización que producen un polvo del ingrediente activo más cualquier ingrediente adicional deseado de una solución estéril previamente filtrada del mismo.

Las composiciones orales generalmente incluyen un diluyente inerte o un portador comestible. Ellas pueden estar incluidas en cápsula de gelatina o comprimidas en tabletas. Para el propósito de administración terapéutica oral, se puede incorporar el compuesto activo con excipientes y utilizarlo en forma de tabletas, pastillas, o cápsulas. También se pueden preparar composiciones orales utilizando un portador líquido para ser usadas como enjuague bucal, en donde el compuesto en el portador líquido se aplica en forma oral y se lo mantiene un momento en la boca y se lo expectora o se traga. Se pueden incluir agentes aglutinantes farmacéuticamente compatibles, y/o materiales adyuvantes como parte de la composición. Las tabletas, píldoras, cápsulas, pastillas y similares pueden contener cualquiera de los siguientes ingredientes, o compuestos de naturaleza similar: un aglutinante tal como celulosa microcristalina, goma tragacanto o gelatina; un excipiente tal como almidón o lactosa, un agente desintegrante tal como ácido alginico, Primogel, o almidón de maíz; un lubricante tal como estearato de magnesio o Sterotes; un deslizante tal como dióxido de silicio coloidal; un agente edulcorante tal como sacarosa sacarina; o un agente saborizante tal como menta, salicilato de metilo, o saborizante de naranja.

Para administración por inhalación, se suministran los compuestos en forma de un aerosol a partir de un contenedor o dispensador presurizado que contiene un propelente adecuado, por ejemplo, un gas tal como dióxido de carbono, o un nebulizador.

La administración sistémica puede ser también por medio transdérmico o transmucosa. Para administración transdérmica o transmucosa, se utilizan penetrantes apropiados para la barrera que va a ser permeada en la formulación. Tales penetrantes son generalmente conocidos en el arte, e incluyen, por ejemplo, para administración transmucosa, detergentes, sales biliares, y derivados de ácido fusídico. La administración transmucosa puede ser lograda a través del uso de aerosoles nasales o supositorios. Para administración transdérmica, se formulan los compuestos activos en ungüentos, bálsamos, geles, o cremas como las generalmente conocidas en el arte.

Los compuestos también pueden ser preparados en la forma de supositorios (por ejemplo, con bases convencionales para supositorios tales como manteca de cacao y otros glicéridos) o enemas de retención para suministro rectal.

Los compuestos activos pueden ser preparados con portadores que protegerán al compuesto contra la eliminación rápida del organismo, tales como una formulación de liberación controlada, incluidos implantes y sistemas de suministro microencapsulados. Se pueden utilizar polímeros biocompatibles biodegradables, tales como acetato de

vinilo etilénico, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y ácido poliláctico. Los métodos para la preparación de tales formulaciones serán evidentes para aquellas personas capacitadas en el arte. Los materiales también pueden ser obtenidos comercialmente de Alza Corporation y Nova Pharmaceuticals, Inc. También se pueden utilizar suspensiones liposomales (incluidos liposomas dirigidos a células infectadas con anticuerpos monoclonales para antígenos virales) como portadores farmacéuticamente aceptables. Estas pueden ser preparadas de acuerdo con métodos conocidos por aquellos capacitados en el arte, por ejemplo, como se describe en la Patente Estadounidense No. 4.522.811.

Es especialmente ventajoso formular composiciones orales o parenterales en forma de dosis unitarias para facilidad de administración y uniformidad de la dosis. La forma de dosis unitarias como se la utiliza aquí se refiere a unidades físicamente discretas adecuadas como dosis unitarias para el individuo que va a ser tratado; cada unidad que contiene una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado junto con el portador farmacéutico requerido. Las especificaciones de las formas de dosificación unitaria están dictadas por y dependen directamente de las características únicas del compuesto activo y del efecto terapéutico particular que se pretende lograr, y por las limitaciones inherentes en el arte respecto a la formación de compuestos tales como un compuesto activo para el tratamiento de los individuos.

La toxicidad y la eficacia terapéutica de tales compuestos se pueden determinar por medio de procedimientos farmacéuticos estándar en cultivos celulares o en animales experimentales, por ejemplo, para determinar la LD50 (la dosis letal para el 50% de la población) y la ED50 (la dosis terapéuticamente efectiva en el 50% de la población). La relación de la dosis entre efectos tóxicos y terapéuticos es el índice terapéutico y puede expresarse como la relación LD50/ED50. Los compuestos que exhiben grandes índices terapéuticos son los preferidos. Aunque se pueden utilizar compuestos que exhiben efectos secundarios tóxicos, se debe tener cuidado al diseñar un sistema de suministro que dirija tales compuestos al sitio del tejido afectado con el propósito de reducir el daño potencial a las células no infectadas y, por lo tanto, reducir los efectos secundarios.

Los datos obtenidos a partir de los ensayos con cultivos celulares y estudios con animales pueden ser utilizados en la formulación de un rango de dosis para uso en humanos. La dosis de tales compuestos cae preferiblemente dentro de un rango de concentraciones circulantes que incluyen la ED50 con poca o ninguna toxicidad. La dosis puede variar dentro de este rango dependiendo de la forma de la dosis empleada y de la ruta de administración utilizada. Para cualquier compuesto utilizado en el método de la invención, la dosis terapéuticamente efectiva puede ser estimada inicialmente a partir de ensayos con cultivos celulares. Se puede formular una dosis en modelos animales para lograr un rango de concentración en el plasma circulante que incluya al IC50 (es decir, la concentración de los compuestos de prueba que logran una inhibición media máxima de los síntomas) como se determina en el cultivo celular. Tal información puede ser utilizada para determinar en forma más precisa las dosis útiles en humanos. Los niveles en plasma se pueden medir, por ejemplo, por medio de cromatografía líquida de alto rendimiento.

Como se define aquí, una cantidad terapéuticamente efectiva de proteína o de polipéptido (es decir, una dosis efectiva) está en el rango aproximadamente de 0,001 a 30 mg/kg de peso corporal, preferiblemente aproximadamente de 0,01 a 25 mg/kg de peso corporal, más preferiblemente aproximadamente de 0,1 a 20 mg/kg de peso corporal, e incluso más preferiblemente aproximadamente de 1 a 10 mg/kg, de 2 a 9 mg/kg, de 3 a 8 mg/kg, de 4 a 7 mg/kg, o de 5 a 6 mg/kg de peso corporal. Es la persona capacitada se darán cuenta que ciertos factores pueden influir sobre la dosis requerida para tratar efectivamente a un individuo, incluyendo pero sin limitarse a la severidad de la enfermedad o el trastorno, tratamientos previos, el estado general de salud y/o la edad del individuo, y otras enfermedades presentes. Además, el tratamiento de un individuo con una cantidad terapéuticamente efectiva de una proteína, polipéptido, o anticuerpo puede incluir un tratamiento único o, preferiblemente, puede incluir una serie de tratamientos.

En un ejemplo preferido, se trata un individuo con un LCPUFA en el rango aproximadamente entre 0,1 a 20 mg/kg de peso corporal, una vez por semana aproximadamente entre 1 a 10 semanas, preferiblemente entre 2 a 8 semanas, más preferiblemente aproximadamente entre 3 a 7 semanas, e incluso más preferiblemente aproximadamente durante 4, 5, ó 6 semanas. Se apreciará también que la dosis efectiva de anticuerpo, proteína, o polipéptido utilizada para el tratamiento puede ser incrementada o disminuida en el transcurso de un tratamiento particular. Pueden haber cambios en las dosis y hacerse evidentes a partir de los resultados de los ensayos de diagnóstico como se describe aquí.

Las composiciones farmacéuticas pueden ser colocadas en un contenedor, empaque, o dispensador junto con las instrucciones para su administración.

Se ilustra adicionalmente la invención por medio de los siguientes ejemplos que no deben ser considerados como limitantes.

Ejemplos

Material: Se adquirió el *Pythium irregulare* a partir de la American Type Culture Collection (12301 Parklawn Drive, Rockville, Maryland, 20852 EUA) y se lo cultivó en un medio (Weete, J. D., et al. (1997) Lipids 32: 839 - 845) a 24°C

durante 7 días. Después de eso se recolectó la biomasa por centrifugación y se la utilizó para el aislamiento del ARN.

Ejemplo 1: Construcción y selección de la biblioteca de ADNc

- 5 Se aisló el ARN total a partir de los materiales anteriores de acuerdo con Qiu y Erickson (Qiu, X. y Erickson, L. (1994) Plant Mol. Biol. Repr. 12: 209 - 214). Se construyó la biblioteca de ADNc a partir del ARN total. Se sintetizó la primera cadena del ADNc por medio de la transcriptasa inversa SuperScript II de Gibco-BRL. Se sintetizó la segunda cadena de ADNc por medio de una ADN polimerasa I de Stratagene. Después del fraccionamiento por tamaño, se ligaron dos insertos de ADNc mayores a 1 kb dentro del vector λ Uni-Zap XR (Stratagene). Se empacaron luego los ADN ~ recombinantes con extracto de empaque Gigapack III Gold (Stratagene) y se sembró sobre placas NZY. La biblioteca resultante representó más de 5×10^6 clones independientes. La selección de la biblioteca de ADNc se realizó de acuerdo con métodos estándar (Sambrook, J, Fritseh, E. F Maniatis, T. (1989) Molecular cloning - A laboratory manual. (Cold Spring Harbor, New York, EUA).

Ejemplo 2: RT-PCR

- 15 Se sintetizó una sola cadena de ADNc por medio de la transcriptasa inversa SuperScript II (Gibco-BRL) a partir del ARN total y se la utilizó luego como molde para la reacción PCR con dos iniciadores degenerados (El iniciador hacia adelante:

- 20 GCXCA/GAXGAXCAC/TCCXGGXGG y el iniciador inverso: ATXTG/TXGGA/GAAXAG/AG/ATGG/ATG). La amplificación por PCR consistió de 35 ciclos con 1 min a 94° C, 1,5 min a 55 °C y 2 min a 72 °C seguido por una etapa de extensión a 72 °C durante 10 min. Se aislaron los productos amplificados desde 800 pb hasta 1000 pb a partir del gel de agarosa y se los purificó por medio de un kit (purificación en el gel Qiaex II, Qiagen), y

posteriormente se los clonó en el vector de coloración TA pCR®2.1 (Invitrogen). Los insertos clonados fueron luego secuenciados por medio de un Sistema Terminador del Ciclo de Secuenciación PRISM DyeDeoxy (Perkin Elmer/Applied Biosystems).

Ejemplo 3: Expresión de FAD5-2 en levadura

- 25 Se amplificó el marco de lectura abierto de Fad5-2 por medio de PCR utilizando la enzima Precision Plus (Stratagene) y clonó en un vector de clonación TA (pGR®2.1, Invitrogen). Habiendo confirmado que los productos de la PCR eran idénticos a los ADNc originales por medio de secuenciación, se liberaron entonces los fragmentos por medio de una digestión doble con BamHI-EcoRI y se los insertó dentro del vector de expresión de levadura pYES2 (Invitrogen) bajo el control del promotor inducible GAL1.
- 30 Se transformaron las cepas de levadura InvSc2 (Invitrogen) con las construcciones de expresión utilizando el método de acetato de litio y se seleccionaron dos transformantes sobre placas con medio mínimo que carecían de uracilo (Gietz, D., et al. (1992) Nucleic Acids Res. 20: 1425; Covello, P. S. y Reed, D. W. (1996) Plant Physiol. 111: 223 - 226).
- 35 Se cultivaron primero los transformantes en medio mínimo que carecía de uracilo y que contenía glucosa a 28°C. Después de cultivar durante la noche, se centrifugan las células, se las lavó y resuspendió en agua destilada. Se inoculó medio mínimo que contenía 2% de galactosa, con o sin sustrato de ácidos grasos 0,3 mM en presencia de 0,1 % de tergitol, con la suspensión de células transformantes de levadura y se incubó a 20°C durante tres días, y luego a 15 °C durante otros tres días.

Ejemplo 4: análisis del ácido graso

- 40 Se cosecharon células de *Pythium irregulare* y de levadura y se lavó dos veces con agua destilada. Luego se añadieron 2 mL de KOH metanólico (7,5% p/v de KOH en metanol al 95%) a los materiales y se selló la mezcla en un tubo de vidrio para cultivo de 12 ml que fue calentado a 80°C durante 2 horas. Se añadieron 0,5 mL de agua y se extrajo la muestra dos veces con 2 mL de hexano para remover los lípidos no saponificables. Se aciduló luego el resto de la fase acuosa por medio de la adición de 1 mL de HCl 6 N y se extrajo dos veces con 2 mL de hexano. Se combinaron las fases de hexano y se secó bajo una corriente de nitrógeno. Se añadieron 2 mL de HCl metanólico 3 N (SUPELCO, Supelco Park, Bellefonte, PA 16823-0048) y se calentó la mezcla a 80°C durante 2 horas. Después de enfriar a temperatura ambiente, se añadió 1 mL de NaCl al 0,9% y se extrajo la mezcla dos veces con 2 x 2 mL de hexano. Se evaporó el hexano combinado bajo atmósfera de nitrógeno. Se analizaron los ésteres metílicos de ácido graso resultantes (los FAME) por medio de GC y GC-MS de acuerdo con Covello & Reed (Covello, P.S. y Reed, D.W. (1996) Plant Physiol. 111: 223 - 226).

- 50 Se llevó a cabo el análisis por GC/MS en modo EI estándar utilizando un espectrómetro de masas Fisons VG TRIO 2000 (VG Analytical, RU) controlado por medio de un software Masslynx versión 2.0, acoplado a un cromatógrafo de gases GC Serie 8000. Se utilizó una columna DB-23 (30M x 0 - 25 mm de diámetro interno, espesor de película 0,25 mm, J&W Scientific, Folsom, CA) que fue programada a una temperatura de 180 °C durante 1 min, luego a 4°C /min a 240 °C y mantenida durante 15 minutos, para el análisis FAME.

Ejemplo 5: Transformación de Brassica juncea y de linaza (*Linum usitatissimum*) y tratamiento con ácido graso exógeno

Se utilizaron los hipocotiledones de plántulas de 5 - 6 días de *B. juncea* y de linaza como explantes para inoculación con el *Agrobacterium tumefaciens* que hospeda vectores binarios con los ADNc de longitud completa bajo el control de los diferentes promotores. Se utilizaron plántulas transgénicas de 20 días para tratamiento con ácido graso exógeno. Se dividieron las plántulas en tres partes: hojas, tallos y raíces. Cada una fue cortada en piezas pequeñas y colocadas en una placa de titulación de 24 pozos. A cada pozo se le añadieron 2 mL de sal de sodio al 0,05% de los sustratos (NuCheck Prep Inc., Elysian, MN). Se incubó luego la placa a 24°C durante 4 h con agitación suave. Después de la incubación, se lavaron los tejidos de la planta tres veces con agua y luego se los utilizó para el análisis del ácido graso.

Ejemplo 6: Perfil de ácidos grasos de *Thrauschytrium sp.*

Pythium irregulare ha llamado recientemente la atención científica dirigido a su habilidad para la producción de los LCPUFA por ejemplo DHA, AA, EPA y DPA. La Figura 8 muestra la composición de ácidos grasos de los lípidos aislados de cultivos de 7 días de *Pythium irregulare*. Como se muestra en la tabla, el microorganismo contiene un amplio rango de poliácidos grasos insaturados, tanto de familias n-3 como n-6.

Ejemplo 7: identificación de los ADNc que codifican para el "extremo frontal" de la desaturasa

Para identificar los genes que codifican para las desaturasas involucradas en la biosíntesis de los LCFCTFA en *Pythium irregulare*, se adoptó una estrategia de clonación con base en la PCR. Se diseñan dos iniciadores degenerados para dirigir el motivo de enlazamiento del hemo de la extensión del terminal N del dominio del tipo citocromo _{b5} en el extremo frontal de las desaturasas y el tercer motivo conservado de histidina en todas las desaturasas microsomales, respectivamente. La razón fundamental detrás del diseño es que las desaturasas involucradas en la biosíntesis de EPA y DHA en *Pythium irregulare*, deben tener una estructura primaria similar al de los otros extremos frontales de las desaturasas, es decir la extensión del terminal N del dominio del tipo del citocromo _{b5} en la desaturasa. Se identificaron cuatro fragmentos de los ADNc de *Pythium irregulare* que codifican las proteínas de fusión que contienen al dominio del tipo del citocromo _{b5} en el terminal N.

Para aislar clones de ADNc de longitud completa, se utilizaron los cuatro insertos como sondas para seleccionar las bibliotecas de ADNc de *Pythium irregulare*, lo cual dio como resultado la identificación de varios clones de ADNc en cada grupo. La secuenciación de todos esos clones identificó los cuatro ADNc de longitud completa que fueron llamados como Fad4, Fad5, Fad5-2 y Fad6. El marco de lectura abierto de Fad5-2 de *Pythium irregulare* es de 1371 pb y codifica para 456 aminoácidos (Figura 1). Fad6 de *Pythium irregulare* es de 1383 pb de longitud y codifica para 460 aminoácidos (Figura 2). La comparación secuencial de las dos secuencias de *Pythium irregulare* mostró una similitud superior al 39% entre las proteínas deducidas (Figura 3).

Una búsqueda por BLASTP™ de la base de datos de proteínas reveló los siguientes aciertos para cada una de las cuatro proteínas, Fad4, Fad5, Fad5-2, y Fad6:

Fad 5-2 (456 residuos aminoácidos)

Blastp nr

Acceso No.	Organismo	Descripción	Longitud	% de identidad
AB029311	<i>Dictostelium discoideum</i>	Desaturasa de ácido graso	443	41
AB022097	<i>Dictostelium discoideum</i>	Δ5 desaturasa de ácido graso	445	39
AF067654	<i>Mortierella alpina</i>	Δ5 desaturasa de ácido graso	441	38
AF054824	<i>Mortierella alpina</i>	Δ5 desaturasa microsomal	441	38
L11421 D90914	<i>Synechocystis sp.</i>	Δ6 desaturasa	361	28

Fad 6 (459 residuos aminoácidos)

Blastp nr

Acceso No.	Organismo	Descripción	Longitud	% de identidad
AF110510	<i>Mortierella alpina</i>	$\Delta 6$ desaturasa de ácido graso	437	38
AB020032	<i>Mortierella alpina</i>	$\Delta 6$ desaturasa de ácido graso	437	38
AF306634	<i>Mortierella isabelina</i>	$\Delta 6$ desaturasa de ácido graso	437	38
AF307940	<i>Mortierella alpina</i>	$\Delta 6$ desaturasa de ácido graso	438	38
AJ250735	<i>Ceratodon purpureus</i>	$\Delta 6$ desaturasa de ácido graso	438	36

Ejemplo 8: Expresión de fad5-2 en levadura

Para confirmar la función de Fad5 y de Fad5-2, se transformó el Invsc2 de *S. cerevisiae* con plásmidos, que contienen el marco de lectura abierto de los Fad5 y Fad5-2 respectivamente bajo el control del promotor inducible por galactosa. Cuando fueron inducidos los transformantes de levadura por medio de galactosa en un medio que contiene ácido homo-gamma-linolénico (HGLA, 20:3 - 8; 11,14), se observó un pico extra en el cromatograma de los FAME que se acumulan en los transformantes comparado con el control. Una comparación del cromatograma con aquella de los estándares reveló que el ácido graso tenía un tiempo de retención idéntico al del estándar de ácido araquidónico (AA, 20:4 - 5, 8, 11, 14). Para confirmar adicionalmente la regioquímica de los productos, los FAME fueron analizados por medio de GC/MS. El espectro de masas del nuevo ácido graso y el estándar de AA son idénticos. Estos resultados demuestran que Fad5 y Fad5-2 convierten HGLA (20:3 - 8, 11, 13) en AA (20:4 - 5, 8, 11, 14) en levadura. Para estudiar adicionalmente la especificidad del sustrato de Fad5-2, el plásmido que contiene Fad5-2 fue transferido en otra cepa de levadura AMY-2 α donde se interrumpe ole1, un gen para la $\Delta 9$ desaturasa. La cepa es incapaz de crecer en medio mínimo sin la suplementación con mono-ácidos grasos insaturados. En este experimento, se cultivó la cepa en medio mínimo suplementado con 17:1(10Z), un sustrato que no es de Fad5-2, que permitió el estudio de la especificidad de la enzima hacia diferentes sustratos, especialmente monoinsaturados. Se analizaron una cantidad de posibles sustratos incluyendo 16:1(9Z), 18:1(9Z), 18:1(11Z), 18:1(11E), 18:1(12E), 18:1(15Z), 18:2(9Z, 12Z), 18:3(9Z, 12Z, 15Z), 20:2(11Z, 14Z) y 20:3(11Z, 14Z, 17Z). Los resultados indicaron que Fad5-2 podría desaturar ácidos grasos insaturados con dobles enlaces $\Delta 9$ etilénicos y $\Delta 11$ etilénicos, así como el ácido graso con un doble enlace $\Delta 8$ etilénico (20:3 - 8, 11, 14). Como se muestra en la Figura 4, convirtió efectivamente Fad5-2 tanto en sustratos 18:1(9Z) como en sustratos 18:1(11Z) en sus correspondientes ácidos grasos $\Delta 5$ desaturados, 18:2 - 5, 9 (el tiempo de retención 10,34 min) y 18:1 - 5, 11 (el tiempo de retención 10,44 min), respectivamente. Fad5-2 también desatura ácidos grasos trans tales como 18:1 (11E) y 18:1(12E).

La Figura 9 es una comparación de la preferencia del sustrato Fad5-2 por sustratos de ácido graso analizados en la cepa de levadura AMY-2 α . Las proporciones relativas de los sustratos y los productos acumulados son un indicador útil de la preferencia del sustrato de la enzima. Como se muestra en la Figure 9, Fad5-2 prefiere ácidos grasos con 20 carbonos como sustratos, tales como 20:3(8Z, 11Z, 14Z), 20:3(11Z, 14Z, 17Z) y 20:2(11Z, 14Z). Mientras el ácido graso de cadena más corta es un sustrato relativamente más débil para la enzima en levadura.

Ejemplo 9: Expresión de Fad5-2 en *B. juncea*

Para determinar si Fad5-2 es funcional en cultivos oleaginosos y si su expresión tiene algún efecto sobre su crecimiento y desarrollo, se transformaron *B. juncea* con un vector binario que contenía ADNc para Fad5-2 detrás de un promotor constitutivo (un promotor 35S del virus del mosaico de la coliflor en tándem). Se obtuvieron seis plantas transgénicas primarias independientes y se determinó el perfil de ácidos grasos de los lípidos de tejidos diferentes. La Figura 5 muestra la composición de ácidos grasos de plántulas de tres semanas de edad de una línea T₁. Comparado con el tipo silvestre, todos los tejidos de plantas transgénicas tienen un perfil alterado de ácidos grasos que contiene cuatro picos adicionales que fueron identificados como cuatro diferentes ácidos grasos interrumpidos por polimetileno $\Delta 5$ no desaturado (los $\Delta 5$ -UPIFA), específicamente, ácidos taxoleico (18:2 - 5, 9); efedrélico (18:2 - 5, 11); pinolénico (18:3 - 5, 9, 12), y coniferónico (18:4 - 5, 9, 12, 15). Por lo tanto *B. juncea*, como la levadura, puede expresar funcionalmente la $\Delta 5$ desaturasa de *P. irregulare* para convertir los sustratos endógenos 18:1 - 9; 18:1 - 11; 18:2 - 9, 12, y 18:3 - 9, 12, 15 a los correspondientes ácidos grasos $\Delta 5$ desaturados. Las raíces produjeron la cantidad más alta de los $\Delta 5$ UPIFA, que representan más del 20% de los ácidos grasos totales, seguido por 6% en tallos y 5% en hojas (Figura 5).

En *B. juncea* no existe sustrato de ácido homo-gamma-linolénico (20:3 - 8, 11, 14) disponible. Por lo tanto, para examinar si la planta transgénica puede producir AA, se suministró en forma exógena el sustrato 20:3(8, 11, 14). En este experimento, se aplicaron tanto de tipo Silvestre como transgénicos con una solución acuosa de homo-gamma-linolenato sódico. Se encontró que los sustratos aplicados en forma exógena fueron fácilmente absorbidos por las raíces, tallo, y hojas de plantas transgénicas y convertidos en AA en las plantas (Figura 6).

No hubo un efecto fenotípico observable sobre el crecimiento y desarrollo en la *B. juncea* transgénica, aunque los $\Delta 5$ -UPIFA se acumularon en todas las partes de la planta. El crecimiento y diferenciación de los tejidos vegetativos tales como las hojas, tallos, y raíces fueron indistinguibles de los correspondientes de tipo silvestre.

Para producir ácidos grasos $\Delta 5$ desaturados en semillas, se transformaron *B. juncea* con la construcción que contiene el ADNc para Fad5-2 detrás de un promotor heterólogo específico de la semilla (promotor de napina de *B. napus*). El análisis de ácidos grasos de semillas transgénicas mostró que había dos nuevos ácidos grasos que aparecen en el cromatograma de gases de transgénicos comparado con el control de tipo silvestre (Figura 7). Fueron identificados como el ácido taxoleico (18:2 - 5, 9) y los ácidos pinolénicos (18:3 - 5, 9, 12). Juntos, estos ácidos grasos representan 9,4 % de los ácidos grasos de la semilla. La acumulación de los $\Delta 5$ -UPIFA no tiene un efecto significativo sobre el contenido de ácido oleico comparado con el control no transformado.

Ejemplo 10: Expresión de fad5-2 en linaza

Para producir ácidos grasos $\Delta 5$ desaturados en semillas de linaza, se transformó el linaza con Fad5-2 bajo el control de dos promotores específicos de la semilla, un promotor heterólogo de napina de *B. napus*, y un promotor endógeno de linaza. Como se muestra en la Figura 26, las plantas transgénicas que contienen al promotor de napina produjeron un ácido graso $\Delta 5$ -desaturado, ácido taxoleico en semillas en un nivel de menos del 1% de los ácidos grasos totales. Mientras que las plantas transgénicas que contenían al promotor específico de la semilla de linaza produjeron tres ácidos grasos $\Delta 5$ desaturados: ácido taxoleico, pinolénico, y coniferónico. De estos, el taxoleico (18:2 - 5, 9) era el más abundante y representó más del 9% de los ácidos grasos totales en un línea élite (FN-10-1), seguido por los ácidos coniferónico y pinolénico. Sorprendentemente, la acumulación de ácidos grasos $\Delta 5$ desaturados en semillas transgénicas tiene un impacto significativo sobre la composición de otros ácidos grasos, especialmente el nivel de ácido oleico. La acumulación de los $\Delta 5$ -UPIFA estaba acompañada por un enorme aumento de los ácidos oleicos en ambos tipos de plantas transgénicas que expresan desaturasa de Fad5-2 bajo el control de los diferentes promotores. El contenido de ácido oleico en plantas transgénicas con los promotores específicos de la semilla de linaza y de napina, en promedio, alcanzó 44,7% y 24,3% de los ácidos grasos totales, respectivamente, con relación al control no transformado de 17,4%.

Equivalentes

Aquellos capacitados en el arte se darán cuenta que, o serán capaces de determinar utilizando únicamente experimentación de rutina, muchos equivalentes a las modalidades específicas de la invención descritas aquí.

LISTADO DE SECUENCIAS

<110> Bioriginal Food & science Corp. et al.

<120> FAD4, FAD5, FAD5-2, Y FAD6, NUEVOS MIEMBROS DE LA FAMILIA DE LA DESATURASA DE ÁCIDO GRASO Y USOS DE LOS MISMOS

<130> BNZ-001PC

<140> PCT/IB01/02346

<141> 2001-09-28

<150> 60/236.303

<151> 2000-09-28

<150> 60/257.562

<151> 2001-06-12

<160> 10

<170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 1560

<212> ADN

<213> Thraustochytrium sp.

<220>

<221> CDS

5 <222> (1)...(1560)

<221> Variante

<222> 462

<223> xaa = Gly

<400> 1

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

cgc Arg	cgc Arg 130	gga Gly	ggc Gly	tac Tyr	gag Glu	ctc Leu 135	tgg Trp	atc Ile	aag Lys	gcg Ala	ttc Phe 140	ctg Leu	ctg Leu	ctc Leu	gtc Val	432
ggc Gly 145	ttc Phe	tgg Trp	agc Ser	tcg Ser	ctg Leu 150	tac Tyr	tgg Trp	atg Met	tgc Cys	acg Thr 155	ctg Leu	gac Asp	ccc Pro	tcg Ser	ttc Phe 160	480
ggg Gly	gcc Ala	atc Ile	ctg Leu	gcc Ala 165	gcc Ala	atg Met	tcg Ser	ctg Leu	ggc Gly 170	gtc Val	ttt Phe	gcc Ala	gcc Ala	ttt Phe 175	gtg Val	528
ggc Gly	acg Thr	tgc Cys	atc Ile 180	cag Gln	cac His	gac Asp	ggc Gly	aac Asn 185	cac His	ggc Gly	gcc Ala	ttt Phe 190	gcc Ala	cag Gln	tcg Ser	576
cga Arg	tgg Trp	gtc Val 195	aac Asn	aag Lys	gtt Val	gcc Ala	ggg Gly 200	tgg Trp	acg Thr	ctc Leu	gac Asp 205	atg Met	atc Ile	ggc Gly	gcc Ala	624
agc Ser	ggc Gly 210	atg Met	acg Thr	tgg Trp	gag Glu	ttc Phe 215	cag Gln	cac His	gtc Val	ctg Leu	ggc Gly 220	cac His	cat His	ccg Pro	tac Tyr	672
acg Thr 225	aac Asn	ctg Leu	atc Ile	gag Glu	gag Glu 230	gag Glu	aac Asn	ggc Gly	ctg Leu	caa Gln 235	aag Lys	gtg Val	agc Ser	ggc Gly	aag Lys 240	720
aag Lys	atg Met	gac Asp	acc Thr	aag Lys 245	ctg Leu	gcc Ala	gac Asp	cag Gln	gag Glu 250	agc Ser	gat Asp	ccg Pro	gac Asp	gtc Val 255	ttt Phe	768
tcc Ser	acg Thr	tac Tyr	ccg Pro 260	atg Met	atg Met	cgc Arg	ctg Leu	cac His 265	ccg Pro	tgg Trp	cac His	cag Gln 270	aag Lys	cgc Arg	tgg Trp	816
tac Tyr	cac His	cgt Arg 275	ttc Phe	cag Gln	cac His	att Ile	tac Tyr 280	ggc Gly	ccc Pro	ttc Phe	atc Ile	ttt Phe 285	ggc Gly	ttc Phe	atg Met	864
acc Thr 290	atc Ile	aac Asn	aag Lys	gtg Val	gtc Val	acg Thr 295	cag Gln	gac Asp	gtc Val	ggt Gly	gtg Val 300	gtg Val	ctc Leu	cgc Arg	aag Lys	912
cgg Arg 305	ctc Leu	ttc Phe	cag Gln	att Ile	gac Asp 310	gcc Ala	gag Glu	tgc Cys	cgg Arg	tac Tyr 315	gcg Ala	agc Ser	cca Pro	atg Met	tac Tyr 320	960
gtg Val	gcg Ala	cgt Arg	ttc Phe	tgg Trp 325	atc Ile	atg Met	aag Lys	gcg Ala	ctc Leu 330	acg Thr	gtg Val	ctc Leu	tac Tyr	atg Met 335	gtg Val	1008
gcc Ala	ctg Leu	ccg Pro	tgc Cys 340	tac Tyr	atg Met	cag Gln	ggc Gly	ccg Pro 345	tgg Trp	cac His	ggc Gly	ctc Leu	aag Lys 350	ctg Leu	ttc Phe	1056
gcg Ala	atc Ile	gcg Ala 355	cac His	ttt Phe	acg Thr	tgc Cys	ggc Gly 360	gag Glu	gtg Val	ctc Leu	gca Ala 365	acc Thr	atg Met	ttc Phe	att Ile	1104
gtg Val	aac Asn 370	cac His	atc Ile	atc Ile	gag Glu	ggc Gly 375	gtc Val	tcg Ser	tac Tyr	gct Ala	tcc Ser 380	aag Lys	gac Asp	gcg Ala	gtc Val	1152
aag Lys 385	ggc Gly	acg Thr	atg Met	gcg Ala	ccg Pro 390	ccg Pro	aag Lys	acg Thr	atg Met	cac His 395	ggc Gly	gtg Val	acg Thr	ccc Pro	atg Met 400	1200

aac aac acg cgc aag gag gtg gag gcg gag gcg tcc aag tct ggc gcc	1248
Asn Asn Thr Arg Lys 405 Glu Val Glu Ala 410 Glu Ala Ser Lys Ser 415 Gly Ala	
gtg gtc aag tca gtc ccg ctc gac gac tgg gcc gtc gtc cag tgc cag	1296
Val Val Lys 420 Ser Val Pro Leu Asp 425 Trp Ala Val Val 430 Gln Cys Gln	
acc tcg gtg aac tgg agc gtc ggc tcg tgg ttc tgg aat cac ttt tcc	1344
Thr Ser Val 435 Asn Trp Ser Val 440 Gly Ser Trp Phe Trp Asn 445 His Phe Ser	
ggc ggc ctc aac cac cag att gag cac cac ctg ttc ccc ggr ctc agc	1392
Gly Gly 450 Leu Asn His 455 Glu Ile 455 Glu His His Leu 460 Phe Pro Xaa Leu Ser	
cac gag acg tac tac cac att cag gac gtc ttt cag tcc acc tgc gcc	1440
His 465 Glu Thr Tyr Tyr 470 His Ile Gln Asp Val 475 Phe Gln Ser Thr Cys Ala 480	
gag tac ggc gtc ccg tac cag cac gag cct tcg ctc tgg acc gcg tac	1488
Glu Tyr Gly Val 485 Pro Tyr Gln His Glu 490 Pro Ser Leu Trp Thr 495 Ala Tyr	
tgg aag atg ctc gag cac ctc cgt cag ctc ggc aat gag gag acc cac	1536
Trp Lys Met 500 Leu Glu His Leu Arg 505 Gln Leu Gly Asn Glu 510 Glu Thr His	
gag tcc tgg cag cgc gct gcc tga	1560
Glu Ser Trp 515 Gln Arg Ala Ala *	

<210> 2

<211> 519

5 <212> PRT

<213> Thraustochytrium sp.

<220>

<221> VARIANTE

<222> 462

10 <223> xaa = Gly

<400> 2

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

Arg	Trp	Val	180 Asn	Lys	Val	Ala	Gly	185 Trp	Thr	Leu	Asp	Met	190 Ile	Gly	Ala
Ser	Gly	Met	195 Thr	Trp	Glu	Phe	200 Gln	His	Val	Leu	Gly	205 His	His	Pro	Tyr
Thr	210 Asn	Leu	Ile	Glu	Glu	215 Glu	Asn	Gly	Leu	Gln	Lys	220 Val	Ser	Gly	Lys
225	Met	Asp	Thr	Lys	230 Leu	Ala	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe
Lys	Met	Asp	Thr	Lys	245 Leu	Ala	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe
Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	His	Pro	Trp	His	Gln	Lys	Arg	Trp
Tyr	His	Arg	260 Phe	Gln	His	Ile	Tyr	265 Gly	Pro	Phe	Ile	Phe	Gly	Phe	Met
Thr	Ile	Asn	Lys	Val	Val	Thr	280 Gln	Asp	Val	Gly	Val	285 Val	Leu	Arg	Lys
Arg	Leu	Phe	Gln	Ile	Asp	Ala	Glu	Cys	Arg	Tyr	Ala	Ser	Pro	Met	Tyr
305	Ala	Arg	Phe	Trp	Ile	Met	Lys	Ala	Leu	Thr	Val	Leu	Tyr	Met	Val
Val	Ala	Arg	Phe	Trp	Ile	Met	Lys	Ala	Leu	Thr	Val	Leu	Tyr	Met	Val
Ala	Leu	Pro	Cys	Tyr	Met	Gln	Gly	Pro	Trp	His	Gly	Leu	Lys	Leu	Phe
Ala	Ile	Ala	His	Phe	Thr	Cys	Gly	Glu	Val	Leu	Ala	Thr	Met	Phe	Ile
Val	Asn	His	Ile	Ile	Glu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val
Lys	Gly	Thr	Met	Ala	Pro	Pro	Lys	Thr	Met	His	Gly	Val	Thr	Pro	Met
385	Asn	Asn	Thr	Arg	Lys	Glu	Val	Glu	Ala	Glu	Ala	Ser	Lys	Ser	Gly
Asn	Asn	Thr	Arg	Lys	Glu	Val	Glu	Ala	Glu	Ala	Ser	Lys	Ser	Gly	Ala
Val	Val	Lys	Ser	Val	Pro	Leu	Asp	Asp	Trp	Ala	Val	Val	Gln	Cys	Gln
Thr	Ser	Val	Asn	Trp	Ser	Val	Gly	Ser	Trp	Phe	Trp	Asn	His	Phe	Ser
Gly	Gly	Leu	Asn	His	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Xaa	Leu	Ser
His	Glu	Thr	Tyr	Tyr	His	Ile	Gln	Asp	Val	Phe	Gln	Ser	Thr	Cys	Ala
465	Glu	Tyr	Gly	Val	Pro	Tyr	Gln	His	Glu	Pro	Ser	Leu	Trp	Thr	Ala
Glu	Tyr	Gly	Val	Pro	Tyr	Gln	His	Glu	Pro	Ser	Leu	Trp	Thr	Ala	Tyr
Trp	Lys	Met	Leu	Glu	His	Leu	Arg	Gln	Leu	Gly	Asn	Glu	Glu	Thr	His
Glu	Ser	Trp	Gln	Arg	Ala	Ala		505					510		

<210> 3

<211> 1320

5 <212> ADN

<213> Thraustochytrium sp.

<220>

<221> CDS

<222> (1)...(1320)

10 <400> 3

atg	ggc	aag	ggc	agc	gag	ggc	cgc	agc	gcg	gcg	cgc	gag	atg	acg	gcc	48
Met	Gly	Lys	Gly	Ser	Glu	Gly	Arg	Ser	Ala	Ala	Arg	Glu	Met	Thr	Ala	
1				5					10					15		
gag	gcg	aac	ggc	gac	aag	cgg	aaa	acg	att	ctg	atc	gag	ggc	gtc	ctg	96
Glu	Ala	Asn	Gly	Asp	Lys	Arg	Lys	Thr	Ile	Leu	Ile	Glu	Gly	Val	Leu	
			20					25					30			
tac	gac	gcg	acg	aac	ttt	aag	cac	ccg	ggc	ggt	tcg	atc	atc	aac	ttc	144
Tyr	Asp	Ala	Thr	Asn	Phe	Lys	His	Pro	Gly	Gly	Ser	Ile	Ile	Asn	Phe	
		35					40					45				
ttg	acc	gag	ggc	gag	gcc	ggc	gtg	gac	gcg	acg	cag	gcg	tac	cgc	gag	192

Leu	Thr	Glu	Gly	Glu	Ala	Gly	Val	Asp	Ala	Thr	Gln	Ala	Tyr	Arg	Glu	
	50					55					60					
ttt	cat	cag	cgg	tcc	ggc	aag	gcc	gac	aag	tac	ctc	aag	tcg	ctg	ccg	240
Phe	His	Gln	Arg	Ser	Gly	Lys	Ala	Asp	Lys	Tyr	Leu	Lys	Ser	Leu	Pro	80
65					70					75						
aag	ctg	gat	gcg	tcc	aag	gtg	gag	tcg	cgg	ttc	tcg	gcc	aaa	gag	cag	288
Lys	Leu	Asp	Ala	Ser	Lys	Val	Glu	Ser	Arg	Phe	Ser	Ala	Lys	Glu	Gln	95
				85					90					95		
gcg	cgg	cgc	gac	gcc	atg	acg	cgc	gac	tac	gcg	gcc	ttt	cgc	gag	gag	336
Ala	Arg	Arg	Asp	Ala	Met	Thr	Arg	Asp	Tyr	Ala	Ala	Phe	Arg	Glu	Glu	
			100					105					110			
ctc	gtc	gcc	gag	ggg	tac	ttt	gac	ccg	tcg	atc	ccg	cac	atg	att	tac	384
Leu	Val	Ala	Glu	Gly	Tyr	Phe	Asp	Pro	Ser	Ile	Pro	His	Met	Ile	Tyr	
		115					120					125				
cgc	gtc	gtg	gag	atc	gtg	gcg	ctc	ttc	gcg	ctc	tcg	ttc	tgg	ctc	atg	432
Arg	Val	Val	Glu	Ile	Val	Ala	Leu	Phe	Ala	Leu	Ser	Phe	Trp	Leu	Met	
	130					135					140					
tcc	aag	gcc	tcg	ccc	acc	tcg	ctc	gtg	ctg	ggc	gtg	gtg	atg	aac	ggc	480
Ser	Lys	Ala	Ser	Pro	Thr	Ser	Leu	Val	Leu	Gly	Val	Val	Met	Asn	Gly	160
145					150					155						
att	gcg	cag	ggc	cgc	tgc	ggc	tgg	gtc	atg	cac	gag	atg	ggc	cac	ggg	528
Ile	Ala	Gln	Gly	Arg	Cys	Gly	Trp	Val	Met	His	Glu	Met	Gly	His	Gly	
				165					170				175			
tcg	ttc	acg	ggc	gtc	atc	tgg	ctc	gac	gac	cgg	atg	tgc	gag	ttc	ttc	576
Ser	Phe	Thr	Gly	Val	Ile	Trp	Leu	Asp	Asp	Arg	Met	Cys	Glu	Phe	Phe	
			180					185					190			
tac	ggc	gtc	ggc	tgc	ggc	atg	agc	ggg	cac	tac	tgg	aag	aac	cag	cac	624
Tyr	Gly	Val	Gly	Cys	Gly	Met	Ser	Gly	His	Tyr	Trp	Lys	Asn	Gln	His	
		195					200					205				
agc	aag	cac	cac	gcc	gcg	ccc	aac	cgc	ctc	gag	cac	gat	gtc	gat	ctc	672
Ser	Lys	His	His	Ala	Ala	Pro	Asn	Arg	Leu	Glu	His	Asp	Val	Asp	Leu	
	210					215					220					
aac	acg	ctg	ccc	ctg	gtc	gcc	ttt	aac	gag	cgc	gtc	gtg	cgc	aag	gtc	720
Asn	Thr	Leu	Pro	Leu	Val	Ala	Phe	Asn	Glu	Arg	Val	Val	Arg	Lys	Val	240
225					230					235						
aag	ccg	gga	tcg	ctg	ctg	gcg	ctc	tgg	ctg	cgc	gtg	cag	gcg	tac	ctc	768
Lys	Pro	Gly	Ser	Leu	Leu	Ala	Leu	Trp	Leu	Arg	Val	Gln	Ala	Tyr	Leu	
				245					250					255		
ttt	gcg	ccc	gtc	tcg	tgc	ctg	ctc	atc	ggc	ctt	ggc	tgg	acg	ctc	tac	816
Phe	Ala	Pro	Val	Ser	Cys	Leu	Leu	Ile	Gly	Leu	Gly	Trp	Thr	Leu	Tyr	
			260					265					270			
ctg	cac	ccg	cgc	tac	atg	ctg	cgc	acc	aag	cgg	cac	atg	gag	ttc	gtc	864
Leu	His	Pro	Arg	Tyr	Met	Leu	Arg	Thr	Lys	Arg	His	Met	Glu	Phe	Val	
		275					280					285				
tgg	atc	ttc	gcg	cgc	tac	att	ggc	tgg	ttc	tcg	ctc	atg	ggc	gct	ctc	912
Trp	Ile	Phe	Ala	Arg	Tyr	Ile	Gly	Trp	Phe	Ser	Leu	Met	Gly	Ala	Leu	
	290					295					300					
ggc	tac	tcg	ccg	ggc	acc	tcg	gtc	ggg	atg	tac	ctg	tgc	tcg	ttc	ggc	960
Gly	Tyr	Ser	Pro	Gly	Thr	Ser	Val	Gly	Met	Tyr	Leu	Cys	Ser	Phe	Gly	320
305					310					315						
ctc	ggc	tgc	att	tac	att	ttc	ctg	cag	ttc	gcc	gtc	agc	cac	acg	cac	1008

Leu	Gly	Cys	Ile	Tyr 325	Ile	Phe	Leu	Gln	Phe 330	Ala	Val	Ser	His 335	Thr	His		
ctg	ccg	gtg	acc	aac	ccg	gag	gac	cag	ctg	cac	tgg	ctc	gag	tac	gcg	1056	
Leu	Pro	Val	Thr 340	Asn	Pro	Glu	Asp	Gln 345	Leu	His	Trp	Leu	Glu 350	Tyr	Ala		
gcc	gac	cac	acg	gtg	aac	att	agc	acc	aag	tcc	tgg	ctc	gtc	acg	tgg	1104	
Ala	Asp	His 355	Thr	Val	Asn	Ile	Ser 360	Thr	Lys	Ser	Trp	Leu	Val 365	Thr	Trp		
tgg	atg	tcg	aac	ctg	aac	ttt	cag	atc	gag	cac	cac	ctc	ttc	ccc	acg	1152	
Trp	Met 370	Ser	Asn	Leu	Asn	Phe 375	Gln	Ile	Glu	His	His 380	Leu	Phe	Pro	Thr		
gcg	ccg	cag	ttc	cgc	ttc	aag	gaa	atc	agt	cct	cgc	gtc	gag	gcc	ctc	1200	
Ala	Pro	Gln	Phe	Arg	Phe 390	Lys	Glu	Ile	Ser	Pro 395	Arg	Val	Glu	Ala	Leu 400		
ttc	aag	cgc	cac	aac	ctc	ccg	tac	tac	gac	ctg	ccc	tac	acg	agc	gcg	1248	
Phe	Lys	Arg	His 405	Asn	Leu	Pro	Tyr	Tyr	Asp 410	Leu	Pro	Tyr	Thr	Ser 415	Ala		
gtc	tcg	acc	acc	ttt	gcc	aat	ctt	tat	tcc	gtc	ggc	cac	tcg	gtc	ggc	1296	
Val	Ser	Thr	Thr 420	Phe	Ala	Asn	Leu	Tyr 425	Ser	Val	Gly	His	Ser 430	Val	Gly		
gcc	gac	acc	aag	aag	cag	gac	tga									1320	
Ala	Asp	Thr 435	Lys	Lys	Gln	Asp	*										

<210> 4

<211> 439

5 <212> PRT

<213> Thraustochytrium sp.

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

Lys Pro Gly Ser Leu Leu Ala Leu Trp Leu Arg Val Gln Ala Tyr Leu
 245 250 255
 Phe Ala Pro Val Ser Cys Leu Leu Ile Gly Leu Gly Trp Thr Leu Tyr
 260 265 270
 Leu His Pro Arg Tyr Met Leu Arg Thr Lys Arg His Met Glu Phe Val
 275 280 285
 Trp Ile Phe Ala Arg Tyr Ile Gly Trp Phe Ser Leu Met Gly Ala Leu
 290 295 300
 Gly Tyr Ser Pro Gly Thr Ser Val Gly Met Tyr Leu Cys Ser Phe Gly
 305 310 315 320
 Leu Gly Cys Ile Tyr Ile Phe Leu Gln Phe Ala Val Ser His Thr His
 325 330 335
 Leu Pro Val Thr Asn Pro Glu Asp Gln Leu His Trp Leu Glu Tyr Ala
 340 345 350
 Ala Asp His Thr Val Asn Ile Ser Thr Lys Ser Trp Leu Val Thr Trp
 355 360 365
 Trp Met Ser Asn Leu Asn Phe Gln Ile Glu His His Leu Phe Pro Thr
 370 375 380
 Ala Pro Gln Phe Arg Phe Lys Glu Ile Ser Pro Arg Val Glu Ala Leu
 385 390 395 400
 Phe Lys Arg His Asn Leu Pro Tyr Tyr Asp Leu Pro Tyr Thr Ser Ala
 405 410 415
 Val Ser Thr Thr Phe Ala Asn Leu Tyr Ser Val Gly His Ser Val Gly
 420 425 430
 Ala Asp Thr Lys Lys Gln Asp
 435

- 5 <210> 5
 <211> 1371
 <212> ADN
 <213> Thraustochytrium sp.
 <220>

<221> CDS

<222> (1)...(1368)

<400> 5

atg acc gag aag gcg agt gac gag ttc acg tgg cag gag gtc gcc aag	48
Met Thr Glu Lys Ala Ser Asp Glu Phe Thr Trp Gln Glu Val Ala Lys	
1 5 10 15	
cac aac acg gcc aag agc gcg tgg gtg atc atc cgc ggc gag gtg tac	96
His Asn Thr Ala Lys Ser Ala Trp Val Ile Ile Arg Gly Glu Val Tyr	
20 25 30	
gac gtg acc gag tgg gcg gac aag cac ccg ggc ggc agc gag ctc atc	144
Asp Val Thr Glu Trp Ala Asp Lys His Pro Gly Gly Ser Glu Leu Ile	
35 40 45	
gtc ctg cac tcc ggt cgt gaa tgc acg gac acg ttc tac tcg tac cac	192
Val Leu His Ser Gly Arg Glu Cys Thr Asp Thr Phe Tyr Ser Tyr His	
50 55 60	
ccg ttc tcg aac cgc gcc gac aag atc ttg gcc aag tac aag atc ggc	240
Pro Phe Ser Asn Arg Ala Asp Lys Ile Leu Ala Lys Tyr Lys Ile Gly	
65 70 75 80	
aag ctc gtg ggc ggc tac gag ttc ccg gtg ttc aag ccg gac tcg ggc	288
Lys Leu Val Gly Gly Tyr Glu Phe Pro Val Phe Lys Pro Asp Ser Gly	
85 90 95	
ttc tac aag gaa tgc tcg gag cgc gtg gcc gag tac ttt aag acg aac	336
Phe Tyr Lys Glu Cys Ser Glu Arg Val Ala Glu Tyr Phe Lys Thr Asn	
100 105 110	
aat ctg gac cca aag gcg gcg ttc gcg ggt ctc tgg cgc atg gtg ttc	384
Asn Leu Asp Pro Lys Ala Ala Phe Ala Gly Leu Trp Arg Met Val Phe	

5

115				120				125								
gtg Val	ttc Phe 130	gcg Ala	gtc Val	gcc Ala	gcg Ala	ctc Leu 135	gcg Ala	tac Tyr	atg Met	ggc Gly 140	atg Met 140	aat Asn	gag Glu	ctc Leu	atc Ile	432
cct Pro 145	gga Gly	aac Asn	gtg Val	tac Tyr	gcg Ala 150	cag Gln	tac Tyr	gcg Ala	tgg Trp	ggc Gly 155	gtg Val	gtg Val	ttc Phe	ggt Gly	gtc Val 160	480
ttc Phe	cag Gln	gcg Ala	ctg Leu	cca Pro 165	ttg Leu	ctg Leu	cac His	gtg Val	atg Met 170	cac His	gac Asp	tcg Ser	tcg Ser	cac His 175	gcg Ala	528
gca Ala	tgc Cys	tcg Ser	agc Ser 180	agc Ser	cca Pro	gcg Ala	atg Met	tgg Trp 185	cag Gln	atc Ile	atc Ile	ggt Gly	cgt Arg 190	ggt Gly	gtg Val	576
atg Met	gac Asp	tgg Trp 195	ttc Phe	gct Ala	ggc Gly	gcc Ala	agc Ser 200	atg Met	gtg Val	tcg Ser	tgg Trp	ttg Leu 205	aac Asn	cag Gln	cac His	624
gtt Val	gtg Val 210	ggc Gly	cac His	cac His	atc Ile	tac Tyr 215	acg Thr	aac Asn	gtc Val	gcg Ala	ggc Gly 220	gcg Ala	gac Asp	ccg Pro	gat Asp	672
ctc Leu 225	ccg Pro	gtc Val	gac Asp	ttt Phe	gag Glu 230	agc Ser	gac Asp	gtg Val	cgc Arg	cgc Arg 235	atc Ile	gtg Val	cac His	cgc Arg	cag Gln 240	720
gtg Val	ctg Leu	ctg Leu	ccg Pro	atc Ile 245	tac Tyr	aag Lys	ttc Phe	cag Gln	cac His 250	atc Ile	tac Tyr	ctg Leu	cca Pro	ccg Pro 255	ctc Leu	768
tac Tyr	ggc Gly	gtg Val	ctg Leu 260	ggc Gly	ctc Leu	aag Lys	ttc Phe	cgc Arg 265	atc Ile	cag Gln	gac Asp	gtg Val	ttc Phe 270	gag Glu	acg Thr	816
ttc Phe	gtg Val	tcg Ser 275	ctc Leu	acg Thr	aac Asn	ggc Gly	ccg Pro 280	gtg Val	cgt Arg	gtg Val	aac Asn	ccg Pro 285	cac His	ccg Pro	gtg Val	864
tcg Ser	gac Asp 290	tgg Trp	gtg Val	caa Gln	atg Met	atc Ile 295	ttc Phe	gcc Ala	aag Lys	gcg Ala	ttc Phe 300	tgg Trp	acg Thr	ttc Phe	tac Tyr	912
cgc Arg 305	atc Ile	tac Tyr	atc Ile	ccg Pro	ttg Leu 310	gcg Ala	tgg Trp	ctc Leu	aag Lys	atc Ile 315	acg Thr	ccg Pro	tcg Ser	acg Thr	ttc Phe 320	960
tgg Trp	ggc Gly	gtg Val	ttt Phe	ttc Phe 325	ctc Leu	gcc Ala	gag Glu	ttc Phe	acc Thr 330	aca Thr	ggt Gly	tgg Trp	tac Tyr	ctc Leu 335	gcg Ala	1008
ttc Phe	aac Asn	ttc Phe	cag Gln 340	gtg Val	agc Ser	cac His	gtc Val	tcg Ser 345	acc Thr	gag Glu	tgc Cys	gag Glu	tac Tyr 350	ccg Pro	tgc Cys	1056
ggt Gly	gat Asp	gcg Ala 355	ccg Pro	tcg Ser	gcc Ala	gag Glu	gtc Val 360	ggt Gly	gac Asp	gag Glu	tgg Trp	gcg Ala 365	atc Ile	tcg Ser	cag Gln	1104
gtc Val	aag Lys 370	tcg Ser	tcg Ser	gtg Val	gac Asp	tac Tyr 375	gcg Ala	cac His	ggc Gly	tcg Ser	ccg Pro 380	ctc Leu	gcg Ala	gcg Ala	ttc Phe	1152
ctc Leu	tgc Cys	ggc Gly	gcg Ala	ctc Leu	aac Asn	tac Tyr	cag Gln	gtg Val	acc Thr	cac His	cac His	ttg Leu	tac Tyr	ccg Pro	ggc Gly	1200

385		390		395		400	
atc tca cag tac cac tac cct gcg atc gcg ccg atc atc atc gac gtg							1248
Ile Ser Gln Tyr His Tyr Pro Ala Ile Ala Pro Ile Ile Ile Asp Val							
		405		410		415	
tgc aag aag tac aac atc aag tac acg gtg ctg ccg acg ttc acc gag							1296
Cys Lys Lys Tyr Asn Ile Lys Tyr Thr Val Leu Pro Thr Phe Thr Glu							
		420		425		430	
gcg ctg ctc gcg cac ttc aag cac ctg aag aac atg ggc gag ctc ggc							1344
Ala Leu Leu Ala His Phe Lys His Leu Lys Asn Met Gly Glu Leu Gly							
		435		440		445	
aag ccc gtg gag atc cac atg ggt taa							1371
Lys Pro Val Glu Ile His Met Gly * * *							
		450		455			

<210> 6

5 <211> 456

<212> PRT

<213> Thraustochytrium sp.

<400> 6

Met Thr Glu Lys Ala Ser Asp Glu Phe Thr Trp Gln Glu Val Ala Lys
 1 5 10 15
 His Asn Thr Ala Lys Ser Ala Trp Val Ile Ile Arg Gly Glu Val Tyr
 20 25 30
 Asp Val Thr Glu Trp Ala Asp Lys His Pro Gly Gly Ser Glu Leu Ile
 35 40 45
 Val Leu His Ser Gly Arg Glu Cys Thr Asp Thr Phe Tyr Ser Tyr His
 50 55 60
 Pro Phe Ser Asn Arg Ala Asp Lys Ile Leu Ala Lys Tyr Lys Ile Gly
 65 70 75 80
 Lys Leu Val Gly Gly Tyr Glu Phe Pro Val Phe Lys Pro Asp Ser Gly
 85 90 95
 Phe Tyr Lys Glu Cys Ser Glu Arg Val Ala Glu Tyr Phe Lys Thr Asn
 100 105 110
 Asn Leu Asp Pro Lys Ala Ala Phe Ala Gly Leu Trp Arg Met Val Phe
 115 120 125
 Val Phe Ala Val Ala Ala Leu Ala Tyr Met Gly Met Asn Glu Leu Ile
 130 135 140
 Pro Gly Asn Val Tyr Ala Gln Tyr Ala Trp Gly Val Val Phe Gly Val
 145 150 155 160
 Phe Gln Ala Leu Pro Leu Leu His Val Met His Asp Ser Ser His Ala
 165 170 175
 Ala Cys Ser Ser Ser Pro Ala Met Trp Gln Ile Ile Gly Arg Gly Val
 180 185 190
 Met Asp Trp Phe Ala Gly Ala Ser Met Val Ser Trp Leu Asn Gln His
 195 200 205
 Val Val Gly His His Ile Tyr Thr Asn Val Ala Gly Ala Asp Pro Asp
 210 215 220
 Leu Pro Val Asp Phe Glu Ser Asp Val Arg Arg Ile Val His Arg Gln
 225 230 235 240
 Val Leu Leu Pro Ile Tyr Lys Phe Gln His Ile Tyr Leu Pro Pro Leu
 245 250 255
 Tyr Gly Val Leu Gly Leu Lys Phe Arg Ile Gln Asp Val Phe Glu Thr
 260 265 270
 Phe Val Ser Leu Thr Asn Gly Pro Val Arg Val Asn Pro His Pro Val
 275 280 285
 Ser Asp Trp Val Gln Met Ile Phe Ala Lys Ala Phe Trp Thr Phe Tyr
 290 295 300
 Arg Ile Tyr Ile Pro Leu Ala Trp Leu Lys Ile Thr Pro Ser Thr Phe
 305 310 315 320
 Trp Gly Val Phe Phe Leu Ala Glu Phe Thr Thr Gly Trp Tyr Leu Ala
 325 330 335
 Phe Asn Phe Gln Val Ser His Val Ser Thr Glu Cys Glu Tyr Pro Cys

Gly Asp Ala Pro Ser Ala Glu Val Gly Asp Glu Trp Ala Ile Ser Gln
 340 345 350
 Val Lys Ser Ser Val Asp Tyr Ala His Gly Ser Pro Leu Ala Ala Phe
 355 360 365
 Leu Cys Gly Ala Leu Asn Tyr Gln Val Thr His His Leu Tyr Pro Gly
 370 375 380 385 390 400
 Ile Ser Gln Tyr His Tyr Pro Ala Ile Ala Pro Ile Ile Ile Asp Val
 405 410 415
 Cys Lys Lys Tyr Asn Ile Lys Tyr Thr Val Leu Pro Thr Phe Thr Glu
 420 425 430
 Ala Leu Leu Ala His Phe Lys His Leu Lys Asn Met Gly Glu Leu Gly
 435 440 445
 Lys Pro Val Glu Ile His Met Gly
 450 455

- 5 <210> 7
 <211> 1380
 <212> ADN
 <213> Thraustochytrium sp.
 <220>

<221> CDS

<222> (1)...(1380)

<400> 7

atg	gtg	gac	ctc	aag	cct	gga	gtg	aag	cgc	ctg	gtg	agc	tgg	aag	gag	48
Met	Val	Asp	Leu	Lys	Pro	Gly	Val	Lys	Arg	Leu	Val	Ser	Trp	Lys	Glu	
1				5					10					15		
atc	cgc	gag	cac	gcg	acg	ccc	gcg	acc	gcg	tgg	atc	gtg	att	cac	cac	96
Ile	Arg	Glu	His	Ala	Thr	Pro	Ala	Thr	Ala	Trp	Ile	Val	Ile	His	His	
			20					25					30			
aag	gtc	tac	gac	atc	tcc	aag	tgg	gac	tcg	cac	ccg	ggg	ggc	tcc	gtg	144
Lys	Val	Tyr	Asp	Ile	Ser	Lys	Trp	Asp	Ser	His	Pro	Gly	Gly	Ser	Val	
		35					40					45				
atg	ctc	acg	cag	gcc	ggc	gag	gac	gcc	acg	gac	gcc	ttc	gcg	gtc	ttc	192
Met	Leu	Thr	Gln	Ala	Gly	Glu	Asp	Ala	Thr	Asp	Ala	Phe	Ala	Val	Phe	
	50					55					60					
cac	ccg	tcc	tcg	gcg	ctc	aag	ctg	ctc	gag	cag	ttc	tac	gtc	ggc	gac	240
His	Pro	Ser	Ser	Ala	Leu	Lys	Leu	Leu	Glu	Gln	Phe	Tyr	Val	Gly	Asp	
65					70				75					80		
gtg	gac	gaa	acc	tcc	aag	gcc	gag	atc	gag	ggg	gag	ccg	gcg	agc	gac	288
Val	Asp	Glu	Thr	Ser	Lys	Ala	Glu	Ile	Glu	Gly	Glu	Pro	Ala	Ser	Asp	
				85					90					95		
gag	gag	cgc	gcg	cgc	cgc	gag	cgc	atc	aac	gag	ttc	atc	gcg	tcc	tac	336
Glu	Glu	Arg	Ala	Arg	Arg	Glu	Arg	Ile	Asn	Glu	Phe	Ile	Ala	Ser	Tyr	
			100					105					110			
cgt	cgt	ctg	cgc	gtc	aag	gtc	aag	ggc	atg	ggg	ctc	tac	gac	gcc	agc	384
Arg	Arg	Leu	Arg	Val	Lys	Val	Lys	Gly	Met	Gly	Leu	Tyr	Asp	Ala	Ser	
		115					120					125				
gcg	ctc	tac	tac	gcg	tgg	aag	ctc	gtg	agc	acg	ttc	ggc	atc	gcg	gtg	432
Ala	Leu	Tyr	Tyr	Ala	Trp	Lys	Leu	Val	Ser	Thr	Phe	Gly	Ile	Ala	Val	
	130					135					140					
ctc	tcg	atg	gcg	atc	tgc	ttc	ttc	ttc	aac	agt	ttc	gcc	atg	tac	atg	480
Leu	Ser	Met	Ala	Ile	Cys	Phe	Phe	Phe	Asn	Ser	Phe	Ala	Met	Tyr	Met	
145					150					155					160	
gtc	gcc	ggc	gtg	att	atg	ggg	ctc	ttc	tac	cag	cag	tcc	gga	tgg	ctg	528

5

val	Ala	Gly	Val	Ile 165	Met	Gly	Leu	Phe	Tyr 170	Gln	Gln	Ser	Gly	Trp 175	Leu	
gcg Ala	cac His	gac Asp	ttc Phe 180	ttg Leu	cac His	aac Asn	cag Gln	gtg Val 185	tgc Cys	gag Glu	aac Asn	cgc Arg	acg Thr 190	ctc Leu	ggc Gly	576
aac Asn	ctt Leu	atc Ile 195	ggc Gly	tgc Cys	ctc Leu	gtg Val	ggc Gly 200	aac Asn	gcc Ala	tgg Trp	cag Gln	ggc Gly 205	ttc Phe	agc Ser	gtg Val	624
cag Gln	tgg Trp 210	tgg Trp	aag Lys	aac Asn	aag Lys	cac His 215	aac Asn	ctg Leu	cac His	cac His	gcg Ala 220	gtg Val	ccg Pro	aac Asn	ctg Leu	672
cac His 225	agc Ser	gcc Ala	aag Lys	gac Asp	gag Glu 230	ggc Gly	ttc Phe	atc Ile	ggc Gly	gac Asp 235	ccg Pro	gac Asp	atc Ile	gac Asp	acc Thr 240	720
atg Met	ccg Pro	ctg Leu	ctg Leu	gcg Ala 245	tgg Trp	tct Ser	aag Lys	gag Glu	atg Met 250	gcg Ala	cgc Arg	aag Lys	gcg Ala	ttc Phe 255	gag Glu	768
tcg Ser	gcg Ala	cac His	ggc Gly 260	ccg Pro	ttc Phe	ttc Phe	atc Ile	cgc Arg 265	aac Asn	cag Gln	gcg Ala	ttc Phe	cta Leu 270	tac Tyr	ttc Phe	816
ccg Pro	ctg Leu	ctg Leu 275	ctg Leu	ctc Leu	gcg Ala	cgc Arg	ctg Leu 280	agc Ser	tgg Trp	ctc Leu	gcg Ala	cag Gln 285	tcg Ser	ttc Phe	ttc Phe	864
tac Tyr	gtg Val 290	ttc Phe	acc Thr	gag Glu	ttc Phe	tcg Ser 295	ttc Phe	ggc Gly	atc Ile	ttc Phe	gac Asp 300	aag Lys	gtc Val	gag Glu	ttc Phe	912
gac Asp 305	gga Gly	ccg Pro	gag Glu	aag Lys	gcg Ala 310	ggt Gly	ctg Leu	atc Ile	gtg Val	cac His 315	tac Tyr	atc Ile	tgg Trp	cag Gln	ctc Leu 320	960
gcg Ala	atc Ile	ccg Pro	tac Tyr	ttc Phe 325	tgc Cys	aac Asn	atg Met	agc Ser	ctg Leu 330	ttt Phe	gag Glu	ggc Gly	gtg Val	gca Ala 335	tac Tyr	1008
ttc Phe	ctc Leu	atg Met	ggc Gly 340	cag Gln	gcg Ala	tcc Ser	tgc Cys	ggc Gly 345	ttg Leu	ctc Leu	ctg Leu	gcg Ala	ctg Leu 350	gtg Val	ttc Phe	1056
agt Ser	att Ile	ggc Gly 355	cac His	aac Asn	ggc Gly	atg Met	tcg Ser 360	gtg Val	tac Tyr	gag Glu	cgc Arg	gaa Glu 365	acc Thr	aag Lys	ccg Pro	1104
gac Asp	ttc Phe 370	tgg Trp	cag Gln	ctg Leu	cag Gln	gtg Val 375	acc Thr	acg Thr	acg Thr	cgc Arg	aac Asn 380	atc Ile	cgc Arg	gcg Ala	tcg Ser	1152
gta Val 385	ttc Phe	atg Met	gac Asp	tgg Trp	ttc Phe 390	acc Thr	ggt Gly	ggc Gly	ttg Leu	aac Asn 395	tac Tyr	cag Gln	atc Ile	gac Asp	cat His 400	1200
cac His	ctg Leu	ttc Phe	ccg Pro	ctc Leu 405	gtg Val	ccg Pro	cgc Arg	cac His	aac Asn 410	ttg Leu	cca Pro	aag Lys	gtc Val	aac Asn 415	gtg Val	1248
ctc Leu	atc Ile	aag Lys	tcg Ser 420	cta Leu	tgc Cys	aag Lys	gag Glu	ttc Phe 425	gac Asp	atc Ile	ccg Pro	ttc Phe	cac His 430	gag Glu	acc Thr	1296
ggc	ttc	tgg	gag	ggc	atc	tac	gag	gtc	gtg	gac	cac	ctg	gcg	gac	atc	1344

Gly Phe Trp Glu Gly Ile Tyr Glu Val Val Asp His Leu Ala Asp Ile
435 440 445

agc aag gaa ttc atc acc gag ttc cca gcg atg taa 1380
Ser Lys Glu Phe Ile Thr Glu Phe Pro Ala Met *

450 455

<210> 8

<211> 459

5 <212> PRT

<213> Thraustochytrium sp.

<400> 8

Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Glu
 1 10 15
 Ile Arg Glu His Ala Thr Pro Ala Thr Ala Trp Ile Val Ile His His
 20 25 30
 Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val
 35 40 45
 Met Leu Thr Gln Ala Gly Glu Asp Ala Thr Asp Ala Phe Ala Val Phe
 50 55 60
 His Pro Ser Ser Ala Leu Lys Leu Leu Glu Gln Phe Tyr Val Gly Asp
 65 70 75 80
 Val Asp Glu Thr Ser Lys Ala Glu Ile Glu Gly Glu Pro Ala Ser Asp
 85 90 95
 Glu Glu Arg Ala Arg Arg Glu Arg Ile Asn Glu Phe Ile Ala Ser Tyr
 100 105 110
 Arg Arg Leu Arg Val Lys Val Lys Gly Met Gly Leu Tyr Asp Ala Ser
 115 120 125
 Ala Leu Tyr Tyr Ala Trp Lys Leu Val Ser Thr Phe Gly Ile Ala Val
 130 135 140
 Leu Ser Met Ala Ile Cys Phe Phe Phe Asn Ser Phe Ala Met Tyr Met
 145 150 155 160
 Val Ala Gly Val Ile Met Gly Leu Phe Tyr Gln Gln Ser Gly Trp Leu
 165 170 175
 Ala His Asp Phe Leu His Asn Gln Val Cys Glu Asn Arg Thr Leu Gly
 180 185 190
 Asn Leu Ile Gly Cys Leu Val Gly Asn Ala Trp Gln Gly Phe Ser Val
 195 200 205
 Gln Trp Trp Lys Asn Lys His Asn Leu His His Ala Val Pro Asn Leu
 210 215 220
 His Ser Ala Lys Asp Glu Gly Phe Ile Gly Asp Pro Asp Ile Asp Thr
 225 230 235 240
 Met Pro Leu Leu Ala Trp Ser Lys Glu Met Ala Arg Lys Ala Phe Glu
 245 250 255
 Ser Ala His Gly Pro Phe Phe Ile Arg Asn Gln Ala Phe Leu Tyr Phe
 260 265 270
 Pro Leu Leu Leu Leu Ala Arg Leu Ser Trp Leu Ala Gln Ser Phe Phe
 275 280 285
 Tyr Val Phe Thr Glu Phe Ser Phe Gly Ile Phe Asp Lys Val Glu Phe
 290 295 300
 Asp Gly Pro Glu Lys Ala Gly Leu Ile Val His Tyr Ile Trp Gln Leu
 305 310 315 320
 Ala Ile Pro Tyr Phe Cys Asn Met Ser Leu Phe Glu Gly Val Ala Tyr
 325 330 335
 Phe Leu Met Gly Gln Ala Ser Cys Gly Leu Leu Leu Ala Leu Val Phe
 340 345 350
 Ser Ile Gly His Asn Gly Met Ser Val Tyr Glu Arg Glu Thr Lys Pro
 355 360 365
 Asp Phe Trp Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Arg Ala Ser
 370 375 380
 Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Asp His
 385 390 395 400
 His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val
 405 410 415
 Leu Ile Lys Ser Leu Cys Lys Glu Phe Asp Ile Pro Phe His Glu Thr
 420 425 430

Gly Phe Trp Glu Gly Ile Tyr Glu Val Val Asp His Leu Ala Asp Ile
 435 440 445
 Ser Lys Glu Phe Ile Thr Glu Phe Pro Ala Met
 450 455

- 5 <210> 9
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 <220>

<223> Primer
 <221> característica nueva
 <222> 3, 8, 11, 18, 21
 <223> n = A, T, C o G
 5 <400> 9 23
 gncaganga ncactcngg ngg 23
 <210> 10
 <211> 25
 <212> ADN
 10 <213> Secuencia artificial
 <220>
 <223> Iniciador
 <221> característica nueva
 <222> 3, 7, 14
 15 <223> n = A, T, C o G
 <400> 10 25
 atntgtngga gaanagagat ggatg

REIVINDICACIONES

1. Una molécula aislada de ácido nucleico seleccionada del grupo que consiste de:
 - (a) una molécula aislada de ácido nucleico que comprende la secuencia de nucleótidos expuesta en la SEQ ID NO: 5, o un complemento de la misma;
 - (b) una molécula aislada de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos expuesta en la SEQ ID NO: 6, o un complemento de la misma; y
 - (c) una molécula aislada de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que es al menos 98% idéntica a la secuencia completa de aminoácidos de la SEQ ID NO: 6, o un complemento de la misma, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa.
2. La molécula de ácido nucleico de la reivindicación 1, en donde la molécula de desaturasa codificada por dicha molécula de ácido nucleico comprende al menos un motivo representado por los residuos aminoácidos seleccionados del grupo que consiste de los residuos 40 - 45 de la SEQ ID NO: 6, los residuos 171 - 176 de la SEQ ID NO: 6, los residuos 208 - 213 de la SEQ ID NO: 6, y los residuos 395 - 400 de la SEQ ID NO: 6.
3. La molécula de ácido nucleico de la reivindicación 1, en donde la molécula de desaturasa codificada por dicha molécula de ácido nucleico comprende al menos un motivo de enlazamiento de hemo o un motivo de histidina que tiene al menos dos residuos aminoácidos de histidina.
4. La molécula de ácido nucleico de la reivindicación 1, en donde la molécula de ácido nucleico codifica una molécula de ácido graso $\Delta 5$ -2 desaturasa capaz de convertir 20:4(n-3) a ácido eicosapentaenóico (EPA) (20:5(n-3)) o 20:3(n-6) a ácido araquidónico (AA) (20:4(n-6)).
5. Un vector que comprende la molécula de ácido nucleico de la reivindicación 1.
6. El vector de la reivindicación 5, en donde el vector es un vector de expresión.
7. Una célula huésped que comprende al vector de expresión de la reivindicación 6.
8. La célula huésped de la reivindicación 7, en donde la célula se selecciona del grupo que consiste de una célula vegetal, una célula microbiana, y una célula animal.
9. La célula huésped de la reivindicación 8, en donde la célula vegetal es una célula obtenida de un cultivo de semillas oleaginosas seleccionado del grupo que consiste de lino (*Linum sp.*), *Brassica napus*, *Brassica juncea*, colza (*Brassica sp.*), soja (*Glycine* y *Soja sp.*), girasol (*Helianthus sp.*), algodón (*Gossypium sp.*), maíz (*Zea mays*), olivo (*Olea sp.*), cártamo (*Carthamus sp.*), cacao (*Theobroma cacao*), y cacahuete (*Arachis sp.*).
10. La célula huésped de la reivindicación 8, en donde la célula microbiana se selecciona del grupo que consiste de *Thraustochytrium*, *Pythium*, *Schizochytrium*, y *Crythecodium*.
11. Un polipéptido aislado seleccionado del grupo que consiste de:
 - a) un polipéptido aislado que comprende la secuencia de aminoácidos de la SEQ ID NO: 6;
 - b) un polipéptido aislado codificado por una molécula de ácido nucleico que consiste de la secuencia de nucleótidos de la SEQ ID NO: 5, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa; y
 - c) un polipéptido aislado que comprende una secuencia de aminoácidos que es al menos 98% idéntica a la secuencia entera de aminoácidos de la SEQ ID NO: 6, en donde el polipéptido tiene una actividad de $\Delta 5$ desaturasa;
12. Un método para producir el polipéptido de acuerdo con la reivindicación 11 que comprende el cultivo de la célula huésped de la reivindicación 7 para producir el polipéptido.
13. El polipéptido de la reivindicación 11, en donde el polipéptido comprende al menos un motivo representado por los residuos aminoácidos seleccionados del grupo que consiste de los residuos 40 - 45 de la SEQ ID NO: 6, los residuos 171 - 176 de la SEQ ID NO: 6, los residuos 208 - 213 de la SEQ ID NO: 6, y los residuos 395 - 400 de la SEQ ID NO: 6.
14. El polipéptido de la reivindicación 11, en donde el polipéptido comprende al menos uno de un motivo de enlazamiento de hemo o un motivo de histidina que tiene al menos dos residuos aminoácidos de histidina.
15. El polipéptido de la reivindicación 11, en donde el polipéptido es una molécula de ácido graso $\Delta 5$ -2 desaturasa capaz de convertir 20:4(n-3) al ácido eicosapentaenóico (EPA) (20:5(n-3)) o 20:3(n-6) al ácido araquidónico (AA) (20:4(n-6)).

16. Un método para producir un ácido graso insaturado que comprende el cultivo de la célula de la reivindicación 7 bajo condiciones tales que se produzca el ácido graso insaturado.
- 5 17. Un método para producir un ácido graso insaturado que comprende transfectar o transformar una célula con la molécula aislada de ácido nucleico de la reivindicación 1 y cultivar la célula bajo condiciones de tal manera que se produzca el ácido graso insaturado.
18. Un método para producir un ácido graso insaturado que comprende poner en contacto una composición que contiene al menos un sustrato de desaturasa con al menos un polipéptido aislado de la reivindicación 11 bajo condiciones de tal manera que se produzca el ácido graso insaturado.
- 10 19. Un método para mejorar la producción de un ácido graso insaturado que comprende cultivar la célula de la reivindicación 7, de tal manera que se produzca la modulación de la producción de un ácido graso insaturado.
20. Un método para producir una célula capaz de generar un ácido graso insaturado que comprende introducir en dicha célula la molécula de ácido nucleico de la reivindicación 1, en donde la molécula de ácido nucleico codifica una desaturasa que tiene una actividad de catalizar la formación de un doble enlace en una cadena de acilo graso.
- 15 21. El método de cualquiera de las reivindicaciones 16 - 20, en donde el ácido graso insaturado se selecciona del grupo que consiste del ácido gamma linoleico (GLA) (18:3(6, 9, 12)), ácido estearidónico (SDA) (18:4(6, 9, 12, 15)), ácido araquidónico (AA) (20:4(5, 8, 11, 14)) y ácido eicosapentaenóico (EPA) (20:5(5, 8, 11, 14, 17)).
22. El método de cualquiera de las reivindicaciones 16 - 20, que comprende además recuperar el ácido graso insaturado.
- 20 23. El método de cualquiera de las reivindicaciones 16, 17 ó 19 - 20, en donde la célula se selecciona del grupo que consiste de una célula vegetal, una célula animal y una célula microbiana.
24. El método de la reivindicación 23, en donde la célula vegetal es de una planta de semillas oleaginosas.
- 25 25. El método de la reivindicación 24, en donde la planta de semillas oleaginosas se selecciona del grupo que consiste de linaza (*Linum sp.*), colza (*Brassica sp.*), soja (*Glycine* y *Soja sp.*), girasol (*Helianthus sp.*), algodón (*Gossypium sp.*), maíz (*Zea mays*), olivo (*Olea sp.*), cártamo (*Carthamus sp.*), cacao (*Theobroma cacao*), y cacahuete (*Arachis sp.*).
26. El método de la reivindicación 19, en donde se mejora la producción del ácido graso insaturado.
27. A composición que comprende al polipéptido de la reivindicación 11.
28. Una composición que comprende la célula producida por medio del método de la reivindicación 20.
29. La composición de la reivindicación 27 ó 28 para uso como un suplemento dietético.
- 30 30. La composición de la reivindicación 27 ó 28 para uso como alimento para animales.
31. La composición de la reivindicación 27 para uso en el tratamiento de un paciente que tiene un trastorno, en donde el trastorno se selecciona opcionalmente del grupo que consiste de estrés, diabetes, cáncer, trastornos inflamatorios, y trastornos cardiovasculares.

ATGACCGAGAAGGCGAGTGACGAGTTCACGTGGCAGGAGGTGCGCCAAGCACAACACGGCCA
 AGAGCGCGTGGGTGATCATCCGCGGCGAGGTGTACGACGTGACCGAGTGGGCGGACAAGCAC
 CCGGGCGGCAGCGAGCTCATCGTCCTGCACTCCGGTCGTGAATGCACGGACACGTTCTACTCG
 TACCACCCGTTCTCGAACC GCGCCGACAAGATCTTGGCCAAGTACAAGATCGGCAAGCTCGTG
 GCGGCTACGAGTTCCCGGTGTTCAAGCCGGA CTGCGGCTTCTACAAGGAATGCTCGGAGCGC
 GTGGCCGAGTACTTTAAGACGAACAATCTGGACCCAAAGGCGGCGTTCCGCGGTCTCTGGCG
 CATGGTGTTCGTGTTTCGCGGTGCGCCGCGCTCGCGTACATGGGCATGAATGAGCTCATCCCTGG
 AAACGTGTACGCGCAGTACGCGTGGGGCGTGGTGTTCGGTGTCTTCCAGGCGCTGCCATTGCT
 GCACGTGATGCACGACTCGTCGCACGCGGCATGCTCGAGCAGCCAGCGATGTGGCAGATCA
 TCGGTGCTGGTGTGATGGACTGGTTCGCTGGCGCCAGCATGGTGTGCTGGTGAACCAGCACG
 TTGTGGGCCACCATCTACACGAACGTGCGGGCGCGGACCCGGATCTCCCGGTCGACTTTG
 AGAGCGACGTGCGCCGCATCGTGCACCGCCAGGTGCTGCTGCCGATCTACAAGTTCAGCACA
 TCTACCTGCCACCGCTCTACGGCGTGCTGGGCCCTCAAGTTCGCGATCCAGGACGTGTTGAGA
 CGTTCGTGTGCTCACGAACGGCCCCGGTGCGTGTGAACCCGCACCCGGTGTCGGACTGGGTGC
 AAATGATCTTCGCCAAGGCGTTCTGGACGTTCTACCGCATCTACATCCCGTTGGCGTGGCTCA
 AGATCACGCCGTCGACGTTCTGGGGCGTGTTCCTTCGCGAGTTCACCACAGGTGGTACCT
 CGCGTTCAACTTCCAGGTGAGCCACGTCTCGACCGAGTGCGAGTACCCGTGCGGTGATGCGCC
 GTCGCGCGAGGTGCGTGACGAGTGGGCGATCTCGCAGGTCAAGTCGTGCGGTGGACTACGCGC
 ACGGCTCGCCGCTCGCGGCGTTCTCTGCGGCGCGCTCAACTACCAGGTGACCCACCACTTGT
 ACCCGGGCATCTACAGTACCCTACCGTGCATCGCGCCGATCATCATCGACGTGTGCAAGA
 AGTACAACATCAAGTACACGGTGCTGCCGACGTTACCGAGGCGCTGCTCGCGCACTTCAAGC
 ACCTGAAGAACATGGGCGAGCTCGGCAAGCCCGTGGAGATCCACATGGGTAA

Fig.1 A

MTEKASDEFTWQEVAKHNTAKSAWVIIRGEVYDVTEWADKHPGGSELIVLHSGRECTDTFYSYH
 PFSNRADKILAKYKIGKLVGGYEFPPVFKPD SGFYKECSERVAEYFKTNNLDPKAAFAGLWRMVFV
 FAVAALAYMGMNELIPGNVYAQYAWGVVFGVFQALPLLHVMHDSSHAACSSSPAMWQIIGRGV
 MDWFAGASMVSWLNQHVVGHHIYTNVAGADPDLPVDFESDVRRIVHRQVLLPIYKFQHIYLPPL
 YGVLGLKFRIQDV FETVSLTNGPVRVNPHVSDWVQMI FAKAFWTFYRIYIPLAWLKITPSTFWG
 VFFLAEFTTGWYLA FNFOVSHVSTECEYPCGDAPSAEVD EWAISQVKSSVDYAHGSPLAAFLCG
 ALNYQVTHHLYPGISQYHYP AIPIIIDVCKKYNIKYTVLP TPTTEALLAHFKHLKNMGELGKPVEIH
 MG

Fig.1 B

ATGGTGGACCTCAAGCCTGGAGTGAAGCGCCTGGTGAGCTGGAAGGAGATCCGCGAGCACGC
 GACGCCCCGCGACCGCGTGGATCGTGATTCACCACAAGGTCTACGACATCTCCAAGTGGGACTC
 GCACCCGGGTGGCTCCGTGATGCTCACGCAGGCCGGCGAGGACGCCACGGACGCCTTCGCGG
 TCTTCCACCCGTCCTCGGCGCTCAAGCTGCTCGAGCAGTTCTACGTGCGCGACGTGGACGAAA
 CCTCCAAGGCCGAGATCGAGGGGGAGCCGGCGAGCGACGAGGAGCGCGCGCGCCGCGAGCG
 CATCAACGAGTTCATCGCGTCTTACCGTCGTCTGCGCGTCAAGGTCAAGGGCATGGGGCTCTA
 CGACGCCAGCGCGCTCTACTACGCGTGGAAGCTCGTGAGCACGTTCGGCATCGCGGTGCTCTC
 GATGGCGATCTGCTTCTTCTTCAACAGTTTCGCCATGTACATGGTCGCCGGCGTGATTATGGGG
 CTCTTCTACCAGCAGTCCGGATGGCTGGCGCACGACTTCTTGACACAACCAGGTGTGCGAGAAC
 CGCACGCTCGGCAACCTTATCGGCTGCCCTCGTGGGCAACGCCCTGGCAGGGCTTCAGCGTGCAG
 TGGTGGAAGAACAAGCACAACTGCACCGCGGTGCCGAACCTGCACAGCGCCAAGGACGA
 GGGCTTCATCGGCGACCCGGACATCGACACCATGCCGCTGCTGGCGTGGTCTAAGGAGATGG
 CGCGCAAGGCGTTCGAGTCGGCGCACGGCCCGTTCTTCATCCGCAACCAGGCGTTCCTATACT
 TCCCGCTGCTGCTGCTCGCGCGCCTGAGCTGGCTCGCGCAGTCGTTCTTCTACGTGTTACCGA
 GTTCTCGTTTCGGCATCTTCGACAAGGTCGAGTTCGACGGACCGGAGAAGGCGGGTCTGATCGT
 GCACTACATCTGGCAGCTCGCGATCCCGTACTTCTGCAACATGAGCCTGTTTGAGGGCGTGGC
 ATACTTCCTCATGGGCCAGGCGTCTGCGGCTTGCTCCCTGGCGCTGGTGTTCAGTATTGGCCAC
 AACGGCATGTCCGGTGTACGAGCGCGAAACCAAGCCGGACTTCTGGCAGCTGCAGGTGACCAC
 GACGCGCAACATCCGCGCGTCCGGTATTCATGGACTGGTTCACCGGTGGCTTGAACCTACCAGAT
 CGACCATCACCTGTTCCCGCTCGTGCCGCGCCACAACCTTGCCAAAGGTCAACGTGCTCATCAA
 GTCGCTATGCAAGGAGTTCGACATCCCGTTCACGAGACCGGCTTCTGGGAGGGCATCTACGA
 GGTGCTGGACCACCTGGCGGACATCAGCAAGGAATTCATCACCAGATTCCAGCGATGTAA

Fig.2 A

MVDLKPGVKRLVSWKEIREHATPATAWIVIHKKVYDISKWDSHPGGSVMLTQAGEDATDAFAVF
 HPSSALKLLEQFYVGDVDETSKAEIEGEPASDEERARRERINEFIASYRRLRVKVKMGMLYDASAL
 YYAWKLVSFTGIAVLSMAICFFFNFSFAMVMVAGVIMGLFYQQSGWLAHDFLHNQVCENRTLGNL
 IGCLVGNWQGFVSQWWKNKHNLHHA VPNLHSAKDEGFIGDPDIDTMPLLAWSKEMARKAFES
 ABGPFFIRNQAFLYFPLLLLARLSWLAQSFFYVFTEFSFGIFDKVEFDGPEKAGLIVHYIWQLAIPYF
 CNMSLFEGVAYFLMGQASCGLLLALVFSIGHNGMSVYERETKPDFWQLQVTTTRNIRASVFMWDW
 FTGGLNYQIDHHLFPLVPRHNLPKVNVLIKSLCKEFDIPFHETGFWEGIYEVVDHLADISKEFITEFP
 AM

Fig.2 B

Fad5-2	MTEKASDE---FTWQEVAKHNTAKSAWVIIRGEVYDVTEWADKHPGGSEL	47
Fad6	MVDLKPGVKRLVSWKEIREHATPATAWIVHHKVYDISKW-DSHPGGSSVM	49
Fad5-2	IVLHSGRECTDTFYSYHPFSNRADKILAKYKIGKLVGGYEFPPV-KPDSG	96
Fad6	LT-QAGEDATDAFAVFHP--SSALKLLEQFYVGDVDETSKAEIEGEPASD	96
Fad5-2	FYKECSERVAEYFKT-NNLDPKAAFAGL-----WRMVVFVAVAALA	136
Fad6	EERARRERINEFIASRRRLRVKVKMGGLYDASALYYAWKLVSTFGIAVLS	146
Fad5-2	YMGMNELIPGNVYAQY-AWGVVFGVFQALPLLHVMHDSSHAACSSSPAMW	185
Fad6	-MAICFFF--NSFAMYMVAGVIMGLFYQQSGW-LAHDFLHNQVCENRTL	192
Fad5-2	QIIIGRGVMDWFAGASMVSWLNQHVVGHHTYTNVAGADPDLPVDFESDVRR	235
Fad6	NLIGCLVGNAWQGFQVQWKNKHNLHHA VPNLHSAKDEGFIGDPDIDTMP	242
Fad5-2	IVHRQVLL-----PIYKFQHIYLP-PLYGVLGLKFRIQDVFETEV	274
Fad6	LLAWSKEMARKAFESAHPFFIRNQAFLYFPLLLARLSWLAQSFFYVFT	292
Fad5-2	SLTNGPVRVNPVSDWVQMFKAFAWTFYRIYIPLAWLKITPSTFWGVF	324
Fad6	EFSFGIFDKVEFDGPEKAGLI-VHYIWQLAIPYFCNMSL-FEGVAY---F	337
Fad5-2	FLAEFTTGWYLAFNFQVSHVSTECEYPCGDAPSAEVGDEWAIQVKSSVD	374
Fad6	LMGQASCGLLLALVFSIGH-NGMSVYERETKP-----DFWQL-QVTTTRN	380
Fad5-2	YAHGSPLAAFLCGALNYQVTHHLYPGISQYHYPATAPIIIDVCKKYNIKY	424
Fad6	I-RASVFMDWFTGGLNYQIDHHLFPILVPRHNLPKVNVLIKSLCKEFDIPF	429
Fad5-2	TVLPTFTEALLAHFKHLKNMGELGKP--VEIHMG	456
Fad6	HE-TGFGEI---YEVVDHLADISKEFITEFPAM	459

Fig.3

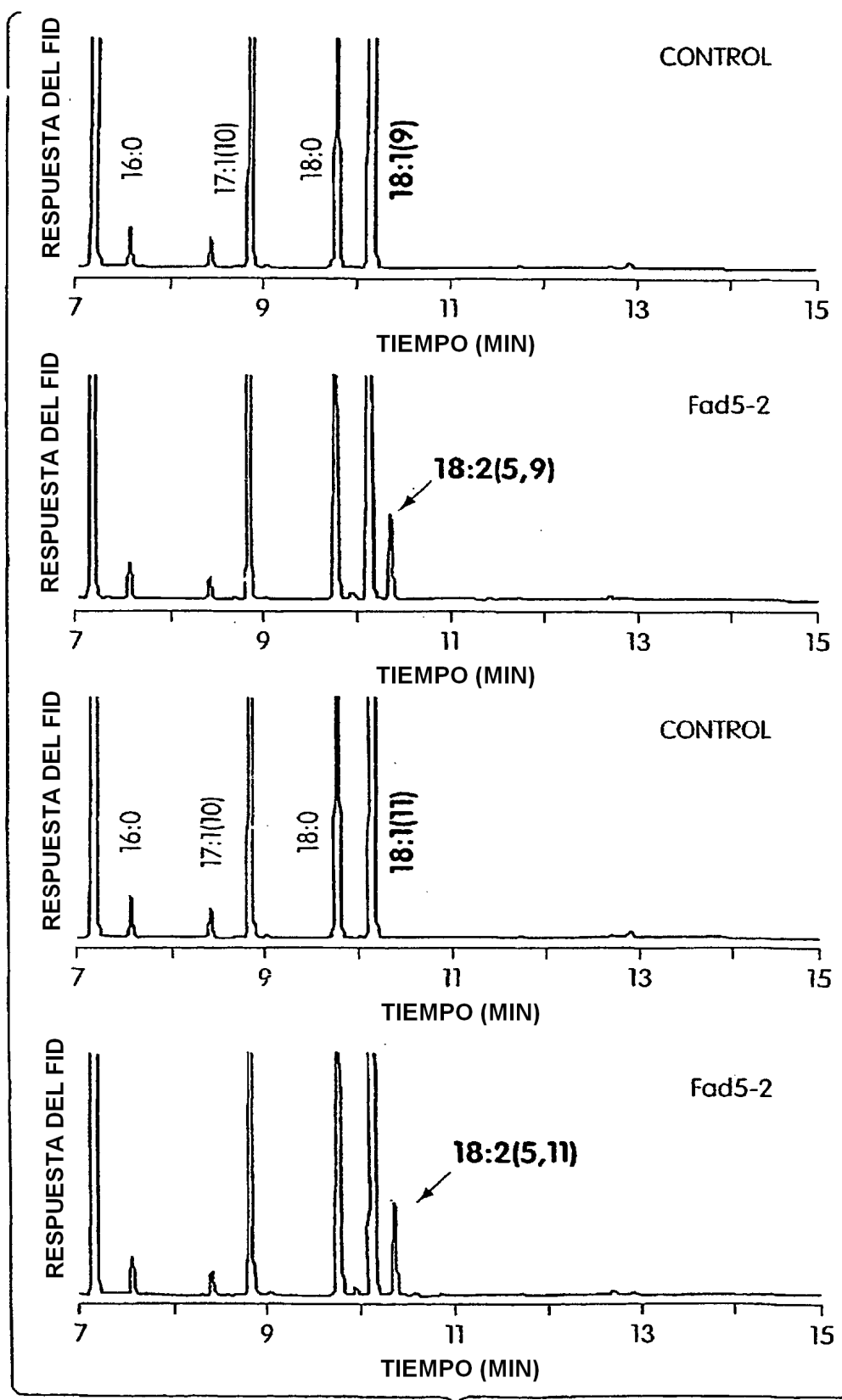


Fig.4

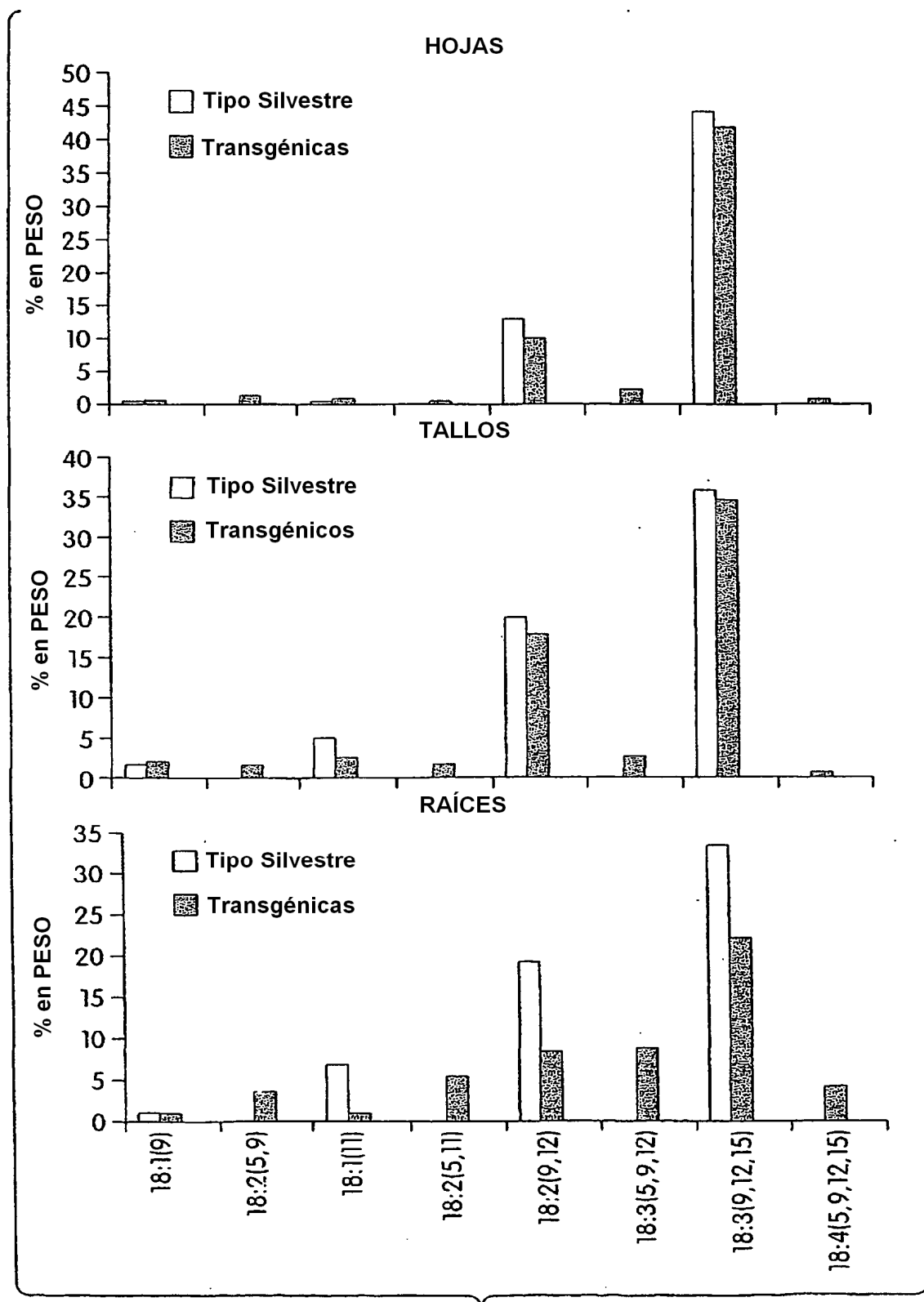


Fig.5

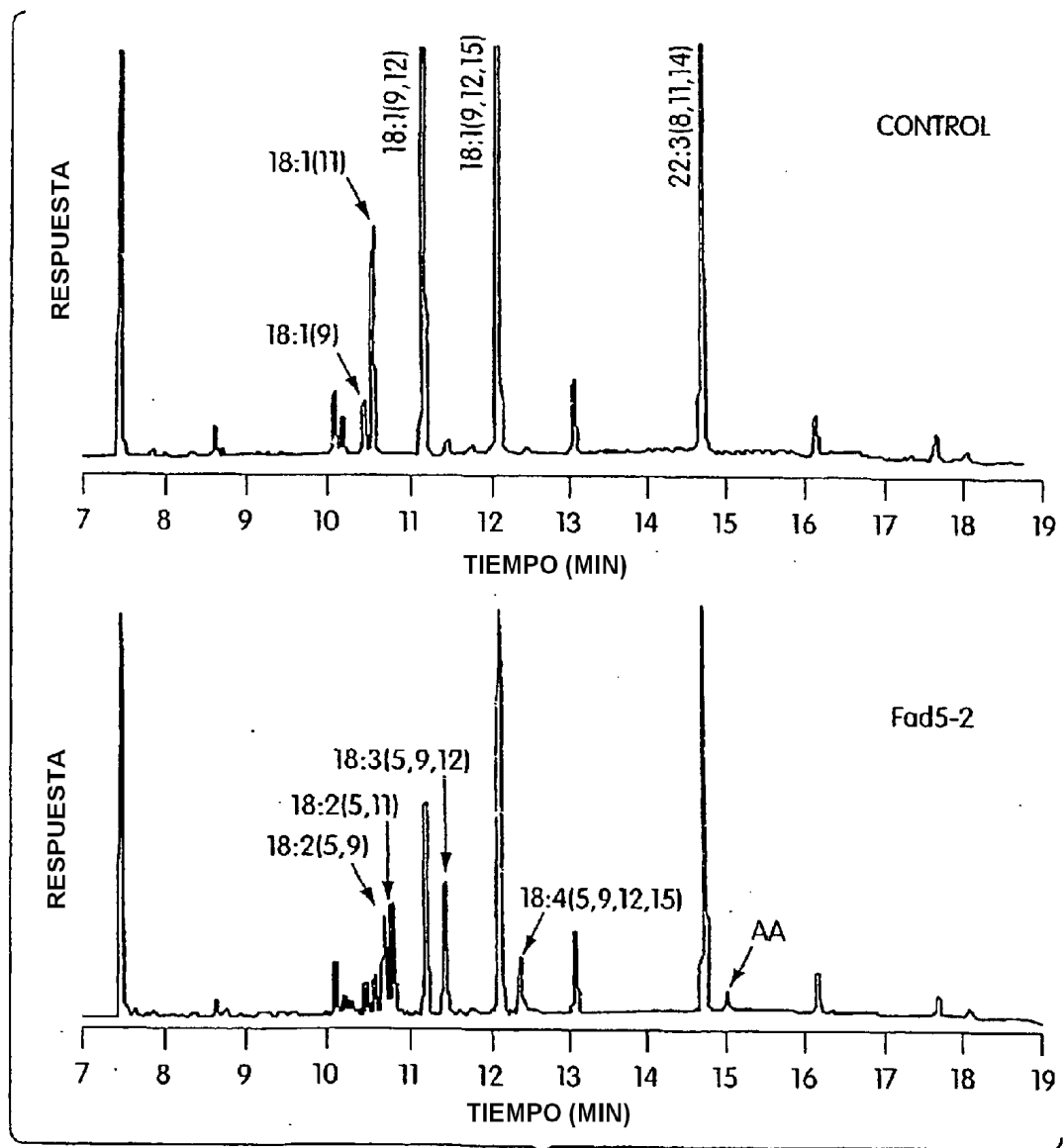


Fig.6

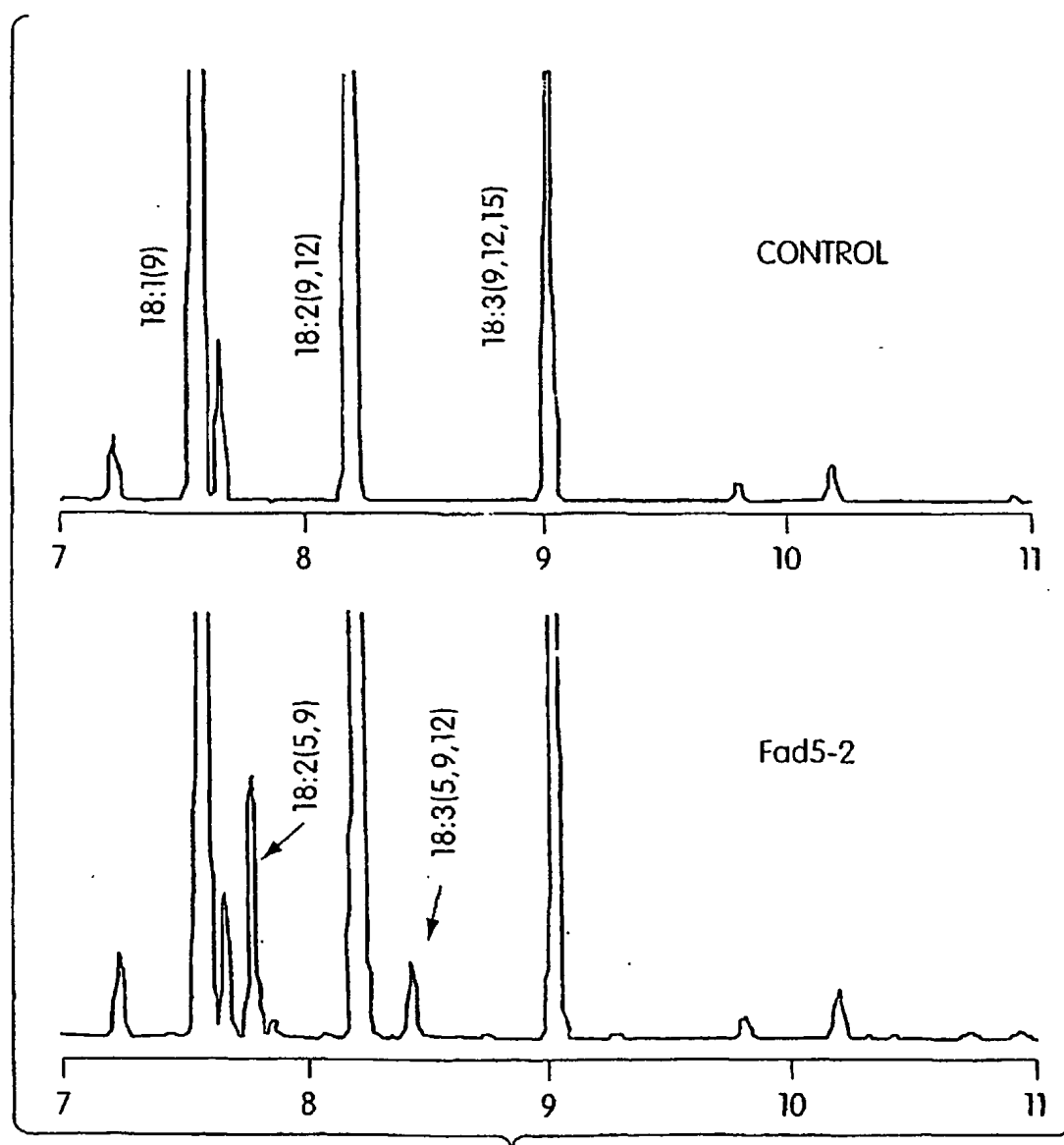


Fig.7

Tabla 2. Perfiles de ácido graso de *Pythium irregulare* (% en peso)

16:0	16:1	18:0	18:1	18:2 (LA)	18:3 (GLA)	18:3 (ALA)	20:4 (AA)	20:5 (EPA)
16.2	5.5	0.9	8.9	15.3	<0.1	<0.1	8.9	20.3

Fig.8

Tabla 3. Conversión de ácidos grasos exógenos en levadura AMY-2apfad5-2.

Sustrato	Acumulación del sustrato (% en peso)	Producto	Acumulación del producto (% en peso)
16:1(9Z)	24.41	16:2(5Z,9Z)	0.47
18:1(9Z)	11.88	18:2(5Z,9Z)	1.26
18:1(11Z)	6.15	18:2(5Z,11Z)	0.85
18:1(11E)	4.57	18:2(5Z,11E)	0.08
18:1(12E)	4.77	18:2(5Z,12E)	0.09
18:1(15Z)	4.55	18:2(5Z,15Z)	0
18:2(9Z,12Z)	3.59	18:3(5Z,9Z,12Z)	0.09
18:3(9Z,12Z,15Z)	30.71	18:4(5Z,9Z,12Z,15Z)	0.07
20:2(11Z,14Z)	8.59	20:3(5Z,11Z,14Z)	2.08
20:3(8Z,11Z,14Z)	14.54	20:4(5Z,8Z,11Z,14Z)	3.68
20:3(11Z,14Z,17Z)	2.29	20:4(5Z,11Z,14Z,17Z)	1.08

Fig.9

Tabla 4. Acumulación de los Δ^5 -UPIFA en semillas de linaza transgénica que expresan Fad5-2 bajo el control de los promotores específicos de napina (Napina) y de semilla de linaza (Cnl). Los niveles de ácido graso se muestran como el porcentaje en peso de los ácidos grasos totales

Ácidos grasos	18:1 (9)	18:2 (5,9)	18:2 9,12)	18:3 (5,9,12)	18:3 (9,12,15)	18:4 (5,9,12,15)
Tipo Silvestre	17.4	0	14.9	0	56.6	0
Cnl-5-2	19.9	6.8	7.1	0.5	55.6	0.7
Cnl-10-1	29.1	9.3	5.1	0.4	47.6	0.4
Cnl-15-3	23.9	7.1	7.1	0.4	51.6	0.5
Napina -5-4	41.0	0.4	7.5	0	42.1	0
Napina -6-1	46.7	0.3	5.3	0	39.1	0
Napina -12-3	46.3	0.6	5.3	0	39.3	0

Fig.10

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Este listado de referencias citado por el solicitante es únicamente para conveniencia del lector. No forma parte del documento europeo de la patente. Aunque se ha tenido gran cuidado en la recopilación, no se pueden excluir los errores o las omisiones y la OEP rechaza toda responsabilidad en este sentido.

5 Documentos de patente citados en la descripción

- WO 9846765 A [0009] • US 5721137 A, Frascotti [0084]
- WO 9846764 A [0009] • US 4522811 A [0122]
- WO 0020602 A [0009] • WO IB0102346 A [0154]
- WO 0020603 A [0009] • US 60236303 B [0154]

10 • WO 9933958 A [0009] • US 60257562 B [0154]**Literatura citada en la descripción que no es de patente:**

- Crawford, M.A. et al. Am. J. Clin. Nutr., 1997, vol. 66, 1032S-1041S [0003]
- Giusto, N.M. et al. Prog. Lipid Res., 2000, vol. 39, 315-391 [0003] [0034]
- Martinetz, M. J. Pediatr., 1992, vol. 120, S129-S138 [0003]
- 15 • Horrocks, L.A. ; Yeo, Y.K. Pharmacol. Res., 1999, vol. 40, 211-215 [0003] [0034]
- Spector, A.A. Lipids, 1999, vol. 34, S1-S3 [0003]
- Horrobin, D.F. Prog. Lipid Res, 1992, vol. 31, 163-194 [0004]
- Sprecher, H. et al. J. Lipid Res., 1995, vol. 36, 2471-2477 [0004]
- Sprecher, H. et al. Lipids, 1999, vol. 34, S153-S156 [0004]
- 20 • Qiu X. et al. J.Biol.Chem., 2001, vol. 276 (34), 31561-31566 [0009]
- Martinetz, M. J. Pediatr., 1992, vol. 120, S129-S1 38 [0034]
- Sambrook, J. et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0047]
- Current Protocols in Molecular Biology. John Wiley & Sons, Inc, 1995 [0066]
- Sambrook et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, 1989 [0066]
- 25 • Church; Gilbert. Proc. Natl. Acad. Sci. USA, 1984, vol. 81, 1991-1995 [0066]
- Needleman; Wunsch. J. Mol. Biol, 1970, 444-453 [0080]
- Meyers; Miller. Comput. Appl. Biosci., 1988, vol. 4, 11-17 [0081]
- Altschul et al. J. Mol. Biol., 1990, vol. 215, 403-10 [0082]
- Altschul et al. Nucleic Acids Res., 1997, vol. 25 (17), 3389-3402 [0082]
- 30 • Sambrook, J.; Fritsh, E. F. ; Maniatis, T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0086]
- Weete, J.D. et al. Lipids, 1997, vol. 32, 839-845 [0130]
- Qiu, X.; Erickson, L. Plant Mol.Biol. Repr., 1994, vol. 12, 209-214 [0131]
- Sambrook, J; Fritseh, E.F ; Maniatis, T. Molecular cloning - A laboratory manual. Cold Spring Harbor, 1989 [0132]
- 35 • Gietz, D. et al. Nucleic Acids Res., 1992, vol. 20, 1425 [0135]
- Covello, P.S.; Reed, D.W. Plant Physiol., 1996, vol. 111, 223-226 [0135] [0137]