



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 343 106**

51 Int. Cl.:  
**C12N 9/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **01985723 .4**

96 Fecha de presentación : **28.09.2001**

97 Número de publicación de la solicitud: **1322752**

97 Fecha de publicación de la solicitud: **02.07.2003**

54 Título: **Fad4, Fad5, Fad5-2 y Fad6, miembros de la familia de desaturasas de ácidos grasos y usos de los mismos.**

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

45 Fecha de publicación de la mención BOPI:  
**23.07.2010**

45 Fecha de la publicación del folleto de la patente:  
**23.07.2010**

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

72 Inventor/es: **Qiu, 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

Fad4, Fad5, Fad5-2 y Fad6, miembros de la familia de desaturasas de ácidos grasos y usos de los mismos.

## 5 Antecedentes de la invención

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, además, se presentan en forma esterificada como componente principal de lípidos. Los lípidos/ácidos grasos son fuentes de energía (por ejemplo,  $\beta$ -oxidación) y son parte integral de membranas celulares que son indispensables para el procesamiento de información bioquímica o biológica.

Los ácidos grasos se pueden dividir 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 se producen por medio de desaturasas terminales que pertenecen a la clase de enzimas de hierro no hemo. Cada una de estas enzimas hace parte de un 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 esas desaturasas que son requeridas para la formación de enlaces dobles particulares en ácidos grasos insaturados. Por lo tanto, los humanos tienen que ingerir algunos ácidos grasos a través de su dieta. Tales ácidos grasos esenciales son, por ejemplo, ácido linoléico (C18:2); ácido linolénico (C18:3); ácido araquidónico (C20:4). En contraste, los insectos y las plantas son capaces de sintetizar una variedad mucho mayor de ácidos grasos insaturados y sus derivados.

Los ácidos grasos poliinsaturados de cadena larga (los LCPUFA) tales como el ácido docosahexaenoico (DHA, 22:6 (4, 7, 10, 13, 16, 19)) son componentes esenciales de las membranas celulares de diferentes tejidos y organelos en mamíferos (células nerviosas, de retina, de cerebro e inmunes). Por ejemplo, más del 30% de los ácidos grasos en los fosfolípidos del cerebro son 22:6 (n-3) y 20:4 (n-6). (Crawford, M. A., y colaboradores (1997) *Am. J-Clin. Nutr.* 66:1032S-1041S). En la retina, DHA representa más del 60% de los ácidos grasos totales en el segmento externo de la barra, la parte fotosensible de la célula fotorreceptora (Giusto, N. M., y colaboradores (2000) *Prog. Lipid Res.* 39: 315-391). Estudios clínicos han mostrado que DHA es esencial para el crecimiento y desarrollo del cerebro de los 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 significativos 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, también se encontraron 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 los humanos para permanecer saludables. Es particularmente importante para los bebés, niños pequeños y adultos mayores la ingesta adecuada de estos ácidos grasos a partir de la dieta ya que no pueden ser eficientemente sintetizados en sus organismos y deben ser suministrados con la alimentación (Spector, A. A. (1999) *Lipids* 34: S1-S3).

El DHA es un ácido graso de la serie n-3 de acuerdo con la ubicación del último doble enlace en el extremo del metilo. 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 de que 20:5 (5,8,11,14,17) se alarga hasta 22:5 (7, 10, 13, 16, 19) y luego se convierte hasta 22:6 (4, 7, 10, 13, 16, 19) por medio de la desaturación final  $\Delta 4$  (Horrobin, D. F. (1992) *Prog. Lipid Res.* 31: 163-194). Sin embargo, Sprecher y colaboradores 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  $\beta$ -oxidación limitada en peroxisomas (Sprecher, H., y colaboradores (1995) *J. Lipid Res.* 36: 2471-2477; Sprecher, H., y colaboradores (1999) *Lipids* 34: S153-S156).

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 está oxidado al momento de 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 a los altos costos de extracción.

EPA y AA son ambos ácidos grasos esenciales  $\Delta 5$ . Ellos forman una clase única de constituyentes alimenticios y de pienso para humanos y animales. EPA pertenece a la serie n-3 con cinco dobles enlaces en la cadena acilo, se encuentra en el alimento de mar y es abundante en pescado azul el Atlántico Norte. AA pertenece a la serie n-6 con cuatro dobles enlaces. La carencia de un doble enlace en la posición  $\omega$ -3 confiere sobre AA propiedades diferentes a aquella encontradas en EPA. Los eicosanoides producidos a partir de AA tienen fuertes propiedades inflamatorias y de agregación plaquetaria, mientras que aquellos derivados de EPA tienen propiedades antiinflamatorias y de anti agregación plaquetaria. Los AA se pueden obtener a partir de algunos alimentos tales como carne, pescado, y huevos, pero la concentración es baja.

El ácido gama-linolénico (GLA) es otro ácido graso esencial encontrado en los mamíferos. El GLA es el intermedio metabólico para los 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 la desaturación  $\Delta 6$ . Muchas condiciones fisiológicas y patológicas tales como el envejecimiento, el estrés, la diabetes, el eczema, y algunas infecciones se ha demostrado que deprimen la etapa de desaturación  $\Delta 6$ . Además, el GLA es fácilmente catabolizado a partir de la oxidación y rápida división celular asociada con ciertos trastornos, por ejemplo, cáncer o inflamación.

5 Por lo tanto, la suplementación de la dieta con GLA puede reducir los riesgos de estos trastornos. Los estudios clínicos han mostrado 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.

10 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, estas fuentes de GLA no son ideales para suplementación de la dieta debido a grandes fluctuaciones en la disponibilidad y en los costos asociados con procesos de extracción.

## 15 Resumen de la invención

La biosíntesis de ácidos grasos es una actividad principal de las plantas y de los microorganismos. Sin embargo, los humanos tienen una capacidad limitada para sintetizar ácidos grasos esenciales, por ejemplo ácidos grasos poliinsaturados de cadena larga (los LCPUFA). La biotecnología ha sido considerada con mucho una forma eficiente para manipular el proceso de producción de ácidos grasos en plantas y microorganismos. Es rentable y renovable con pocos efectos secundarios. En consecuencia, se ha producido un tremendo esfuerzo industrial dirigido a la producción de diferentes compuestos incluyendo 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 ruta atractiva para la producción de ácidos grasos insaturados especialmente los LCPUFA, en una forma segura y rentable para acumular el máximo valor terapéutico a partir de estos ácidos grasos.

La presente invención se basa, al 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 Fad 4 ( $\Delta 4$  desaturasa) y Fad 5 ( $\Delta 5$  desaturasa), que están involucradas en la biosíntesis de ácidos grasos poliinsaturados de cadena larga DHA (ácido docosahexanoico, 22:6, n-3) y DPA (ácido docosapentanoico, 22:5, n-6); más específicamente, Fad 4 desaturada 22:5 (n-3) y 22:4 (n-6) resultando en DHA y DPA; Fad5 desaturada 20:4 (n-3) y 20:3 (n-6) resultando en EPA y AA.

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

- (a) una molécula de ácido nucleico que comprende la secuencia de nucleótidos expuesta en la SEQ ID NO: 1, o en la SEQ ID NO: 3, o un complemento de la misma;
- 40 (b) una molécula de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos expuesta en la SEQ ID NO: 2, o en la SEQ ID NO: 4, o un complemento de la misma; y

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

- (a) una molécula de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que es aproximadamente al menos 50% idéntica a la secuencia entera de aminoácidos de la SEQ ID NO: 2 ó 4, en donde el polipéptido tiene una actividad de  $\Delta 4$  (en el caso de la SEQ ID NO: 2) o  $\Delta 5$  desaturasa (en el caso de la SEQ ID NO: 4), o un complemento de las mismas;
- 50 (b) una molécula de ácido nucleico que hibrida bajo condiciones rigurosas hasta un complemento de una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la SEQ ID NO: 1 ó 3, en donde la molécula de ácido nucleico codifica un polipéptido que tiene una actividad de  $\Delta 4$  (en el caso de la SEQ ID NO: 1) o  $\Delta 5$  desaturasa (en el caso de la SEQ ID NO: 3), o un complemento de las mismas; y
- 55

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

- (a) un polipéptido que comprende una secuencia de aminoácidos que es al menos 50% idéntica a la secuencia entera de aminoácidos de la SEQ ID NO: 2 ó 4 y que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 2 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 4;
- 60 (b) un polipéptido que es codificado por una molécula de ácido nucleico que hibrida bajo condiciones rigurosas hasta un complemento de una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEQ ID NO: 1 ó 3 y que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 1 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 3; y
- 65

- (c) un polipéptido que comprende la secuencia de aminoácidos de la SEQ ID NO: 2 ó 4 que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 2 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 4.

5

En una modalidad, la invención presenta una molécula aislada de ácido nucleico que incluye la secuencia de nucleótidos expuesta en la SEQ ID NO: 1 o en la SEQ ID NO: 3. En otra modalidad, la invención presenta una molécula aislada de ácido nucleico que codifica un polipéptido que incluye la secuencia de aminoácidos expuesta en la SEQ ID NO: 2 ó 4.

10

En aún otras modalidades, la invención proporciona moléculas aisladas de ácido nucleico que incluyen secuencias de nucleótidos que son sustancialmente idénticas (por ejemplo, 70% idénticas) a la secuencia de nucleótidos expuesta como la SEQ ID NO: 1 o la SEQ ID NO: 3. Aquí 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: 1 o la SEQ ID NO: 3. En otra modalidad, la invención presenta moléculas aisladas de ácido nucleico que codifican un polipéptido que incluye una secuencia de aminoácidos que es sustancialmente idéntica (por ejemplo, 50% idéntica) a la secuencia de aminoácidos expuesta como la SEQ ID NO: 2 ó 4. 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: 2 ó 4. 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 de aminoácidos contiguos de la secuencia de aminoácidos de la SEQ ID NO: 2 ó 4). En aún otras modalidades, la invención presenta moléculas de ácido nucleico que son complementarias a, o que hibridan bajo condiciones rigurosas con las moléculas aisladas de ácido nucleico descritas aquí.

25

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

30

Como se describió anteriormente, la invención presenta polipéptidos aislados de desaturasa. Los ejemplos de modalidades presentan un polipéptido que incluye la secuencia de aminoácidos expuesta como la SEQ ID NO: 2 ó 4, un polipéptido que incluye una secuencia de aminoácidos al menos 50% idéntica a la secuencia de aminoácidos expuesta como la SEQ ID NO: 2 ó 4, un polipéptido codificado por una molécula de ácido nucleico que incluye una secuencia de nucleótidos al menos 70% idéntica a la secuencia de nucleótidos expuesta como la SEQ ID NO: 1 o la SEQ ID NO: 3. También se describen fragmentos de los polipéptidos de longitud completa descritos aquí (por ejemplo, fragmentos que incluyen al menos 10 residuos de aminoácidos contiguos de la secuencia expuesta como la SEQ ID NO: 2 ó 4) así como variantes alélicas del polipéptido que tiene la secuencia de aminoácidos expuesta como la SEQ ID NO: 2 ó 4.

40

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

45

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 tales como DHA, EPA, AA, SDA, y GLA. Los ejemplos de plantas utilizadas para expresión de las desaturasas de la presente invención incluyen, entre otras, plantas y semillas de plantas de cultivos de semillas oleaginosas, por ejemplo, lino (*Linum* sp.), colza (*Brassica* sp.), soja (*Glycine* y *Soja* sp.), girasol (*Helianthus* sp.), algodón (*Gossypium* sp.), maíz (*Zea mays*), oliva (*Olea* sp.), cártamo (*Carthamus* sp.), cacao (*Theobroma cacao*), y maní (*Arachis* sp.).

55

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 de ácidos grasos insaturados utilizando células, por ejemplo, células de plantas, células animales, y/o células microbianas en las cuales se ha manipulado la ruta biosintética de ácidos grasos insaturados de tal manera que se produzcan los LCPUFA u otros compuestos deseados de ácidos grasos insaturados.

60

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 de ácidos grasos insaturados manipulada de tal manera que se produzcan ácidos grasos insaturados (por ejemplo, producidos a un nivel mayor). Por ejemplo, la invención presenta métodos para producir un ácido graso insaturado (por ejemplo, DHA) en células que contienen al menos una molécula aislada de ácido nucleico para desaturasa, por ejemplo Fad4 y/o Fad5, 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 recuperar el LCPUFA.

65

En otra modalidad, la presente invención proporciona métodos para producir ácidos grasos insaturados, por ejemplo, los LCPUFA, por ejemplo, DHA, que comprenden poner en contacto una composición que contiene al menos una

molécula objetivo de desaturasa, como se define aquí, con al menos un polipéptido desaturasa aislado, por ejemplo, Fad4 y/o Fad5, como se describió anteriormente, bajo condiciones tales que se produzca un ácido graso insaturado, por ejemplo LCPUFA, por ejemplo DHA. Tales métodos pueden incluir además la etapa de recuperar el LCPUFA.

Los ácidos nucleico, proteínas, y vectores descritos anteriormente son particularmente útiles en las metodologías de la presente invención. En particular, la invención presenta métodos para mejorar la producción de ácido graso insaturado (por ejemplo, la producción de DHA) que incluye el cultivo de una planta, animal, y/o microorganismo recombinante que comprende un ácido nucleico para desaturasa, por ejemplo, Fad4 y Fad5, bajo condiciones tales que se mejore la producción de ácido graso.

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

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

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

En otra modalidad, se utilizan las composiciones de la presente invención como suplemento para la dieta, por ejemplo, en pienso para animales o como un neutracéutico. Las composiciones de la presente invención pueden permitir un método de tratamiento de un paciente que tenga un trastorno. 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 las reivindicaciones.

### Breve descripción de los dibujos

La Figura 1 muestra la secuencia de ADN y de proteína de Fad4 de *Thraustochytrium* sp.; (A) la secuencia de ADNc del marco de lectura abierto (SEQ ID NO: 1); y (B) la secuencia traducida de proteína (SEQ ID NO: 2).

La Figura 2 muestra la secuencia de AND y de proteína de Fad5 de *Thraustochytrium* sp.; (A) la secuencia de ADNc del marco de lectura abierto (SEQ ID NO: 3); y (B) la secuencia traducida de proteína (SEQ ID NO: 4).

La Figura 3 muestra una comparación de las secuencias de proteína de Fad4 y Fad5 de *Thraustochytrium* sp. (SEQ ID NO: 2 y 4, respectivamente). La barra vertical indica la identidad del aminoácido. Los motivos conservados tales como los motivos ricos en bistidina 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 cromatografía de gases (GC) de metil ésteres de ácido graso (los FAME) de la cepa de levadura Invsc2 que expresa Fad4 con sustrato exógeno 22:5 (n-3).

La Figura 5 es un análisis por cromatografía de gases/espectroscopía de masas (MS) de los FAME del pico nuevo en la Figura 4; (A) el producto Fad4; (B) el DHA (22:6, n-3) estándar.

La Figura 6 es un análisis por GC de los FAME de la cepa de levadura Invsc2 que expresa Fad4 con sustrato exógeno 22:4 (n-6).

La Figura 7 es un análisis por GC/MS de los FAME del pico nuevo en la Figura 9; (A) el producto Fad4; (B) el DPA (22:5, n-6) estándar.

La Figura 8 es un análisis por GC de los FAME de la cepa de levadura Invsc2 que expresa Fad5 con sustrato exógeno 20:3 (n-6).

La Figura 9 es un análisis por GC/MS de los FAME del pico nuevo en la Figura 11; (A) el producto Fad5; (B) el AA (20: 4-5, 8, 11, 14) estándar.

La Figura 10 es un análisis por GC de los FAME de hojas de *Brassica juncea* que expresa Fad4 bajo el control del promotor 35S con sustrato suministrado en forma endógena 22:5 (n-3).

La Figura 11 es una tabla que muestra el perfil de ácidos grasos de *Thraustochytrium* sp.

## Descripción detallada de la invención

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 “desaturasa” y moléculas de proteína (por ejemplo, Fad4 y Fad5). 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, Fad4 o Fad5. 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 docosahexaénico (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 hemo del citocromo b5. Como se utiliza aquí, el término “motivo de enlazamiento hemo” es una extensión N-terminal del dominio del tipo del citocromo b5 encontrado en desaturasas del front-end.

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.

- 5 Los ejemplos de motivos de enlazamiento hemo del citocromo b5 y los motivos de histidina incluyen los residuos aminoácidos 41-44, 182-186, 216-223, y 453-462 de la SEQ ID NO: 2, y los residuos aminoácidos 40- 43, 171-175, 207-213, y 375-384 de la SEQ ID NO: 4 como se muestra en la Figura 3.

- 10 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: 2 ó 3. Las proteínas de la invención son aquellas definidas en la reivindicación 10. 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.

- 25 En una modalidad preferida, una proteína desaturasa incluye al menos uno o más de los siguientes dominios o motivos: un motivo de enlazamiento hemo y/o un motivo de histidina y tiene una secuencia de aminoácidos de aproximadamente al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más homóloga o idéntica a la secuencia de aminoácidos de la SEQ ID NO: 2 ó 4. En aún otra modalidad preferida, una proteína desaturasa incluye al menos uno o más de los siguientes dominios: un motivo de enlazamiento 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: 1 ó 3. En otra modalidad preferida, una proteína desaturasa incluye al menos un motivo de enlazamiento hemo y/o al menos aproximadamente tres motivos de histidina, y tiene una actividad de desaturasa.

- 35 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.

- 45 La secuencia de nucleótidos de la  $\Delta 4$  desaturasa aislada de *Thraustochytrium sp.*, Fad4, 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 Fad4 de *Thraustochytrium sp.* (el marco de lectura abierto), que tiene una longitud aproximada de 1560 nucleótidos, codifica una proteína que tiene un peso molecular de aproximadamente 59,1 kD y que tiene aproximadamente 519 residuos aminoácidos de longitud.

- 55 La secuencia de nucleótidos de la  $\Delta 5$  desaturasa de *Thraustochytrium sp.*, Fad5, el ADNc y la secuencia predicha de aminoácidos codificada por el ADNc de Fad5 son mostrados en la Figura 2 y en las SEQ ID NOS: 3 y 4, respectivamente. El gen para Fad5 de *Thraustochytrium sp.*, que tiene una longitud aproximada de 1320 nucleótidos, codifica una proteína que tiene un peso molecular de aproximadamente 49,8 kD y que tiene aproximadamente 439 residuos aminoácidos de longitud.

Se describen diferentes aspectos de la invención con más detalle en las siguientes subsecciones:

## 60 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 lo 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: 1 ó 3, 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: 1 ó 3, 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* 2<sup>nd</sup> 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: 1 ó 3, 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: 1 ó 3.

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: 1 ó 3.

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: 1 ó 3. Una molécula de ácido nucleico que sea complementaria a la secuencia de nucleótidos mostrada en la SEQ ID NO: 1 ó 3 es una que sea suficientemente complementaria a la secuencia de nucleótidos mostrada en la SEQ ID NO: 1 ó 3, de tal manera que puede hibridar a la secuencia de nucleótidos mostrada en la SEQ ID NO: 1 ó 3, formando así un dúplex estable.

En aún otra modalidad, una molécula aislada de ácido nucleico de la presente invención incluye una secuencia de nucleótidos que es al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más idéntica a la secuencia de nucleótidos mostrada en la SEQ ID NO: 1 ó 3.

Las moléculas de ácido nucleico que incluyen únicamente una porción de la secuencia de ácido nucleico de la SEQ ID NO: 1 ó 3, 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: 1 ó 3, de una secuencia antisentido de la SEQ ID NO: 1 ó 3 o de una variante alélica de ocurrencia natural o mutante de la SEQ ID NO: 1 ó 3.

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: 1 ó 3, 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: 1 ó 3. 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: 2 ó 4. 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 lugar), homólogos (diferente lugar), 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 codificada. 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: 2 ó 4, en donde la molécula de ácido nucleico hibrida hasta un complemento de una molécula de ácido nucleico que comprende la SEQ ID NO: 1 ó 3, por ejemplo, bajo condiciones rigurosas de hibridación.

Las variantes alélicas de la desaturasa, por ejemplo, Fad4 o Fad5 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 una sustitución conservadora de uno o más aminoácidos de la SEQ ID NO: 2 ó 4, 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, Fad4 o Fad5, 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: 2 ó 4, 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 *Thraustochytrium sp.* 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 *Thraustochytrium sp.* Pueden ser fácilmente identificados por contener una secuencia de aminoácidos que es sustancialmente homóloga a la SEQ ID NO: 2 ó 4.

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: 1 ó 3. Por ejemplo, se pueden identificar otros ADNc de desaturasa con base en la secuencia de nucleótidos de Fad4 o Fad5. 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: 1 ó 3. Por ejemplo, se puede identificar el ADNc de la desaturasa de *Schizochytrium* o *Cryptocodinium* con base en la secuencia de nucleótidos de una Fad4 o Fad5.

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 (1xSSPE es NaCl 0,15 M, NaH<sub>2</sub>PO<sub>4</sub> 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<sub>m</sub>) del híbrido, donde T<sub>m</sub> de acuerdo a las ecuaciones siguientes. Para híbridos de menos de 18 pares de bases de longitud, T<sub>m</sub> (°C) = 2(# de A + T bases) + 4(# de G + C bases). Para híbridos entre 18 y 49 pares de bases de longitud, T<sub>m</sub> (°C) = 81,5 + 16,6 (log<sub>10</sub>[Na<sup>+</sup>]) + 0,41 (% de G + C) - (600/N), donde N es el número de bases en el híbrido, y [Na<sup>+</sup>] es la concentración de iones sodio en el regulador de hibridación ([Na<sup>+</sup>] 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 NaH<sub>2</sub>PO<sub>4</sub> 0,25-0,5 M, 7% de SDS aproximadamente a 65°C, seguido por uno o más lavados con NaH<sub>2</sub>PO<sub>4</sub> 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: 1, 3, 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: 1 ó 3, 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: 1 ó 3. Un residuo aminoácido “no esencial” es un residuo que puede ser alterado a partir de la secuencia de tipo silvestre de Fad4 o Fad5 (por ejemplo, la secuencia de la SEQ ID NO: 2) 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: 2 ó 4, 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 50%, 55%, 60%, 65%, 76%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 99% o más homóloga con la SEQ ID NO: 2 ó 4, por ejemplo, con la longitud completa de la SEQ ID NO: 2 ó 4.

Se puede crear una molécula aislada de ácido nucleico que codifica una proteína desaturasa homóloga a la proteína de la SEQ ID NO: 2 ó 4, 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: 1 ó 3, 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: 1 ó 3, 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: 1 ó 3, se puede expresar la proteína codificada en forma recombinante y se puede determinar la actividad de la proteína.

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*

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.

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 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 producida en forma recombinante, está también  
 5 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.

La expresión “sustancialmente libre de precursores químicos o de otros compuestos químicos” incluye preparacio-  
 10 nes de proteína desaturasa en las cuales la proteína esta 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 apro-  
 15 ximadamente menos del 30% (en peso seco) de precursores químicos o de compuestos químicos sin desaturasa, más preferiblemente aproximadamente menos del 20% de precursores químicos o de compuestos químicos sin desaturasa, aún más preferiblemente aproximadamente menos del 10% de precursores químicos o de compuestos químicos sin desaturasa, y lo más preferible aproximadamente menos del 5% de precursores químicos o de compuestos químicos sin desaturasa. Debe entenderse que las proteínas de esta invención pueden estar también en una forma que sea di-  
 20 ferente de sus correspondientes proteínas de ocurrencia natural y/o que están aún asociadas al menos con algunos componentes celulares. Por ejemplo, la proteína puede estar asociada con una membrana celular.

Como se lo utiliza aquí, una “porción biológicamente activa” de una proteína desaturasa incluye un fragmento de una proteína de desaturasa que participa en una interacción entre una molécula desaturasa y una molécula sin desaturasa (por ejemplo, un sustrato de desaturasa tal como un ácido graso). Las porciones biológicamente activas de una proteína desaturasa incluyen péptidos que contienen secuencias de aminoácidos suficientemente homólogas con  
 25 o derivadas de las secuencias de aminoácidos de desaturasa, por ejemplo, las secuencias de aminoácidos mostradas en la SEQ ID NO: 2 ó 4 que incluyen suficientes residuos aminoácidos para exhibir al menos una actividad de una proteína desaturasa. Típicamente, porciones biológicamente activas incluyen un dominio o motivo con al menos una actividad de la proteína desaturasa; 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  
 30 sustrato de desaturasa o una molécula objetivo. Una porción biológicamente activa de una proteína desaturasa puede ser un polipéptido que tiene, por ejemplo, 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 incluye un motivo de enlazamiento hemo y/o al  
 35 menos un motivo de histidina, preferiblemente aproximadamente tres motivos de histidina. Además, se pueden preparar otras porciones biológicamente activas, en las cuales están suprimidas 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  
 40 NO: 2 ó 4. En otras modalidades, la proteína desaturasa es sustancialmente idéntica a la SEQ ID NO: 2 ó 4 y retiene la actividad funcional de la proteína de la SEQ ID NO: 2 ó 4, 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 incluye una secuencia de aminoácidos al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más idéntica a la SEQ ID NO: 2 ó 4.

En otra modalidad, la invención presenta una proteína desaturasa que es codificada por una molécula de ácido nucleico que consiste de una secuencia de nucleótidos al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%,  
 45 95%, 96%, 97%, 98%, 99% o más idéntica a una secuencia de nucleótidos de la SEQ ID NO: 1 ó 3, o un complemento de la misma. Esta invención presenta además una proteína desaturasa que es codificada por una molécula de ácido nucleico que consiste de una secuencia de nucleótidos que hibrida bajo condiciones estrictas de hibridación con un  
 50 complemento de una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEQ ID NO: 1 ó 3, o un complemento de la misma.

Para determinar en porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácido nucleico, se alinean las secuencias para lograr una comparación optima (por ejemplo, se pueden introducir huecos en uno  
 55 o ambos de una primera y una segunda secuencia de ácido nucleico o de aminoácidos para alineación optima 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 del 30%, preferiblemente al menos del 40%, más preferiblemente al menos del 50%, incluso más preferiblemente al menos del 60% e incluso más preferiblemente al menos del  
 60 70%, 80% o 90% de la longitud de la secuencia de referencia (por ejemplo, cuando se alinea una segunda secuencia con la secuencia de aminoácidos de Fad4 de la SEQ ID NO: 2 que tiene 519 residuos aminoácidos, se alinean al menos 156, preferiblemente al menos 208, más preferiblemente al menos 260, incluso más preferiblemente al menos 311, e incluso más preferiblemente al menos 363, 415 ó 467 residuos aminoácidos; cuando se alinea una segunda  
 65 secuencia con la secuencia de aminoácidos de Fad5 de la SEQ ID NO: 4 que tiene 439 residuos aminoácidos, se alinean al menos 132, preferiblemente al menos 176, más preferiblemente al menos 220, incluso más preferiblemente al menos 263, e incluso más preferiblemente al menos 307, 351, ó 395 residuos aminoácidos). Se comparan luego los residuos aminoácidos o los nucleótidos en las correspondientes posiciones de los aminoácidos o posiciones de los nucleótidos. Cuando una posición en la primera secuencia es 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, una “identidad” de aminoácido o de ácido nucleico es equivalente a “homología” de aminoácido o de 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 necesitan ser introducidos para una alineación óptima de las dos secuencias.

La comparación de las secuencias y la determinación del porcentaje de identidad entre dos secuencias se pueden lograr 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 dentro del programa GAP en el paquete de software de 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 longitud de hueco 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 GAWP en el paquete de software de GCG (disponible en <http://www.gcg.com>), utilizando una matriz CMP de NWSgapdna y un peso de hueco de 40, 50, 60, 70, u 80 y un peso de longitud de hueco 1, 2, 3, 4, 5, ó 6. Un ejemplo preferido no limitante de parámetros para ser utilizados junto con el programa GAP incluye una matriz de puntuación Blossum 62 con una penalización por hueco de 12, una penalización extendida de hueco de 4, y una penalización por hueco en marco de lectura de 5.

En otra modalidad, se determina el porcentaje de identidad entre dos secuencias de aminoácidos o nucleótidos 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 de PAM120, una penalización por longitud de hueco de 12 y una penalización de hueco de 4.

Se pueden utilizar adicionalmente las secuencias de ácido nucleico y de proteína de la presente invención como una “secuencias pregunta” para llevar a cabo una búsqueda contra bases de datos públicas, por ejemplo, para identificar otros miembros de la familia o secuencias relacionadas. Se pueden llevar a cabo tales búsquedas utilizando los programas NBLAST y XBLAST (versión 2.0) de Altschul y colaboradores (1990) J. Mol. Biol. 215: 403-10. Se pueden llevar a cabo las búsquedas de nucleótidos en BLAST con el programa NBLAST, puntaje = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico para desaturasa de la invención. Las búsquedas de proteína en BLAST se pueden llevar a cabo con el programa XBLAST, puntaje = 50, longitud de palabra = 3 para obtener una secuencia de aminoácidos homóloga a las moléculas de proteína de desaturasa de la invención. Para obtener alineaciones con huecos para propósitos de comparación se puede utilizar, Gapped BLAST como se describe en Altschul y colaboradores (1997) Nucleic Acids Res. 25(17): 3389-3402. Cuando se utilizan los programas BLAST y Gapped BLAST, se pueden utilizar los parámetros por defecto 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 proporciona métodos nuevos y mejorados para la producción de ácidos grasos insaturados, por ejemplo, los LCPUFA, tales como, DHA (ácido docosahexanoico, 22:6 (n-6)), DPA (ácido docosapentanoico, 22:5 (n-6)), AA (ácido Araquidónico, 20:4 (n-6)) y EPA (ácido eicosapentanoico, 20:5 (n-3)).

#### A. Células Recombinantes y Métodos para el Cultivo de las Células

La presente invención presenta además vectores recombinantes que incluyen secuencias de ácido nucleico que codifican los productos génicos como se describe aquí, preferiblemente los productos génicos Fad4 y Fad5. El término vector recombinante incluye un vector (por ejemplo, plásmido) que ha sido alterado o modificado por ingeniería genética para que contenga secuencias mayores, menores o diferentes 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 secuencia(s) reguladora(s)” significa que la secuencia de nucleótidos de interés está enlazada a la secuencia(s) reguladora(s) en una forma que permite la expresión (por ejemplo una expresión mejorada, incrementada, basal constitutiva, atenuada, disminuida o reprimida) de la secuencia de nucleótidos, preferiblemente la expresión de un producto génico codificado por la secuencia de nucleótidos (por ejemplo, cuando se introduce el vector recombinante en una célula). Se describen con mayor detalle aquí ejemplos de vectores, así como, por ejemplo, en Frascotti y colaboradores, Patente Estadounidense No. 5.721.137.

El término “secuencia reguladora” 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, se incluye una secuencia reguladora en un vector recombinante en una posición y/o orientación idéntica o similar con relación a un gen particular de interés como se observa para la secuencia reguladora y el gen de interés como aparece 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 Fad4 o Fad5) 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 Fad4 o Fad5) (por ejemplo, el promotor “nativo” de Fad4 o Fad5). Alternativamente, se puede incluir un gen de interés (por ejemplo un gen para Fad4 o Fad5) 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,

se puede incluir un gen para Fad4 o Fad5 en un vector operativamente enlazado a secuencias no reguladoras de Fad4 o Fad5. Alternativamente, se puede incluir un gen de interés (por ejemplo un gen para Fad4 o Fad5) en un vector operativamente enlazado a una secuencia reguladora de otro organismo. Por ejemplo, las secuencias reguladoras de otros microbios (por ejemplo otras secuencias reguladoras bacterianas, secuencias reguladoras de bacteriófagos y similares) pueden estar operativamente enlazadas 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, Gold 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), aquellos que dirigen la expresión inducible de una secuencia de nucleótidos en una célula (por ejemplo, promotores inducibles, por ejemplo, promotores inducibles de xilosa) y aquellos 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 la regulación de la expresión de un gen de interés por medio de la remoción o supresión de secuencias reguladoras. Por ejemplo, se pueden remover las secuencias involucradas en la regulación negativa de la transcripción de tal manera que se refuerce 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, Fad4 o Fad5) operativamente enlazado a un promotor o 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 de ARNm. Las secuencias terminadoras (o terminadoras de la transcripción en tándem) pueden servir además para estabilizar 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 que confieren resistencia a los antibióticos sobre el organismo huésped. En una modalidad, las secuencias de resistencia a los antibióticos se seleccionan del grupo que consiste de secuencias *car* (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 espectinomicina). Los vectores recombinantes de la presente invención pueden incluir además secuencias homólogas de recombinación (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 genéticamente (por ejemplo, modificada por ingeniería genética) o modificada para que la célula tenga al menos una desaturasa de ácido graso, por ejemplo, Fad4 y/o Fad5, de tal manera que se produzca un ácido graso insaturado. La modificación o manipulación por ingeniería genética de tales microorganismos puede hacerse de acuerdo con cualquier 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 altere o modifique al menos un precursor secuencia arriba o secuencia abajo, sustrato o producto de la enzima, por ejemplo, comparada con una enzima correspondiente de ocurrencia natural o de tipo silvestre.

El término “sobreexpresado(a)” o “sobreexpresión” incluye la expresión de un producto génico (por ejemplo, una desaturasa de ácido graso) en un nivel superior que la que aquel expresado antes de la manipulación de la célula o en una célula comparable que no haya sido manipulada. En una modalidad, se puede manipular genéticamente la célula (por ejemplo, modificarla por ingeniería genética) 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. La manipulación genética puede incluir, pero no se limita a, la alteración o modificación de secuencias reguladoras o de los 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 sea constitutiva), modificando la ubicación en el cromosoma de un gen particular, alterando las secuencias de ácido nucleico adyacentes a un gen particular tal como un sitio de enlazamiento de 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) involucrados en la transcripción de un gen particular y/o en la traducción de un producto génico particular, o cualquier otro medio convencional para desregular la expresión de una rutina génica particular en el arte (incluyendo, pero sin limitarse a el uso de moléculas de ácido nucleico antisentido, por ejemplo, para bloquear la expresión de proteínas represoras).

En otra modalidad se puede manipular física o ambientalmente la célula 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 haya sido

manipulada. Por ejemplo, se puede tratar una célula con, o cultivarla en presencia de un agente conocido o que se sospeche que incrementa la transcripción de un gen particular y/o la traducción de un producto génico particular de tal manera que se refuerce o incrementa la transcripción y/o la traducción. Alternativamente, se puede cultivar una célula 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 refuerce o incrementa la transcripción y/o la traducción.

El término “desregulado(a)” o “desregulación” incluye la alteración o modificación de al menos un gen en una célula que codifica una enzima en una ruta biosintética, de tal manera que se altere o modifique el nivel o la actividad de la enzima biosintética en la célula. Preferiblemente, se altera o modifica un gen que codifica una enzima en una ruta biosintética de tal manera que se refuerce 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 codifique una enzima en una ruta biosintética para que se altere o modifique el nivel o la 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 se codifique más de una enzima (por ejemplo, dos o tres enzimas biosintéticas) por parte de los genes que se presentan 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 de proteínas reguladoras que se enlazan al elemento regulador o por medio de antiterminación de la transcripción. Se puede transcribir los genes estructurales para producir un solo ARNm que codifica todas las proteínas estructurales. Debido a la regulación coordinada de los genes incluidos en un operón, la alteración o modificación del promotor único y/o del elemento coordinador puede resultar en 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), la adición de promotores fuertes, promotores inducibles o promotores múltiples o la remoción de secuencias reguladoras de tal manera que se modifique la expresión de los productos génicos, modificando la ubicación en el cromosoma del operón, alterando 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) involucrados en la transcripción del operón y/o en la traducción de los productos génicos del operón, o cualquier otro medio convencional rutinario en el arte para desregulación de la expresión de los genes (incluyendo, pero sin limitarse al uso de moléculas de ácido nucleico antisentido, por ejemplo, para bloquear la expresión de proteína represora). La desregulación puede involucrar también 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 mayor o menor actividad específica.

Una célula “recombinante” particularmente preferida de la presente invención ha sido modificada genéticamente para sobreexpresar un gen derivado de una planta o un producto génico o un gen derivado a través de un microorganismo o producto génico. El término “derivado de una planta”, “derivado a través 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 de semillas oleaginosas, o un producto génico (por ejemplo, Fad4 o Fad5) o que es codificado por un gen de una planta o un gen de un microorganismo (por ejemplo, codificado por la SEQ ID NO: 1 o la SEQ ID NO: 3).

Las metodologías de la presente invención presentan células recombinantes que sobreexpresan al menos una desaturasa de ácido graso. En una modalidad, se ha modificado genéticamente una célula recombinante de la presente invención para sobreexpresar una desaturasa de ácido graso de *Thrauschytrium sp.* (por ejemplo, ha sido modificada genéticamente para sobreexpresar al menos un  $\Delta 4$  o  $\Delta 5$  desaturasa de *Thrauschytrium sp.* (el producto génico Fad4 o Fad5) (por ejemplo, una desaturasa de ácido graso que tiene la secuencia de aminoácidos de la SEQ ID NO: 2 ó 4 o codificada por la secuencia de ácido nucleico de la SEQ ID NO: 1 ó 3).

En otra modalidad, la invención presenta una célula (por ejemplo una célula microbiana) que ha sido transformada con un vector que comprende una secuencia de ácido nucleico para la desaturasa de ácido graso (por ejemplo, una secuencia de ácido nucleico para desaturasa de ácido graso como la expuesta en la SEQ ID NO: 1 ó 3).

Otro aspecto de la presente invención presenta un método para modular la producción de ácidos grasos que comprende el cultivo de células transformadas por medio de las moléculas de ácido nucleico de la presente invención (por ejemplo, una desaturasa) para que ocurra modulación de la producción de ácido graso (por ejemplo, se refuerza la producción de ácidos grasos insaturados). El método para cultivar células transformadas por medio de las moléculas de ácido nucleico de la presente invención (por ejemplo, Fad4 y Fad5) para modular la producción de ácidos grasos denominados aquí como “biotransformación”. El proceso de biotransformación puede utilizar células recombinantes y/o las desaturasas descritas aquí. El término “proceso de biotransformación”, también denominado aquí como “proceso de bioconversión”, incluye procesos biológicos que resultan en la producción (por ejemplo, transformación o conversión) de cualquier compuesto (por ejemplo, sustrato, intermediario, o producto) que está: secuencia arriba de una desaturasa de ácido graso con respecto a un compuesto (por ejemplo, sustrato, intermediario, o producto) que está secuencia debajo de una desaturasa de ácido graso, en particular, un ácido graso insaturado. En una modalidad, la invención presenta 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 tales que se produzca un ácido graso insaturado y, opcionalmente, recuperar el ácido graso. En la presente modalidad, la invención se relaciona con un proceso de biotransformación para la producción de ácidos grasos insaturados que comprende poner en contacto una célula que sobreexpresa Fad4 o Fad5 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) 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 utilizadas en las reacciones de biotransformación están en una forma que les 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. Se pueden suspender las células (por ejemplo, en una solución apropiada tal como soluciones o medios amortiguados), lavadas (por ejemplo, lavados sin medios de cultivo de la célula), secadas con acetona, inmovilizadas (por ejemplo, con gel de poliacrilamida o k-carragenina o sobre soportes sintéticos, por ejemplo, cuentas, matrices y similares), fijadas, entrelazadas o permeabilizadas (por ejemplo, tienen membranas permeabilizadas y/o paredes de tal manera que los compuestos, por ejemplo, sustratos, intermediarios 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, de 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 “cultivar” incluye mantener y/o desarrollar 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 de la invención en medio líquido. En otra modalidad, se cultiva un microorganismo de la invención en medio sólido o en medio semisólido. En una modalidad preferida, se cultiva un microorganismo de la invención en 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 sustrato de carbono, por ejemplo, carbohidratos complejos tales como harina de habichuela o 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, habichuelas 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 del mismo; elementos traza, por ejemplo, sales de magnesio, hierro, manganeso, calcio, cobre, cinc, boro, molibdeno, y/o cobalto; así como factores de crecimiento tales como aminoácidos, vitaminas, promotores de crecimiento y similares).

Preferiblemente, se cultivan microorganismos de la presente invención bajo 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 en un pH de aproximadamente 7. En otra modalidad, se cultivan los microorganismos en un pH entre 6,0 y 8,5. Se puede mantener el pH deseado por medio de cualquier cantidad de métodos conocidos por aquellos capacitados en el arte.

Preferiblemente también, se cultivan los microorganismos de la presente invención bajo aireación controlada. El término “aireación controlada” incluye aireación suficiente (por ejemplo, oxígeno) para resultar 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, por medio de la regulación de la cantidad de oxígeno disuelto en el medio de cultivo. Preferiblemente, se controla la aireación del cultivo por medio de la agitación del mismo. Se puede suministrar agitación por medio de un equipo mecánico de agitación con hélice o similar, revolviendo o agitando en recipiente de crecimiento (por ejemplo, un fermentador) o por medio de diferentes equipos de bombeo. La aireación puede ser controlada adicionalmente 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 microorganismos de la presente invención sin formación excesiva de espuma (por ejemplo, por medio de la adición de agentes antiespumantes).

Además, se pueden cultivar plantas o microorganismos de la presente invención 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 temperaturas 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.

Se pueden cultivar microorganismos (por ejemplo, mantenerlos y/o desarrollarlos) en medio líquido y preferiblemente cultivarlos, ya sea en forma continua o intermitente, por medio de métodos convencionales de cultivo tales como cultivo en posición vertical, cultivo en tubos de ensayo, cultivo con agitación (por ejemplo, cultivo con agitación rotatoria, cultivo en matraz con agitación, etc.), cultivo en centrifugadora con aireación, 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, un proceso de fermentación). Los procesos de fermentación de la presente invención incluyen pero no se limitan a métodos de fermentación continuos y por lotes, o por lotes con alimentación adicional. La frase “proceso por lotes” o “fermentación por lotes” se refiere a un sistema cerrado en el



cual la composición del medio, los nutrientes, los aditivos para suplementación y similares se ajustan al comienzo de la fermentación y no están sometidos a alteración durante la fermentación, sin embargo, pueden 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 lotes con alimentación adicional” o fermentación “por lotes con alimentación adicional” se refiere a una fermentación por lotes con la diferencia de que se añaden uno o más sustratos o suplementos (por ejemplo, añadidos en suplementos 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 continuamente un medio definido de fermentación a un fermentador y se remueve simultáneamente una cantidad igual de medio utilizado o “acondicionado”, preferiblemente para recuperar el 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 tales que se produce un compuesto deseado (por ejemplo, un ácido graso insaturado, por ejemplo, DHA)” incluye el mantenimiento y/o el cultivo de plantas o microorganismos bajo condiciones (por ejemplo, de temperatura, presión, pH, duración, etc.) apropiadas o suficientes para obtener la producción del compuesto deseado o para obtener los rendimientos deseados del compuesto particular que está siendo producido. Por ejemplo, se continúa un cultivo durante un tiempo suficiente para producir la cantidad deseada de un ácido graso insaturado (por ejemplo, DHA). Preferiblemente, se continúa el cultivo durante un 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 tiempo suficiente para alcanzar rendimientos en la producción de ácidos grasos insaturados, por ejemplo, se cultivan células de tal manera que se produzcan aproximadamente al menos 15 a 20 g/L de ácidos grasos insaturados, que se produzcan aproximadamente al menos 20 a 25 g/L de ácidos grasos insaturados, que se produzcan aproximadamente al menos 25 a 30 g/L de ácidos grasos insaturados, que se produzcan aproximadamente al menos 30 a 35 g/L de ácidos grasos insaturados, que se produzcan aproximadamente al menos 35 a 40 g/L de ácidos grasos insaturados (por ejemplo, aproximadamente al menos 37 g/L de ácidos grasos insaturados) o que se produzcan aproximadamente al menos 40 a 50 g/L de ácidos grasos insaturados. En aún otra modalidad, se cultivan los microorganismos bajo condiciones tales que se produzca un rendimiento preferido de ácidos grasos insaturados, por ejemplo, un rendimiento dentro de un rango como el expuesto anteriormente, aproximadamente en 24 horas, aproximadamente en 36 horas, aproximadamente en 48 horas, aproximadamente en 72 hora, o aproximadamente en 96 horas.

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

La mitología de la presente invención puede incluir además una etapa de recuperación de un compuesto deseado (por ejemplo, un ácido graso insaturado). El término “recuperación” de un compuesto deseado incluye la extracción, recolección, aislamiento o purificación del compuesto a partir del medio de cultivo. La recuperación del compuesto se puede llevar a cabo de acuerdo a cualquier metodología convencional de aislamiento o purificación conocida en el arte incluyendo, pero sin limitarse a, tratamiento con una resina convencional (por ejemplo, una resina aniónica o catiónica de intercambio, una resina no iónica de adsorción, etc.), tratamiento con un adsorbente convencional (por ejemplo, carbón activado, ácido silícico, gel de sílice, celulosa, alúmina, etc.), alteración del pH, extracción con solventes (por ejemplo, con un solvente convencional tal como un alcohol, acetato de etilo, hexano y similares), diálisis, filtración, concentración, cristalización, recristalización, ajuste de pH, liofilización y similares. Por ejemplo, se puede recuperar un compuesto a partir 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 los cationes no deseados y luego a través o sobre una resina de intercambio aniónico para remover los aniones inorgánicos no deseados y los ácidos orgánicos que tienen una acidez mayor que los ácidos grasos no saturados 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 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 otros componentes del medio o subproductos de la fermentación), más preferiblemente aproximadamente más del 90% del compuesto deseado (por ejemplo, aproximadamente menos del 10% de 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 otros componentes del medio o subproductos de la fermentación), y lo más preferible aproximadamente más del 98-99% del compuesto deseado (por ejemplo, aproximadamente menos del 1-2% de otros componentes del medio o subproductos de la fermentación). Cuando el compuesto deseado es un ácido graso insaturado que ha sido derivatizado hasta una sal, el compuesto está preferiblemente libre además (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 derivatizado hasta un alcohol, el compuesto está preferiblemente libre además (por ejemplo, sustancialmente libre) de contaminantes químicos asociados con la formación del alcohol.

- 5 En una modalidad alternativa, el ácido graso insaturado deseado no es purificado a partir de la planta o el microorganismo, por ejemplo, cuando la planta o el microorganismo no es biológicamente peligroso (por ejemplo, seguro). Por ejemplo, se pueden utilizar la planta entera o el cultivo (o el sobrenadante del cultivo) como fuente del producto (por ejemplo, un producto crudo). En una modalidad, se utiliza la planta o 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  
10 modalidad, se pulveriza, seca, o liofiliza la planta o el cultivo (o el sobrenadante del cultivo).

#### B. Metodologías para la Producción con Alto Rendimiento

Una modalidad particularmente preferida de la presente invención es un método de producción con alto rendimiento para la producción de ácidos grasos insaturados, por ejemplo, DHA, que comprende el cultivo de una planta o microorganismo manipulados bajo condiciones tales que se produce el ácido graso insaturado con un rendimiento significativamente alto. La frase “método de producción con alto rendimiento”, por ejemplo, un método de producción con 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 en un nivel elevado o que está  
20 por encima de lo usual para métodos de producción comparables. Preferiblemente, un método de producción con 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 de rendimiento que es significativamente elevado o que está por encima de lo usual para métodos de producción comparables, por ejemplo, que se eleva hasta un nivel suficiente para la producción comercial del producto deseado (por ejemplo, la producción del producto con un costo  
25 comercialmente razonable). En una modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 2 g/L. En otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso  
30 insaturado en un nivel superior a 10 g/L. En otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 20 g/L. En aún otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce  
35 un ácido graso insaturado en un nivel superior a 30 g/L. En aún otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 40 g/L.

La invención presenta además un método de producción con 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 manipulado bajo condiciones tales que se produce un nivel suficientemente elevado de compuesto dentro de un período de tiempo comercialmente deseable. En un ejemplo de modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo  
45 de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 15-20 g/L en 36 horas. En otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 25-30 g/L en 48 horas. En otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 35-40 g/L en 72 horas, por ejemplo, superior a 37 g/L en 72 horas. En otra modalidad, la invención presenta un método de producción con alto rendimiento para la producción de ácidos grasos insaturados que incluye el cultivo de una planta o microorganismo manipulado bajo condiciones tales que se produce un ácido graso insaturado en un nivel superior a 30-40 g/L en 60 horas; por ejemplo,  
50 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 se encuentren dentro del alcance de la presente invención. Por ejemplo, se pretende que queden incluidos niveles de producción de ácido graso insaturado de al menos 31, 32, 33, 34, 35, 36, 37, 38 y 39 g/L en 60 horas dentro del rango de 30-40 g/L en 60 horas. En otro ejemplo, se pretende que queden incluidos rangos de 30-35 g/L o 35-40 g/L dentro del rango de 30-40 g/L en 60 horas. Además, la persona entrenada en el arte  
60 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 adicionales de tiempo (por ejemplo, períodos de tiempo mayores a 60 horas), resultando opcionalmente en rendimientos aún mayores de un ácido graso insaturado que está siendo producido.

#### 65 IV. Composiciones

Se pueden utilizar las moléculas de ácido nucleico para desaturasa y las proteínas desaturasa de la invención para producir ácidos grasos insaturados que pueden ser incorporados en composiciones. Las composiciones incluyen, por

ejemplo, composiciones para uso como alimento para animales, composiciones para uso como nutracéuticos (por ejemplo, suplementos para la dieta), y composiciones farmacéuticas adecuadas para administración.

Tales composiciones farmacéuticas típicamente incluyen un ácido graso insaturado y un portador farmacéuticamente aceptable. Como se la utiliza aquí, la expresión “portador farmacéuticamente aceptable” pretende incluir a cualquiera y todos los solventes, medios de dispersión, recubrimientos, agentes antibacteriales y antifúngicos, agentes isotónicos y para retraso de la absorción, y similares, compatibles con la administración farmacéutica. El uso de tales medios y agentes para sustancias farmacéuticamente activas es bien conocido en el arte. Salvo porque cualquier medio o agente convencional sea incompatible con el compuesto activo, se contempla el uso de los mismos en las composiciones. Se pueden incorporar también compuestos activos suplementarios en las composiciones.

Se formula una composición farmacéutica para ser compatible con su ruta pretendida de administración. Los ejemplos de rutas de administración incluyen administración parenteral, por ejemplo, intravenosa, intradérmica, subcutánea, 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 estéril 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 de quelación tales como ácido etilendiaminotetraacético; amortiguadores tales como acetatos, citratos o fosfatos y agentes para ajustar 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. Se puede colocar la preparación parenteral en ampollitas, jeringas desechables o viales de vidrio o plástico para múltiples dosis.

Las composiciones farmacéuticas adecuadas para uso como inyectables incluyen soluciones acuosas estériles (cuando son solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones estériles inyectables. Para administración intravenosa, los portadores adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor EL<sup>TM</sup> (BASF, Parsippany, NJ) o solución salina amortiguada con sulfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida en una medida en que pueda ser utilizada fácilmente para ser inyectada. Debe ser estable bajo las condiciones de fabricación y almacenamiento y debe ser preservada contra la acción contaminante de microorganismos tales como bacterias hongos. El portador puede ser un solvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilén glicol, y polietilén glicol líquido, y similares), y mezclas adecuadas de los mismos. Se puede mantener una fluidez adecuada, 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. Se puede lograr la prevención de la acción de los microorganismos por medio de diferentes agentes antibacteriales y antifúngicos, por ejemplo, parabenos, clorbutanol, fenol, ácido ascórbico, timerosal, y similares. En muchos casos será 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 provocada incluyendo en la composición un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina. Se pueden preparar soluciones inyectables estériles por medio de la incorporación del compuesto activo (por ejemplo, un LCPUFA, o un fragmento del mismo, producido por las moléculas de ácido nucleico y proteína de la presente invención) en la cantidad requerida en un solvente apropiado con uno o con una combinación de ingredientes enumerados anteriormente, según se requiera, seguido por esterilización por filtración. Generalmente, se preparan dispersiones por medio de la incorporación del compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los otros ingredientes requeridos a partir de aquellos denominados más arriba. En el caso de polvos estériles para la preparación de soluciones estériles inyectables, los métodos preferidos de preparación son secado al vacío y liofilización que producen un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de la solución previamente filtrada en forma estéril del mismo.

Las composiciones orales generalmente incluyen un diluyente inerte o un portador comestible. Pueden ser colocadas en cápsulas de gelatina o comprimidas en tabletas. Para los propósitos de administración terapéutica oral se puede incorporar el compuesto activo con excipientes y utilizarlo en la forma de tabletas, trociscos, o cápsulas. Se pueden preparar también composiciones orales utilizando un fluido portador para uso como un enjuague bucal, en donde el compuesto en el portador fluido se aplica oralmente y se hacen buches y se expectora o se traga. Se pueden incluir agentes aglutinantes farmacéuticamente compatibles, y/o materiales adyuvantes como parte de la composición. Las tabletas, píldoras, cápsulas, trociscos 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 algínico Primogel, o almidón de maíz; un lubricante tal como estearato de magnesio o Esteroatos; un deslizante tal como dióxido de silicio coloidal; un agente edulcorante tal como sacarosa o sacarina; o un agente saborizante tal como menta, salicilato de metilo, o saborizante de naranja.

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

La administración sistémica puede ser también por medio transmucosa o transdérmica. Para administración transmucosa o transdérmica, se utilizan en la formulación penetrantes apropiados para la barrera que va a ser permeada. 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 se puede lograr a través del uso de atomizadores nasales o supositorios. Para administración transdérmica, se formulan los compuestos activos en ungüentos, bálsamos, geles o cremas como generalmente se conoce en el arte.

- 5 Se pueden preparar también los compuestos 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.

10 Se pueden preparar los compuestos activos con portadores que protejan al compuesto contra la eliminación rápida por parte del organismo, tales como una formulación de liberación controlada, incluidos implantes y sistemas de suministro microencapsulados. Se pueden utilizar polímeros biocompatibles biodegradables, tales como etilén vinil acetato, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y ácido poliláctico. Los métodos para la preparación de tales formulaciones serán evidentes para aquellos capacitados en el arte. Se pueden obtener también los materiales comercialmente con Alza Corporation y Nova Phramaceutical, Inc. Se pueden utilizar también suspensiones liposomales (incluidos liposomas destinados a células infectadas con anticuerpos monoclonales como antígenos virales) como portadores farmacéuticamente aceptables. Estos se pueden preparar de acuerdo a métodos conocidos por aquellos capacitados en el arte, por ejemplo, como se describe en la Patente Estadounidense No. 4.522.811.

20 Es especialmente conveniente la formulación de composiciones morales o parenterales en forma de dosis unitarias para una fácil administración y uniformidad de las dosis. Una forma unitaria de dosificación como se la utiliza aquí se refiere a unidades físicamente discretas adecuadas como dosis unitarias para el individuo que va a ser tratado, conteniendo cada unidad una cantidad predeterminada de compuesto activo calculada para producir el efecto terapéutico deseado junto con el portador farmacéutico requerido. Las especificaciones para las formas unitarias de dosificación están dictadas por, y dependen directamente de las características únicas del compuesto activo y del efecto terapéutico particular que se pretende lograr, y de las limitaciones inherentes en el arte para la elaboración de tales compuestos activos para el tratamiento de individuos.

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

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

40 Como se define aquí, una cantidad terapéuticamente efectiva de proteína o de polipéptido (es decir una dosis efectiva) está en un rango aproximadamente desde 0,001 hasta 30 mg/kg de peso corporal, preferiblemente aproximadamente desde 0,01 hasta 25 mg/kg de peso corporal, más preferiblemente aproximadamente desde 0,1 hasta 20 mg/kg de peso corporal y aún más preferiblemente aproximadamente desde 1 hasta 10 mg/kg desde 2 hasta 9 mg/kg, desde 3 hasta 8 mg/kg, desde 4 hasta 7 mg/kg, o desde 5 hasta 6 mg/kg de peso corporal. La persona capacitada se dará 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 del trastorno, a tratamientos previos, a la salud general y/o la edad del individuo, y a otras enfermedades presentes. Además, el tratamiento de un individuo con una cantidad terapéuticamente efectiva de una proteína, polipéptido, o antibiótico puede incluir un único tratamiento o, preferiblemente, puede incluir una serie de tratamientos.

60 En un ejemplo preferido se trata a un individuo con un LCPUFA en el rango aproximadamente entre 0,1 hasta 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 3 a 7 semanas, e incluso más preferiblemente aproximadamente durante 4, 5 ó 6 semanas. Se apreciara también que la dosis efectiva de anticuerpo, proteína, o polipéptido utilizada para el tratamiento, puede incrementarse o disminuirse durante el transcurso de un tratamiento particular. Los cambios en las dosis pueden resultar y hacerse evidentes a partir de los resultados de ensayos de diagnóstico como los descritos aquí.

65 Las composiciones farmacéuticas se pueden incluir en un contenedor, empaque, o dispensador junto con las instrucciones para su administración.

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

## Ejemplos

**Materiales:** Se adquirió *Thraustochytrium sp.* ATCC 21685 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., y colaboradores (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 una genoteca de ADNc*

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 genoteca de ADNc a partir del ARN total. Se sintetizó la primera hebra de ADNc por medio de la transcriptasa inversa superscript II de Gibco-BRL. Se sintetizó la segunda hebra de ADNc por medio de la ADN polimerasa de Stratagene. Después del fraccionamiento de tamaño, se ligaron los insertos de ADNc mayores a 1 kb dentro del vector  $\lambda$  Uni-Zap XR (Stratagene). Se empacaron luego los ADN recombinantes con extracto de empaquetamiento Gigapack III Gold (Stratagene) y se sembró sobre placas NZY. La genoteca resultante representó más de  $5 \times 10^6$  clones independientes. Se llevó a cabo la selección de la genoteca ADNc 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*

Se sintetizó el ADNc monocatenario por medio de la transcriptasa inversa superscript II (Gibco-BRL) a partir del ARN total y se lo utilizó luego como molde para la reacción PCR con dos iniciadores degenerados (El iniciador hacia adelante: GCXCA/GAXGAXCAC/TCCXGGXGG y el iniciador inverso: ATXTG/TXGGA/GAAXAG/AG/ATGG/ATG). La amplificación por medio de 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 amplificación a 72°C durante 10 min. Se aislaron los productos amplificados de 800 pb a 1000 pb a partir del gel de agarosa y se purificó por medio de un kit (purificación a través del gel Qiaex II, Qiagen), y posteriormente se clonó en el reactor de clonación TA pCR<sup>®</sup>2.1. (Invitrogen). Se secuenciaron luego los insertos clonados por medio del PRISM DyeDeoxy Terminator Cycle Sequencing System (Perkin Elmer/Applied Biosystems).

### Ejemplo 3

#### *Expresión de Fad4 y Fad5 en levadura*

Se amplificaron los marcos de lectura abiertos de Fad4 y Fad5 por medio de PCR utilizando la enzima Precision Plus (Stratagene) y se clonó en un vector de clonación TA (pCR<sup>®</sup>2.1, Invitrogen). Habiendo confirmado que los productos de la PCR eran idénticos a los ADNc originales por medio de secuenciación, se liberaron luego 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*.

Se transformaron cepas de levadura InvSc2 (Invitrogen) con las construcciones para expresión utilizando el método de acetato de litio y se seleccionaron los transformantes sobre placas con medio mínimo que carecía de uracilo (Gietz, D., y colaboradores (1992) *Nucleic Acids Res.* 20: 1425; Covello, P. S. y Reed, D. W. (1996) *Plant Physiol.* 111: 223-226).

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 centrifugaron las células, se las lavó y resuspendió en agua destilada. Se inoculó medio mínimo que contenía 2% de galactosa, con o sin sustrato 0,3 mM de ácidos grasos 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 de ácidos grasos*

Se recolectaron dos veces con agua destilada células de levadura y de *Thraustochytrium*. Luego se añadieron 2 mL de KOH metanólico (KOH al 7,5% p/v en metanol al 95%) a los materiales y se calentó la mezcla sellada en un tubo

## ES 2 343 106 T3

de vidrio para cultivo de 12 ml 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 la fase acuosa restante 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 en hexano y se secó bajo una corriente de nitrógeno. Se añadieron 2 mL de HCl metabólico 3 N (SUPELCO, Supelco Park, Bellefonte, PA 16823-0048) y se calentó la mezcla a 80°C durante 2 horas. Después de enfriar hasta 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 una corriente de nitrógeno. Se analizaron los ésteres metílicos resultantes de ácido graso (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).

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

### Ejemplo 5

*Transformación de Brassilca juncea y de lino (Linum usitatissimum) y tratamiento exógeno del ácido graso*

Se utilizaron los hipocotiledones de plántulas de 5-6 días de *B. juncea* y de lino 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 exógeno del ácido graso. Se dividieron las plántulas en tres partes, hojas, tallos y raíces. Cada una fue cortada en pedazos muy pequeños y colocada en una placa de titulación de 24 pozos. A cada pozo se le añadieron 2 mL de la sal sódica 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 incubación, se lavaron tres veces los tejidos de la planta con agua y luego se los utilizó para análisis de ácidos grasos.

### Ejemplo 6

*Perfil de ácidos grasos de la Thrauschytrium sp.*

*Thraustochytrium* ha llamado recientemente la atención de los científicos debido a su habilidad para producir LCFUFA tal como DHA, AA, EPA y DPA. La Figura 11 muestra la composición de ácidos grasos de los lípidos aislados a partir de cultivos de 7 días de *Thraustochytrium sp.*. Como se muestra en las tablas, los microorganismos contienen un amplio rango de ácidos grasos poliinsaturados, tanto de familias n-3 como n-6, de ácidos grasos A6 de 18 átomos de carbono (ácido gama-linolénico y ácido estearidónico) hasta ácidos grasos  $\Delta 4$  de 22 átomos de carbono (DHA y DPA). Los organismos, especialmente *Thraustochytrium sp.*, parecen contener un juego completo de enzimas de desaturación y alargamiento requeridas para la biosíntesis de DHA y DPA. La cepa carece de ácidos grasos poliinsaturados de 24 átomos de carbono, de los precursores propuestos para DHA, y de la síntesis de DPA en la ruta de Precher (Voss, A., y colaboradores (1991) J. Biol. Chem. 266: 19995-20000; Mohammed, B. S., y colaboradores (1997) Biochem. J. 326: 425-430). Los ácidos grasos de 24 átomos de carbono puede que no estén involucrados en la síntesis *in vivo* de ácidos grasos  $\Delta 4$  de 22 átomos de carbono tales como DHA y DPA en *Thraustochytrium sp.*

### Ejemplo 7

*Identificación de los ADNc que codifican para la desaturasa "Front-end"*

Para identificar los genes que codifican para desaturasas involucradas en la biosíntesis de los LCFUFA en *Thraustochytrium sp.*, se adoptó una estrategia de clonación con base en la PCR. Se diseñaron dos iniciadores degenerados para dirigir el motivo de enlazamiento hemo de extensión N-terminal del dominio tipo cyt b5 en desaturasas front-end y el tercer motivo conservado de histidina en todas las desaturasas microsomales, respectivamente. El razonamiento detrás del diseño es que las desaturasas involucradas en la biosíntesis de EPA y DHA en *Thraustochytrium sp.*, debe tener estructura de primacía similar como otras desaturasas front-end, es decir extensión N-terminal del dominio tipo cyt b5 en la desaturasa. Se identificaron cuatro fragmentos de los ADNc de *Thraustochytrium sp.* que codifican proteínas de fusión que contienen al dominio tipo cyt b5 en el terminal N.

Para aislar clones de ADNc de longitud completa, se utilizaron los cuatro insertos como sondas para seleccionar genotecas de ADN de *Thraustochytrium sp.* lo cual resultó en la identificación de varios clones de ADNc en cada grupo. La secuenciación de todos esos clones identificó cuatro ADNc de longitud completa que fueron llamados como Fad4, Fad5, Fad5-2 y Fad6. El marco de lectura abierto de Fad4 es de 1560 pb y codifica para 519 aminoácidos con peso molecular de 59,1 kDa (Figura 1). Fad5 es de 1230 pb de longitud y codifica para 439 aminoácidos con peso molecular de 49,8 kDa (Figura 2). Una comparación de secuencias de estas dos secuencias de *Thraustochytrium sp.* mostró únicamente un 16% de identidad de aminoácidos entre las proteínas deducidas. Un análisis detallado reveló

## ES 2 343 106 T3

que Fad4 es 80 aminoácidos más largo que Fad5, que se presentan entre el segundo y el tercer motivos conservados de histidina (Figura 3).

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

Fad 4 (519 residuos aminoácidos)				
Blastp nr				
Acceso No.	Organismo	Descripción	Longitud	% de Identidad
AF067654	<i>Mortierella alpina</i>	$\Delta 5$ desaturasa de ácido graso	509	29
AF054824	<i>Mortierella alpina</i>	$\Delta 5$ desaturasa microsomal	509	28
A022097	<i>Dictyostelium discoideum</i>	$\Delta 5$ desaturasa de ácido graso	507	27
AB029311	<i>Dictyostelium discoideum</i>	desaturasa de ácido graso	519	26
L11421 D90914	<i>Synechocystis sp.</i>	$\Delta 6$ desaturasa	410	25

Fad 5 (439 residuos aminoácidos)				
Blastp nr				
Acceso No.	Organismo	Descripción	Longitud	% de Identidad
AF139720	<i>Euglena gracilis</i>	$\Delta 8$ desaturasa de ácido graso	404	29
AF007561	<i>Borago officinalis</i>	$\Delta 6$ desaturasa	421	27
U79010	<i>Borago officinalis</i>	$\Delta 6$ desaturasa	421	27
A309556	<i>Danio rerio</i>	$\Delta 6$ desaturasa de ácido graso	422	26
AF110510	<i>Mortierella alpina</i>	$\Delta 6$ desaturasa de ácido graso	463	25

### Ejemplo 8

#### Expresión de Fad4 y Fad5 en levadura

Para confirmar la función de Fad4, se expresó el ADNc de longitud completa en la cepa de levadura InvSc2 bajo el control del promotor inducible. La Figura 4 muestra que con suplementación del medio con 22:5 (7, 10, 13, 16, 19), las células de levadura que contenían ADNc de Fad4 tenían un ácido graso extra comparadas con el control del vector. El pico tiene un tiempo de retención idéntico al del estándar de DHA. El análisis por LC/MS del ácido graso libre mostró que produce iones moleculares desprotonados ( $m/z = 279$ ) idénticos al estándar de DHA en electrospray de iones negativos. Además, el análisis por medio de GC/MS del FAME confirmó que el espectro del pico es idéntico a aquel del estándar de DHA (Figura 8). Estos resultados indican que Fad4 es una  $\Delta 4$  desaturasa de ácido graso que es capaz de introducir un doble enlace en posición 4 del sustrato 22:5 (7, 10, 13, 16, 19), resultando en un ácido graso  $\Delta 4$  desaturado, DHA (22:6-4, 7, 10, 13, 16, 19).

Para estudiar adicionalmente la especificidad del sustrato de la Fad4, se suministraron separadamente una cantidad de sustratos incluyendo 18:2 (9, 12), 18:3 (9, 12, 15), 20:3 (8, 11, 14) y 22:4 (7, 10, 13, 16) a los transformantes de las levaduras. Los resultados indican que Fad4 podría utilizar también 22:4 (7, 10, 13, 16) como sustrato (Figura 6) para producir otro ácido graso  $\Delta 4$  desaturado, DPA (22:5-4, 7, 10, 13, 16) (Figura 7). El resto de los ácidos grasos examinados no fueron sustratos efectivos.

Para confirmar la función de Fad5, se transformó InvSc2 de *S. cerevisiae* con plásmidos, que contenían el marco de lectura abierto de la Fad5 bajo el control del promotor inducible por galactosa. Cuando se indujeron los transformantes de levadura por medio de galactosa en un medio que contenía ácido homo-gama-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 (Figura 8). Una comparación del cromatograma con aquel de los estándares reveló que el ácido graso tenía un

tiempo de retención idéntico al estándar de ácido araquidónico (AA, 20:4-5, 8, 11, 14). Para confirmar adicionalmente la regioquímica de los productos, se analizaron los FAME por medio de GC/MS. La Figura 19 indica que el espectro de masas del nuevo ácido graso y el estándar de AA son idénticos. Estos resultados demuestran que Fad5 convierte HGLA (20:3-8, 11, 13) en AA (20:4-5, 8, 11, 14) en levadura.

#### Ejemplo 9

##### *Expresión de Fad4 en B. juncea*

Para determinar si Fad4 de *Thraustochytrium* es funcional en cultivos de semillas oleaginosas, se transformaron *B. juncea* con la construcción que contenía Fad4 bajo el control de un promotor constitutivo. Se obtuvieron ocho plantas transgénicas independientes. En *B. Juncea* no hay sustratos disponibles de  $\Delta 4$  desaturasa de ácido graso. Por lo tanto, para examinar la actividad de la enzima transgénica en las plantas, se debe suministrar en forma exógena el sustrato 22:5 (n-3). En este experimento, se aplicaron tanto el tipo silvestre como los transgénicos con soluciones acuosas de docosapentanoato sódico. Se encontró que los sustratos aplicados en forma exógena eran fácilmente fijados por las raíces, los tallos y las hojas de ambos tipos de plantas, pero convertidos en DRA únicamente en los transgénicos. Las hojas tiene un nivel más alto de producción de que las raíces y los tallos. En las hojas se incorporó sustrato exógeno hasta un nivel de 10-20% de los ácidos grasos totales y se produjo ácido graso  $\Delta 4$  desaturado (22:6, n-3) en un rango de 3-6% de los ácidos grasos totales (Figura 16). Estos resultados indican que la  $\Delta 4$  desaturasa de ácido grado de *Thraustochytrium* es funcional en cultivos de semillas oleaginosas.

#### 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.

#### Documentos de patente citados en la descripción

- US 5721137 A, Frascotti [0085]
- US 4522811 A [0122]

#### Literatura citada en la descripción que no es de patente:

- Crawford, M. A. y colaboradores, *Am. J-Clin. Nutr.*, 1997, vol. 66, 1032S-1041S [0003]
- Giusto, N. M. y colaboradores, *Prog. Lipid Res.*, 2000, vol. 39, 315-391 [0003]
- Martinetz, M. J. *Pediatr.*, 1992, vol. 120, S129-S138 [0003] [0034]
- Giusto, N. M. y colaboradores, *Prog. Lipid Res.*, 2000, vol. 39, 315-391 [0003] [0034]
- 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. y colaboradores, *J. Lipid Res.*, 1995, vol. 36, 2471-2477 [0004]
- Sprecher, H. y colaboradores, *Lipids*, 1999, vol. 34, S153-S156 [0004]
- Sambrook, J. y colaboradores, *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press*, 1989 [0047]
- Current Protocols in Molecular Biology. *John Wiley & Sons, Inc.*, 1995 [0067]
- Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Press*, 1989 [0067]
- Church; Gilbert. *Proc. Natl. Acad. Sci. USA*, 1984, vol. 81, 1991-1995 [0067]
- Needleman; Wunsch. *J. Mol. Biol.*, 1970, 444-453 [0081]
- Meyers; Miller. *Comput. Appl. Biosci.*, 1988, vol. 4, 11-17 [0082]
- Altschul y colaboradores, *J. Mol. Biol.*, 1990, vol. 215, 403-10 [0083]



## ES 2 343 106 T3

• **Altschul y colaboradores**, *Nucleic Acids Res.*, 1997, vol. 25 (17), 3389-3402 [0083]

• **Sambrook, J.; Fritsh, E. F.; Maniatis, T.** Molecular Cloning: A Laboratory Manual. *Gold Spring Harbor Laboratory Press*, 1989 [0087]

• **Weete, J. D. y colaboradores**, *Lipids*, 1997, vol. 32, 839-845 [0130]

• **Qiu, X.; Eriekson, 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* [0131]

• **Gietz, D. y colaboradores**, *Nucleic Acids Res.*, 1992, vol. 20, 1425 [0134]

• **Covello, P. S.; Reed, D. W.** *Plant Physiol.*, 1996, vol. 111, 223-226 [0134] [0136]

• **Voss, A. y colaboradores**, *J. Biol. Chem.*, 1991, vol. 266, 19995-20000 [0139]

• **Mohammed, B. S. y colaboradores**, *Biochem. J.*, 1997, vol. 326, 425-430 [0139].

# REIVINDICACIONES

1. Una molécula aislada de ácido nucleico seleccionada del grupo que consiste de:

- (a) una molécula de ácido nucleico que comprende la secuencia de nucleótidos expuesta en la SEQ ID NO: 1, o en la SEQ ID NO: 3, o un complemento de la misma; y
- (b) una molécula de ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos expuesta en la SEQ ID NO: 2, o en la SEQ ID NO: 4, o un complemento de la misma.

2. Una molécula aislada de ácido nucleico seleccionada del grupo que consiste de:

- (a) una molécula de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que es aproximadamente al menos 50% idéntica a la secuencia entera de aminoácidos de la SEQ ID NO: 2 ó 4, en donde el polipéptido tiene una actividad de  $\Delta 4$  (en el caso de la SEQ ID NO: 2) o  $\Delta 5$  desaturasa (en el caso de la SEQ ID NO: 4), o un complemento de las mismas; y
- (b) una molécula de ácido nucleico que hibrida bajo condiciones rigurosas hasta un complemento de una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la SEQ ID NO: 1 ó 3, en donde la molécula de ácido nucleico codifica un polipéptido que tiene una actividad de  $\Delta 4$  (en el caso de la SEQ ID NO: 1) o  $\Delta 5$  desaturasa (en el caso de la SEQ ID NO: 3), o un complemento de las mismas.

3. Un vector que comprende la molécula de ácido nucleico de la reivindicación 1 ó 2.

4. El vector de la reivindicación 3, que es un vector de expresión.

5. Una célula huésped transfectada con el vector de expresión de la reivindicación 4.

6. La célula huésped de la reivindicación 5, en donde la célula es seleccionada del grupo que consiste de una célula vegetal, una célula microbiana, y una célula animal.

7. La célula huésped de la reivindicación 6, en donde la célula vegetal es de una planta de semillas oleaginosas seleccionada del grupo que consiste de lino (*Linum* sp.), colza (*Brassica* sp.), soja (*Glycine* y *Soja* sp.), girasol (*Helianthus* sp.), algodón (*Gossypium* sp.), maíz (*Zea mays*), oliva (*Olea* sp.), cártamo (*Carthamus* sp.), cacao (*Theobroma cacao*), y maní (*Arachis* sp.).

8. La célula huésped de la reivindicación 6, en donde la célula microbiana es seleccionada del grupo que consiste de *Thraustochytrium*, *Pythium irregulare*, *Schizichytrium*, y *Cryptocodinium*.

9. Un método para producir un polipéptido que comprende el cultivo de la célula huésped de la reivindicación 5 para producir el polipéptido.

10. Un polipéptido aislado seleccionado del grupo que consiste de:

- (a) un polipéptido que comprende una secuencia de aminoácidos que es al menos 50% idéntica a la secuencia entera de aminoácidos de la SEQ ID NO: 2 ó 4 y que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 2 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 4;
- (b) un polipéptido que es codificado por una molécula de ácido nucleico que hibrida bajo condiciones rigurosas hasta un complemento de una molécula de ácido nucleico que comprende la secuencia de nucleótidos de la SEQ ID NO: 1 ó 3 y que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 1 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 3; y
- (c) un polipéptido que comprende la secuencia de aminoácidos de la SEQ ID NO: 2 ó 4 que tiene una actividad de  $\Delta 4$  desaturasa en el caso de la SEQ ID NO: 2 y que tiene una actividad de  $\Delta 5$  desaturasa en el caso de la SEQ ID NO: 4.

11. El polipéptido aislado de la reivindicación 10 que comprende la secuencia de aminoácidos de las SEQ ID NO: 2 ó 4.

12. Un método para transformar un ácido graso insaturado que comprende transfectar o transformar una célula con la molécula aislada de ácido nucleico de la reivindicación 1 ó 2 y cultivar la célula bajo condiciones tales que se produzca el ácido graso insaturado.

13. El método de la reivindicación 12, en donde se selecciona la célula del grupo que consiste de una célula vegetal, una célula animal, y una célula microbiana.

## ES 2 343 106 T3

14. El método de la reivindicación 13, en donde la célula vegetal es seleccionada de una planta de semillas oleaginosas.

15. El método de la reivindicación 14, en donde se selecciona la planta de semillas oleaginosas del grupo que consiste de lino (*Linum* sp.), colza (*Brassica* sp.), soja (*Glycine* y *Soja* sp.), girasol (*Helianthus* sp.), algodón (*Gossypium* sp.), maíz (*Zea mays*), oliva (*Olea* sp.), cártamo (*Carthamus* sp.), cacao (*Theobroma cacao*), y maní (*Arachis* sp.).

16. El método de la reivindicación 12, que comprende además la etapa de recuperación del ácido graso insaturado.

17. El método de la reivindicación 12, en donde se selecciona el ácido graso insaturado del grupo que consiste de AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

18. Un método para producir un ácido graso insaturado que comprende poner en contacto una composición que contiene al menos una molécula objetivo de desaturasa con al menos un polipéptido aislado de la reivindicación 10 u 11 bajo condiciones tales que se produce el ácido graso insaturado.

19. El método de la reivindicación 18, que comprende además la etapa de recuperación del ácido graso insaturado.

20. El método de la reivindicación 18, en donde se selecciona al ácido graso insaturado del grupo que consiste de AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

21. Un método para producir una célula vegetal o microbiana capaz de generar un ácido graso insaturado que comprende, la introducción en dicha célula de la molécula de ácido nucleico de la reivindicación 1 ó 2, en donde la molécula de ácido nucleico codifica una desaturasa que tiene actividad para catalizar la formación de un doble enlace en una cadena acilo grasa.

22. El método de la reivindicación 21, en donde la célula es una célula vegetal.

23. El método de la reivindicación 22, en donde la célula vegetal es de una planta de semillas oleaginosas.

24. El método de la reivindicación 23, en donde la planta de semillas oleaginosas es lino (*Linum* sp.), colza (*Brassica* sp.), soja (*Glycine* y *Soja* sp.), girasol (*Helianthus* sp.), algodón (*Gossypium* sp.), maíz (*Zea mays*), oliva (*Olea* sp.), cártamo (*Carthamus* sp.), cacao (*Theobroma cacao*), y maní (*Arachis* sp.).

25. El método de la reivindicación 21, en donde se selecciona al ácido graso insaturado del grupo que consiste de AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

26. Un método para modular la producción de un ácido graso insaturado que comprende el cultivo de la célula de la reivindicación 5, de tal manera que ocurre la modulación de la producción de un ácido graso insaturado.

27. El método de la reivindicación 26, en donde se refuerza la producción del ácido graso insaturado.

28. El método de la reivindicación 26, en donde se selecciona al ácido graso insaturado del grupo que consiste de AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

29. El método de la reivindicación 26, que comprende además la recuperación del ácido graso insaturado.

30. Un método para la producción a gran escala de un ácido graso insaturado, que comprende el cultivo de la célula de la reivindicación 5, de tal manera que se produce el ácido graso insaturado.

31. El método de la reivindicación 30, en donde se refuerza la producción del ácido graso insaturado.

32. El método de la reivindicación 30, en donde se selecciona al ácido graso insaturado del grupo que consiste de AA 20:4 (5, 8, 11, 14), EPA 20:5 (5, 8, 11, 14, 17), DPA 22:5 (4, 7, 10, 13, 16), y DHA 22:6 (4, 7, 10, 13, 16, 19).

33. El método de la reivindicación 28, que comprende además la recuperación del ácido graso insaturado.

34. Una composición que comprende al polipéptido de la reivindicación 10 u 11.

35. Un método para producir un suplemento dietético que comprende al método de la reivindicación 12, 18, 26 ó 30.

36. Una composición que comprende a la célula producida por medio del método de la reivindicación 21.

37. La composición de la reivindicación 34, en donde se utiliza la composición en alimento para animales.

38. Un método de acuerdo a la reivindicación 35, en donde el suplemento dietético es suplemento para animales.

## ES 2 343 106 T3

39. Un suplemento dietético que comprende a la composición de la reivindicación 34.

40. Un método para suplementar la dieta de un humano o de un animal, que comprende la adición de la composición de la reivindicación 34.

5

10

15

20

25

30

35

40

45

50

55

60

65

ATGACGGTCGGCTACGACGAGGAGATCCCGTTCCAGCAGGTCCGCGCGCACAAACAGCCGGA  
 TGACGCCTGGTGC GCGATCCACGGGCACGTGTACGATGTGACCAAGTTCCGCGAGCGTGACCCC  
 GGGCGGCGACATTATCCTGCTGGCCGCAGGCAAGGAGGCCACCGTGCTGTACGAGACTTACC  
 ATGTGCGGGGCGTCTCGGACGCGGTGCTGCGCAAGTACCGCATCGGCAAGCTGCCGGACGGC  
 CAAGGCGGCGCGAACGAGAAGGAAAAGCGGACGCTCTCGGGCCTCTCGTCCGGCCTCGTACTA  
 CACGTGGAACAGCGACTTTTACAGGGTAATGCGCGAGCGCGTCGTGGCTCGGCTCAAGGAGC  
 GCGGCAAGGCCCGCCGCGGAGGCTACGAGCTCTGGATCAAGGCGTTCCTGCTGCTCGTCCGGCT  
 TCTGGAGCTCGCTGTACTGGATGTGCACGCTGGACCCCTCGTTCCGGGGCCATCCTGGCCGCCA  
 TGTGCTGGGCGTCTTTGCCGCCCTTTGTGGGCACGTGCATCCAGCACGACGGCAACCACGGCG  
 CCTTTGCCAGTCCGATGGGTCAACAAGGTTGCCGGGTGGACGCTCGACATGATCGGCGCCA  
 GCGGCATGACGTGGGAGTTCCAGCACGTCTGGGCCACCATCCGTACACGAACCTGATCGAG  
 GAGGAGAACGGCCTGCAAAAGGTGAGCGGCAAGAAGATGACACCAAGCTGGCCGACCAGG  
 AGAGCGATCCGGACGTCTTTTCCACGTACCCGATGATGCGCCTGCACCCGTGGCACCAGAAGC  
 GCTGGTACCACCGTTTCCAGCACATTTACGGCCCTTCATCTTTGGCTTCATGACCATCAACAA  
 GGTGGTCACGCAGGACGTCCGTGTGGTGCTCCGCAAGCGGCTCTTCCAGATTGACGCCGAGTG  
 CCGGTACGCGAGCCCAATGTACGTGGCGCGTTTCTGGATCATGAAGGCGCTCACGGTGCTCTA  
 CATGGTGGCCCTGCCGTGCTACATGCAGGGCCCGTGGCACGGCCCTCAAGCTGTTCCGATCGC  
 GCACTTTACGTGCGGCGAGGTGCTCGCAACCATGTTCAATTGTGAACCACATCATCGAGGGCGT  
 CTCGTACGCTTCCAAGGACGCGGTCAAGGGCACGATGGCGCCGCCGAAGACGATGCACGGCG  
 TGACGCCCATGAACAACACGCGCAAGGAGGTGGAGGCGGAGGCGTCCAAGTCTGGCGCCGTG  
 GTCAAGTCAGTCCCGCTCGACGACTGGGCCGTGCTCCAGTGCCAGACCTCGGTGAAGTGGAGC  
 GTCGGCTCGTGGTTCTGGAATCACTTTTCCGGCGGCCCTCAACCACCAGATTGAGCACCACCTG  
 TTCCCCGGRCTCAGCCACGAGACGTACTACCACATTGAGGACGTCTTTCAGTCCACCTGCGCC  
 GAGTACGGCGTCCCGTACCAGCACGAGCCTTCGCTCTGGACCGCGTACTGGAAGATGCTCGAG  
 CACCTCCGTACGCTCGGCAATGAGGAGACCCACGAGTCCTGGCAGCGCGCTGCCTGA

Fig. 1A

MTVGYDEEIPFEQVRAHNKPDDAWCAIHGHVYDVTKFASVHPGGDIILLAAKEATVLYETVHV  
 RGVSDAVLRKYRIGKLPDGGGANEKEKRTL SGLSSASYTTWNSDFYRVMRERVVARLKERGKA  
 RRGGYELWIKAFLLLVGFWSLYWMCTLDPSFGAILAAMSLGVFAAFVGTCTQEDGNHGAFAQS  
 RWNKVAGWTLDMIGASGMTWBFQHVVLGHHPYTNLIEENGLQKVS GKKMDTKLADQESDPD  
 VFSTYPMRLHPWHQKRWYHRFQHTYGPFI FGFMTINKVVTQDVGVVLRKRLFQIDAECRYASP  
 MYVARFWIMKALTVLYMVALPCYMQGPWHGLKLF AIAHFTCGEVLATMFIVNHIIEGVSYASKD  
 AVKGTMAPPKTMHGVTMNNTRKEVEAEASKSGAVVKSVP LDDWAVVQCQTSVNWSVGSWF  
 WNHFSGLNHQIEHHLFPGLSHETYYHIQDVFQSTCAEYGVPHYQHEPSLWTAYWKMLEHLRQLG  
 NEETHESWQRAA

Fig. 1B

ATGGGCAAGGGCAGCGAGGGCCGCGAGCGCGCGCGAGATGACGGCCGAGGCGAACGGCG  
 ACAAGCGGAAAACGATTCTGATCGAGGGCGTCCTGTACGACGCGACGAACTTTAAGCACCCG  
 GCGGTTTCGATCATCAACTTCTTGACCGAGGGCGAGGCGGGCGTGGACGCGACGAGGCGTA  
 CCGCGAGTTTCATCAGCGGTCCGGCAAGGCCGACAAGTACCTCAAGTCGCTGCCAAGCTGG  
 ATGCGTCCAAGGTGGAGTCGCGGTTCTCGGCCAAAGAGCAGGCGGGCGCGACGCCATGACG  
 CGCGACTACGCGGCCTTTTCGCGAGGAGCTCGTCGCGCGAGGGGTACTTTGACCCGTCGATCCCG  
 CACATGATTTACCGCGTCGTGGAGATCGTGGCGCTCTTCGCGCTCTCGTTCGCTCATGTCCA  
 AGGCCTCGCCACCTCGCTCGTGTGGCGTGGTGATGAACGGCATTGCGCAGGGCCGCTGCG  
 GCTGGGTTCATGCACGAGATGGGCCACGGGTCTTCACGGGCGTCATCTGGCTCGACGACCGG  
 ATGTGCGAGTTCTTCTACGGCGTCGGCTGCGGCATGAGCGGGCACTACTGGAAGAACCAGCA  
 CAGCAAGCACCACGCGCGGCCAACCGCCTCGAGCACGATGTCGATCTCAACACGCTGCCCT  
 GGTGCGCTTTAACGAGCGCGTCGTGCGCAAGGTCAAGCCGGGATCGCTGCTGGCGCTCTGGCT  
 GCGCGTGCAGGCGTACCTCTTTGCGCCCGTCTCGTGCCTGCTCATCGGCCCTGGCTGGACGCTC  
 TACCTGCACCCGCGCTACATGCTGCGCACCAAGCGGCACATGGAGTTCTGCTGGATCTTCGCG  
 CGCTACATTGGCTGGTTCTCGCTCATGGGCGCTCTCGGCTACTCGCCGGGCACCTCGGTGCGG  
 ATGTACCTGTGCTCGTTTCGGCCTCGGCTGCATTTACATTTTCTGCGAGTTGCGCGTCAGCCACA  
 CGCACCTGCGCGTGACCAACCCGGAGGACCAGCTGCACTGGCTCGAGTACGCGGGCCGACCAC  
 ACGGTGAACATTAGCACCAAGTCTGGCTCGTCACGTGGTGGATGTCGAACCTGAACTTTCAG  
 ATCGAGCACCACTCTTCCCCACGGCGCCGCGAGTTCCGCTTCAAGGAAATCAGTCTCGCGTC  
 GAGGCCCTCTTCAAGCGCCACAACCTCCCGTACTACGACCTGCCCTACACGAGCGCGGTCTCG  
 ACCACCTTTGCCAATCTTTATTCCGTGCGCCACTCGGTGCGCGCCGACACCAAGAAGCAGGAC  
 TGA

Fig. 2A

MKGSEGRSAAREMTAEANGDKRKTILIEGVLYDATNFKHPGGSIIINFLTEGEAGVDATQAYREF  
 HQRSGKADKYLKSLPKLDASKVESRFSAKEQARRDAMTRDYAAPREELVAEGYFDPSIPHMIYRV  
 VEIVALFALSFWLMSKASPTSLVLGVVMNGIAQGRCGWVMHEMGBGSFTGVIWLDLDRMCFFY  
 GVGCGMSGHYWKNQHSKHHAAPNRLEHDVDLNTLPLVAFNERVVVRKVKPGSLLALWLRVQAY  
 LFAPVSCLLIGLWTLYLHPRYMLRTKRMEFWWIFARYIGWPSLMGALGYSPGTSVGMYLCSFG  
 LGCIYIFLQFAVSHTHLPVTNPEDQLHWLEYAADHTVNIISTKSWLVTTWMSNLNFQIEHHLFPTA  
 PQFRFKBISPRVEALFKRHNLFPYYDLPYTSAVSTTFANLYSVGHSVGADTKKQD

Fig. 2B

FAD4 - MTVGYDEEIPFEQVRAHNKPDDAWCAIHGHVYDVTKFASVHPGGDIIL-L -50  
 FAD5 - MGKGSEGRSAAREMTAEANGDKRKTIIEGVLYDATNFK-HPGSSIINFL -50  
 FAD4 - AGKEATVLYETYHVRGVSDAVLRKYRIGKLPDGGANEKEKRTLSGLSS -100  
 FAD5 - EGEAGVDATQAYREFHQRSKGADKY-LKSLPKLDAS---KVESRFSACEQ -96  
 FAD4 - ASYYTWNSTFYRVMRERVVARLKERGKARRGGYELWIKAFLLLVGFWSSL -150  
 FAD5 - ARRDAMTRDYAAFREELVAEGYFDPSIPHMI-----YRVVEIVALFALSE -141  
 FAD4 - YWMCTLDPSFGAILAAMSLGVFAAFVGTICIQHDGNECAFAQSRWVNKVAG -200  
 FAD5 - WLMSKASPTSLVLGVVMN-G-IAQRCGWVMEHMGHGSFTGVIWLDDRMC -185  
 FAD4 - WTLDMIGASCMTWEFOHVLGHHPYTNLIEEENGLQKVSQKMDTKLADQE -250  
 FAD5 - FYGVGCGMSGHYWKNQHSK-HHAAPNRLEHDVDLNT----- -226  
 FAD4 - SDPDVFSTYPMRLHPWHQKRWYHRFOHIYGPFI FGFM TINKVVTQDVGV -300  
 FAD5 - LPLVAFNERVVRKVKPGSLLALWLRVQ-----AYLFAPVSCLLIGLGWT -270  
 FAD4 - VLRKRLFQIDAECRYASPMYVARFWIMKALTVLYMVALPCYMQGPWHGLK -350  
 FAD5 - LYLHPRYMLRTRHMEFVWIFARYIGWFSLMGALGYSPGT----- -310  
 FAD4 - LEAIAHFTCGEVLATMFIVNHIIEGVSYASKDAVKGTMAPPKTMEGVTPM -400  
 FAD5 - --SVGMYLCSFGLGCIYIFLQF-----AVSHTHLPVTNP -342  
 FAD4 - NNTRKEVEAEASKSGAVVKSVP LDDWAVVQCQTSVNWSVGSWFVWNHFSGG -450  
 FAD5 - EDQLHWLEYAADHT-----VNISTKSWLVTWWMNS -372  
 FAD4 - LNEQIEHHLFPGLSHETYYHIQDVFOSTCAEYGVYPYQHEPSLWTAYWKML -500  
 FAD5 - LNEQIEHHLFPPTAPQFRFKEISPRVEALFKRHNL PY-YDLPYTS AVSTTF -421  
 FAD4 - EHLRQLGNEETHESWQRAA -519  
 FAD5 - ANLYSVGHSVGADT-KKQD -439

Fig. 3

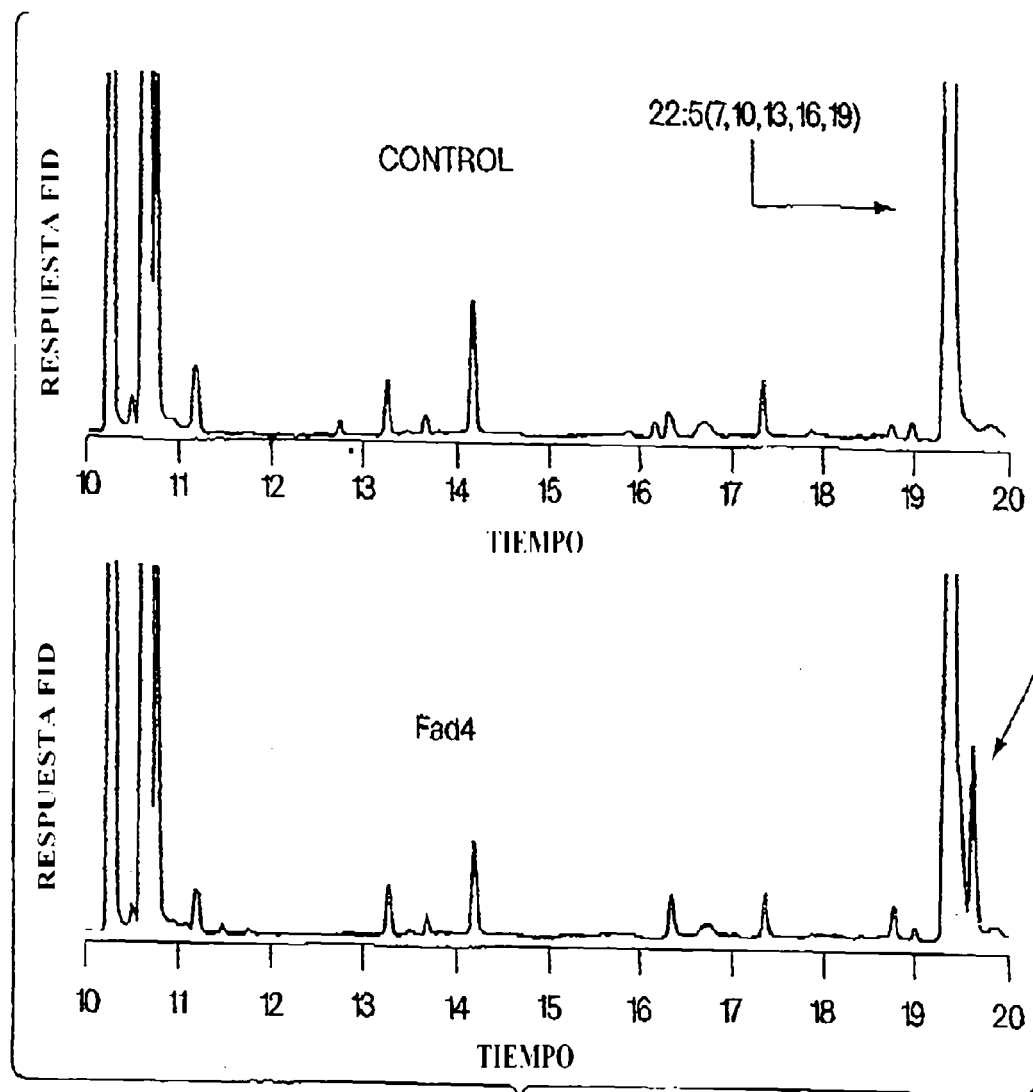


Fig. 4



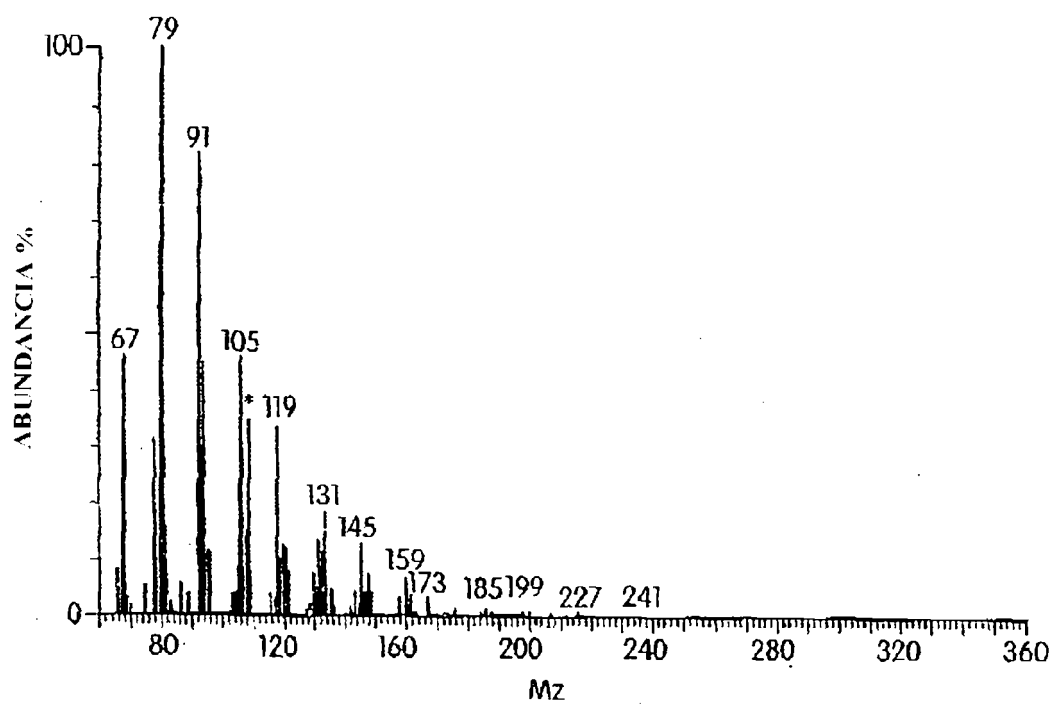


Fig. 5A

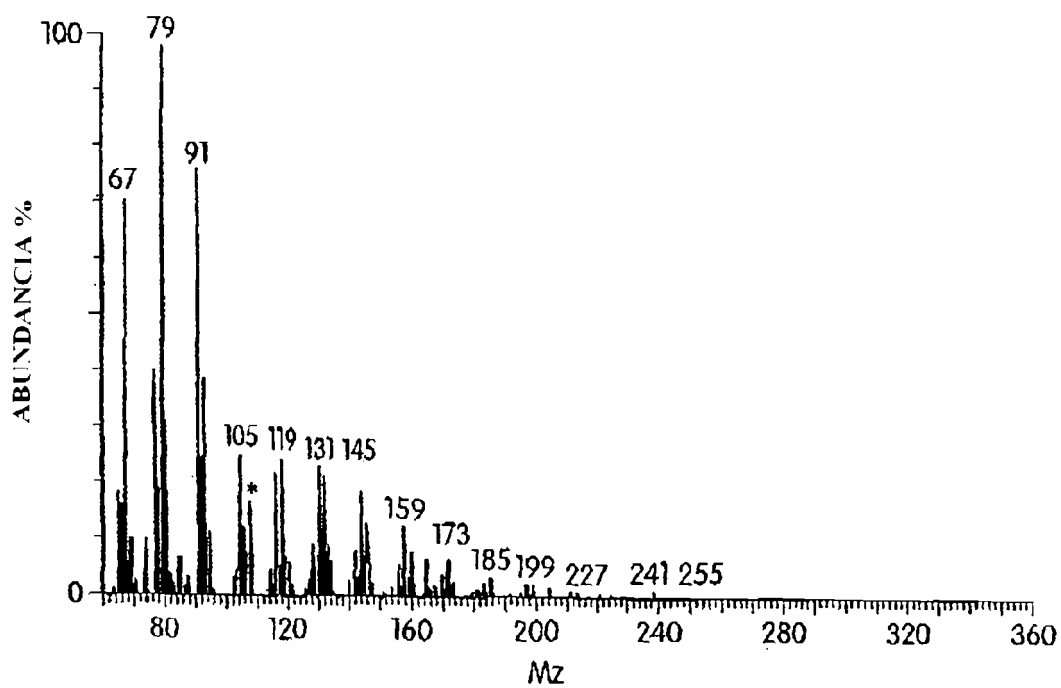


Fig. 5B

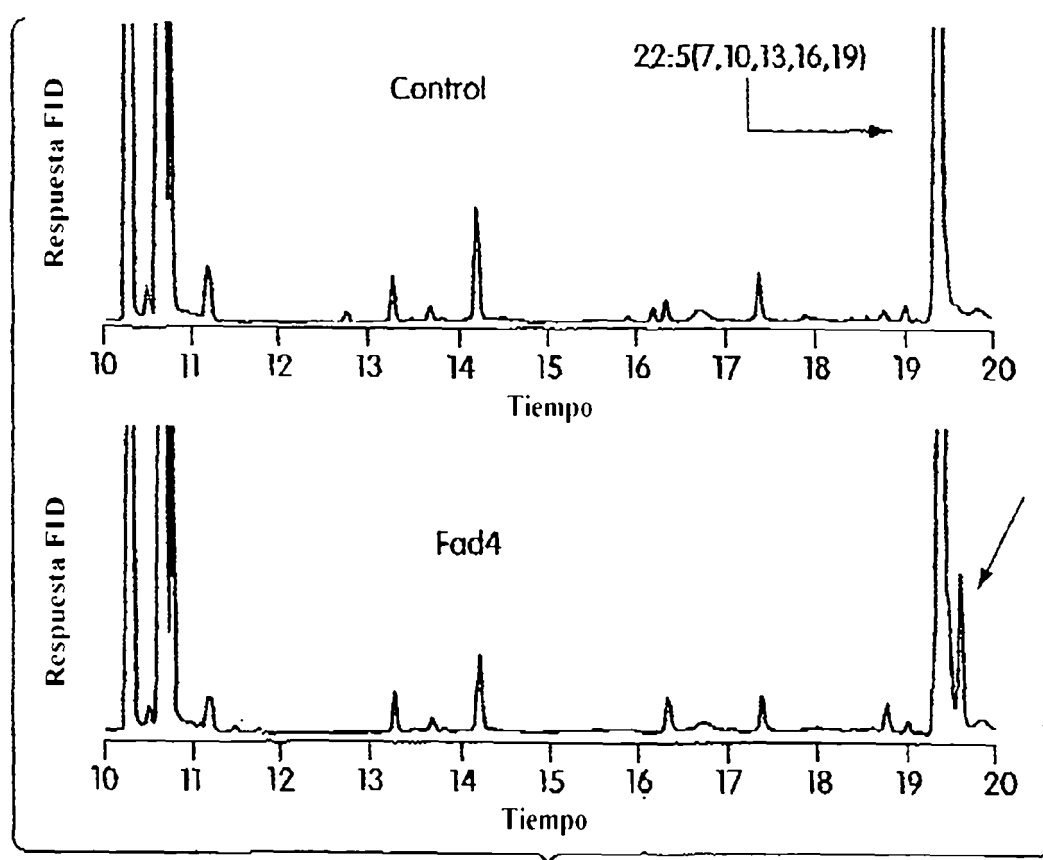


Fig. 6

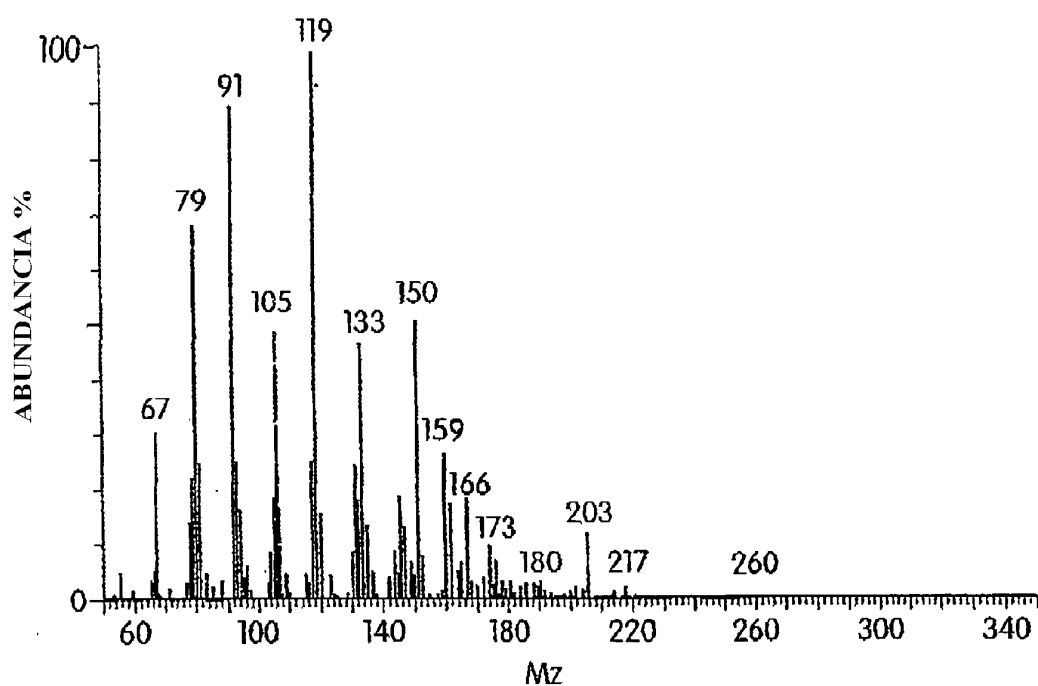


Fig. 7A

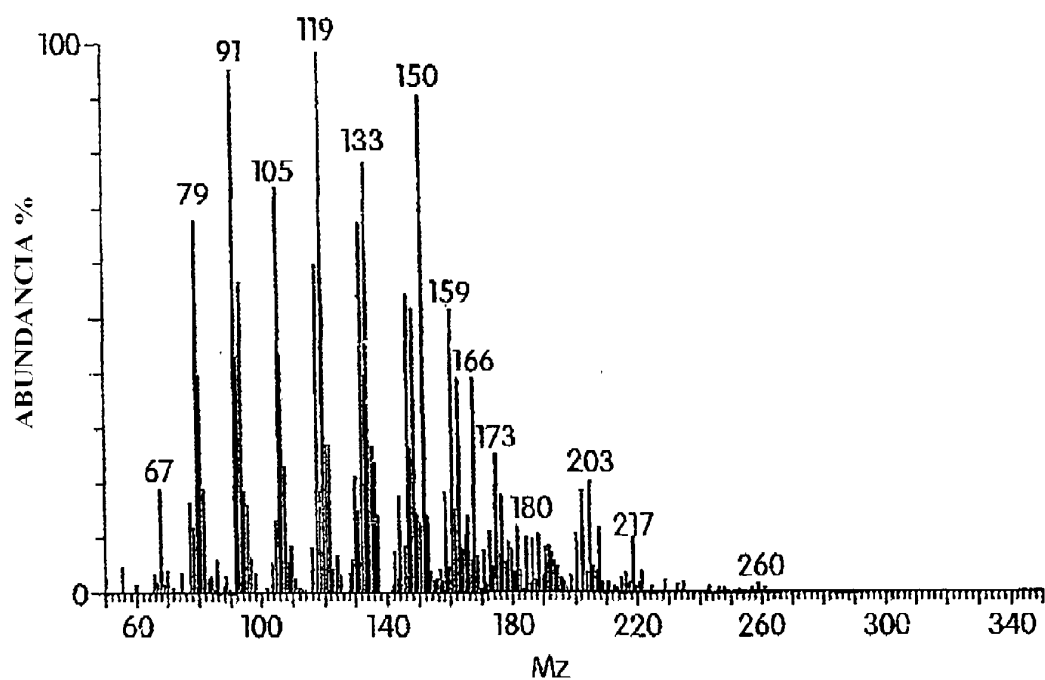


Fig. 7B

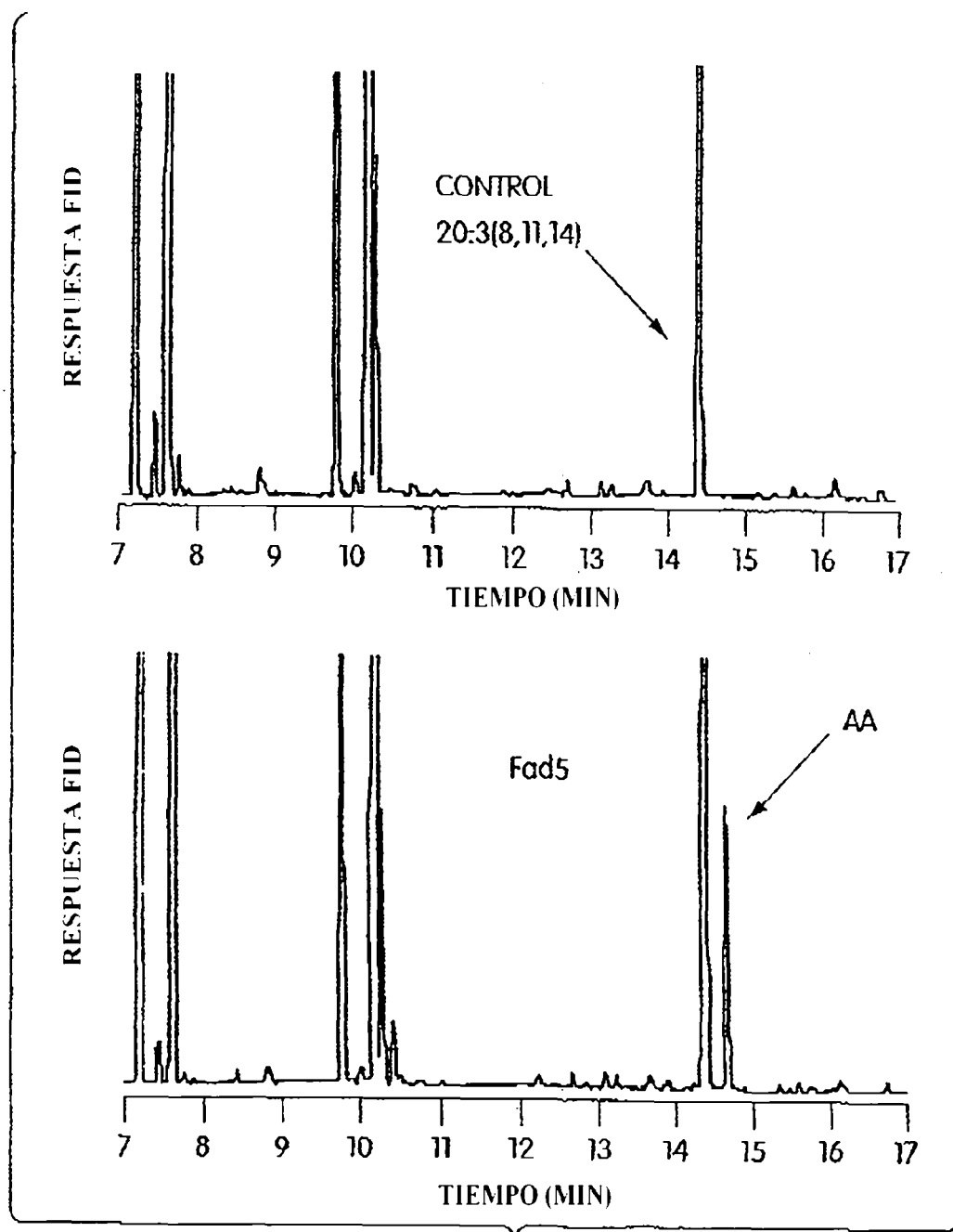


Fig. 8

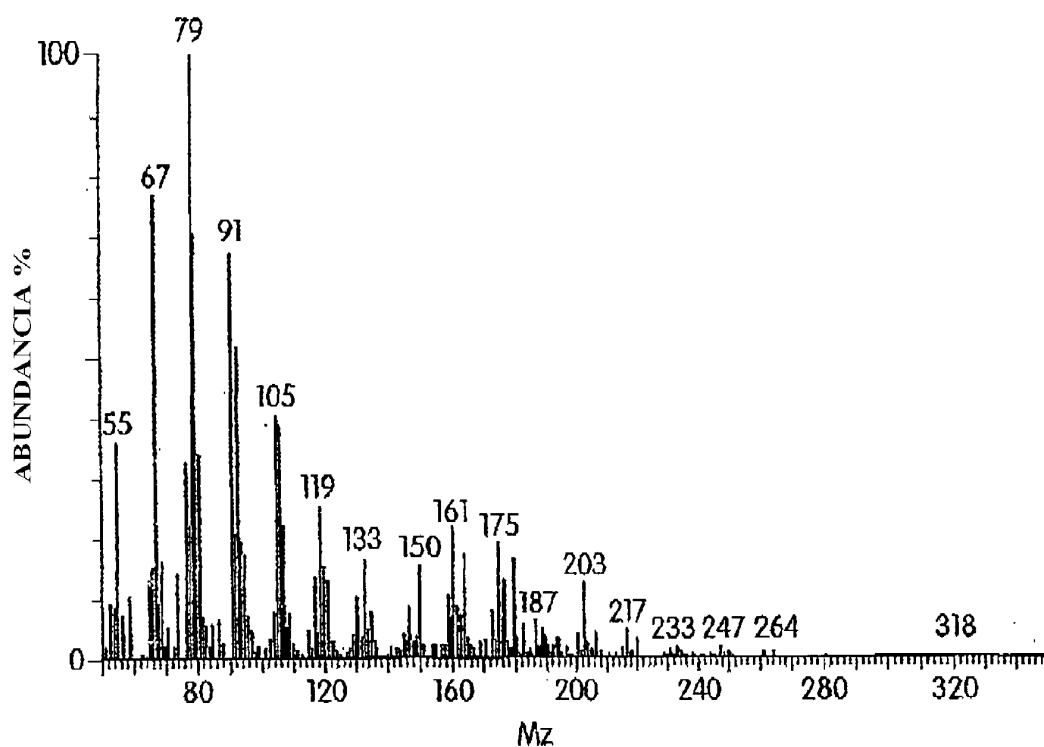


Fig. 9A

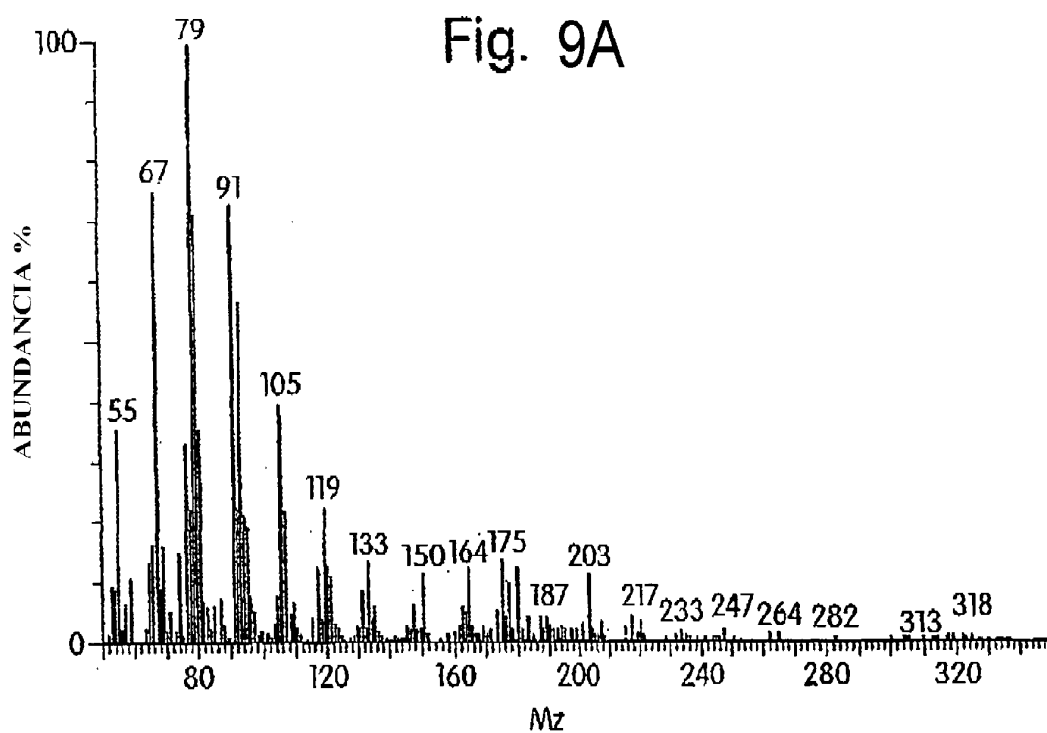


Fig. 9B

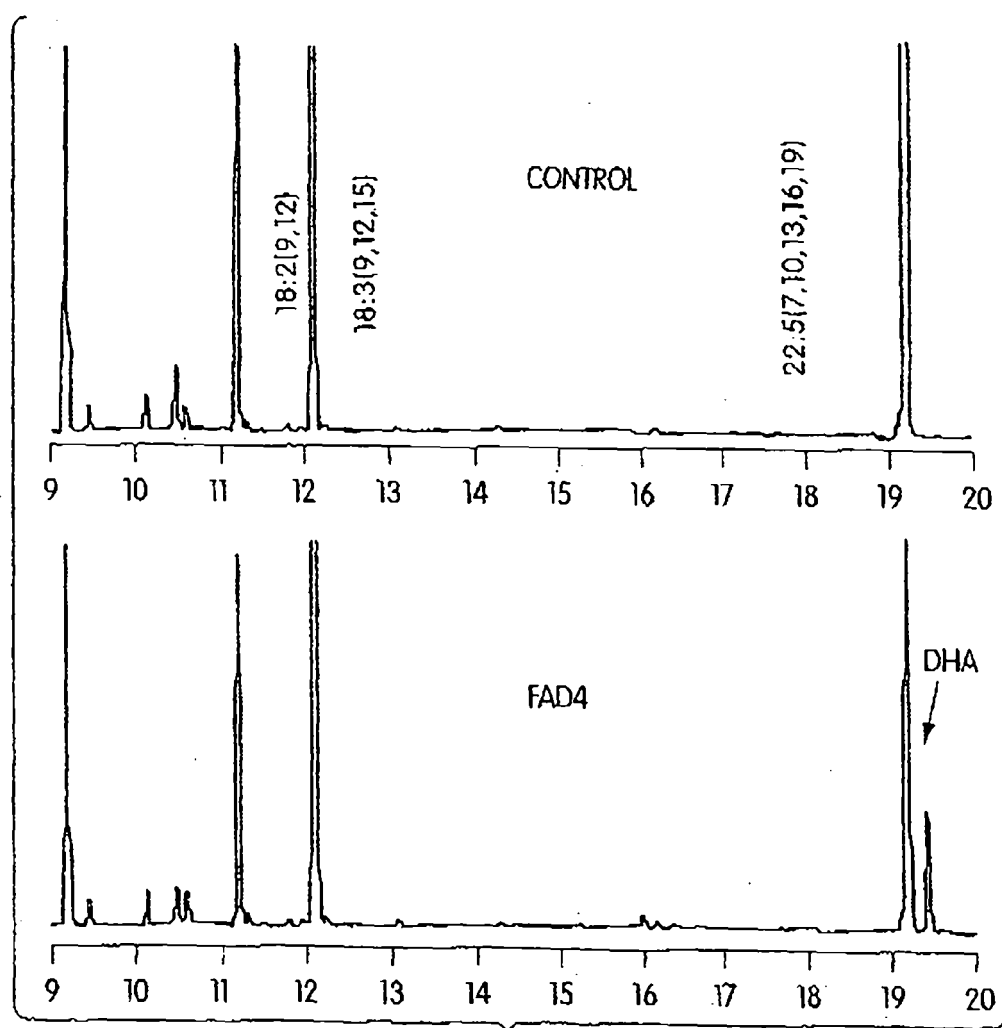


Fig. 10

Tabla 1. Perfiles de Ácido Graso de *Thraustochytrium sp.* (% Peso)

Ácidos Grasos	18:3 (n-6)	18:4 (n-3)	20:4 (n-6)	20:5 (n-3)	22:4 (n-6)	22:5 (n-6)	22:5 (n-3)	DHA (n-3)
% Peso	0.1	0.2	0.3	1.0	0.2	5.3	0.6	16.7

Fig. 11

# ES 2 343 106 T3

## LISTA DE SECUENCIAS

<110> Bioriginal Food & Science Corp.

5 <120> Fad4, Fad5, Fad5-2, Y Fad6, MIEMBROS DE LA FAMILIA DE DESATURASAS DE ÁCIDOS GRASOS Y USOS DE LOS MISMOS

<130> BNZ-001PC

<150> 60/236.303

10 <151> 2000-09-28

<150> 60/297.562

<151> 2001-06-12

<160> 8

15 <170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 1560

20 <212> ADN

<213> *Thraustochytrium* sp.

<220>

<221> CDS

25 <222> (1)...(1560)

<400> 1 .

30 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

35 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

40 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

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

50 cgg gcc 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

55 ccg gac gcc caa ggc gcc 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

60 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

65 gta atg cgc gag cgc gtc gtg gct cgg ctc aag gag cgc gcc aag gcc 384  
Val Met Arg Glu Arg Val Val Ala Arg Leu Lys Glu Arg Gly Lys Ala  
115 120 125



# ES 2 343 106 T3

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

# ES 2 343 106 T3

	gtg aac cac atc atc gag ggc gtc tcg tac gct tcc aag gac gcg gtc	1152
	Val Asn His Ile Ile Glu Gly Val Ser Tyr Ala Ser Lys Asp Ala Val	
	370 375 380	
5	aag ggc acg atg gcg ccg ccg aag acg atg cac ggc gtg acg ccc atg	1200
	Lys Gly Thr Met Ala Pro Pro Lys Thr Met His Gly Val Thr Pro Met	
	385 390 395 400	
10	aac aac acg cgc aag gag gtg gag gcg gag gcg tcc aag tct ggc gcc	1248
	Asn Asn Thr Arg Lys Glu Val Glu Ala Glu Ala Ser Lys Ser Gly Ala	
	405 410 415	
15	gtg gtc aag tca gtc ccg ctc gac gac tgg gcc gtc gtc cag tgc cag	1296
	Val Val Lys Ser Val Pro Leu Asp Trp Ala Val Val Gln Cys Gln	
	420 425 430	
20	acc tcg gtg aac tgg agc gtc ggc tcg tgg ttc tgg aat cac ttt tcc	1344
	Thr Ser Val Asn Trp Ser Val Gly Ser Trp Phe Trp Asn His Phe Ser	
	435 440 445	
25	ggc ggc ctc aac cac cag att gag cac cac ctg ttc ccc ggr ctc agc	1392
	Gly Gly Leu Asn His Gln Ile Glu His His Leu Phe Pro Xaa Leu Ser	
	450 455 460	
30	cac gag acg tac tac cac att cag gac gtc ttt cag tcc acc tgc gcc	1440
	His Glu Thr Tyr Tyr His Ile Gln Asp Val Phe Gln Ser Thr Cys Ala	
	465 470 475 480	
35	gag tac ggc gtc ccg tac cag cac gag cct tcg ctc tgg acc gcg tac	1488
	Glu Tyr Gly Val Pro Tyr Gln His Glu Pro Ser Leu Trp Thr Ala Tyr	
	485 490 495	
40	tgg aag atg ctc gag cac ctc cgt cag ctc ggc aat gag gag acc cac	1536
	Trp Lys Met Leu Glu His Leu Arg Gln Leu Gly Asn Glu Glu Thr His	
	500 505 510	
45	gag tcc tgg cag cgc gct gcc tga	1560
	Glu Ser Trp Gln Arg Ala Ala *	
	515	
50	<210> 2	
	<211> 519	
	<212> PRT	
55	<213> <i>Thraustochytrium</i> sp.	
	<220>	
	<221> VARIANT	
	<222> (1)...(519)	
60	<223> Xaa = Cualquier Aminoácido	
	<400> 2	
65	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	
70	Asp Val Thr Lys Phe Ala Ser Val His Pro Gly Gly Asp Ile Ile Leu	
	35 40 45	
75	Leu Ala Ala Gly Lys Glu Ala Thr Val Leu Tyr Glu Thr Tyr His Val	
	50 55 60	

# ES 2 343 106 T3

	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
5				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
10				115					120					125		
	Arg	Arg	Gly	Gly	Tyr	Glu	Leu	Trp	Ile	Lys	Ala	Phe	Leu	Leu	Leu	Val
							135						140			
	Gly	Phe	Trp	Ser	Ser	Leu	Tyr	Trp	Met	Cys	Thr	Leu	Asp	Pro	Ser	Phe
	145					150					155					160
15	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
				180					185					190		
20	Arg	Trp	Val	Asn	Lys	Val	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala
			195					200					205			
	Ser	Gly	Met	Thr	Trp	Glu	Phe	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr
		210					215					220				
	Thr	Asn	Leu	Ile	Glu	Glu	Asn	Gly	Leu	Gln	Lys	Val	Ser	Gly	Lys	
25	225					230				235						240
	Lys	Met	Asp	Thr	Lys	Leu	Ala	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe
					245					250					255	
	Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	His	Pro	Trp	His	Gln	Lys	Arg	Trp
				260					265					270		
30	Tyr	His	Arg	Phe	Gln	His	Ile	Tyr	Gly	Pro	Phe	Ile	Phe	Gly	Phe	Met
			275					280					285			
	Thr	Ile	Asn	Lys	Val	Val	Thr	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys
		290					295					300				
35	Arg	Leu	Phe	Gln	Ile	Asp	Ala	Glu	Cys	Arg	Tyr	Ala	Ser	Pro	Met	Tyr
	305					310					315					320
	Val	Ala	Arg	Phe	Trp	Ile	Met	Lys	Ala	Leu	Thr	Val	Leu	Tyr	Met	Val
					325					330					335	
40	Ala	Leu	Pro	Cys	Tyr	Met	Gln	Gly	Pro	Trp	His	Gly	Leu	Lys	Leu	Phe
				340					345					350		
	Ala	Ile	Ala	His	Phe	Thr	Cys	Gly	Glu	Val	Leu	Ala	Thr	Met	Phe	Ile
			355					360					365			
	Val	Asn	His	Ile	Ile	Glu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val
		370					375					380				
45	Lys	Gly	Thr	Met	Ala	Pro	Pro	Lys	Thr	Met	His	Gly	Val	Thr	Pro	Met
						390					395					400
	Asn	Asn	Thr	Arg	Lys	Glu	Val	Glu	Ala	Glu	Ala	Ser	Lys	Ser	Gly	Ala
					405					410					415	
50	Val	Val	Lys	Ser	Val	Pro	Leu	Asp	Asp	Trp	Ala	Val	Val	Gln	Cys	Gln
				420					425					430		
	Thr	Ser	Val	Asn	Trp	Ser	Val	Gly	Ser	Trp	Phe	Trp	Asn	His	Phe	Ser
			435					440					445			
55	Gly	Gly	Leu	Asn	His	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Xaa	Leu	Ser
		450					455					460				
	His	Glu	Thr	Tyr	Tyr	His	Ile	Gln	Asp	Val	Phe	Gln	Ser	Thr	Cys	Ala
	465					470					475					480
	Glu	Tyr	Gly	Val	Pro	Tyr	Gln	His	Glu	Pro	Ser	Leu	Trp	Thr	Ala	Tyr
					485					490					495	
60	Trp	Lys	Met	Leu	Glu	His	Leu	Arg	Gln	Leu	Gly	Asn	Glu	Glu	Thr	His
				500					505					510		
	Glu	Ser	Trp	Gln	Arg	Ala	Ala									
				515												

<210> 3

<211> 1320

# ES 2 343 106 T3

<212> ADN

<213> *Thraustochytrium* sp.

<220>

5 <221> CDS

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

<400> 3

```

10      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

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

20      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

25      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

30      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
        65              70              75              80

35      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
                85              90              95

40      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

45      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

50      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

55      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
        145              150              155              160

60      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

65      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

70      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

```

# ES 2 343 106 T3

	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	
5	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	
	225 230 235 240	
10	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	
15	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	
20	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	
25	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	
30	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	
	305 310 315 320	
35	ctc ggc tgc att tac att ttc ctg cag ttc gcc gtc agc cac acg cac	1008
	Leu Gly Cys Ile Tyr Ile Phe Leu Gln Phe Ala Val Ser His Thr His	
	325 330 335	
40	ctg ccg gtg acc aac ccg gag gac cag ctg cac tgg ctc gag tac gcg	1056
	Leu Pro Val Thr Asn Pro Glu Asp Gln Leu His Trp Leu Glu Tyr Ala	
	340 345 350	
45	gcc gac cac acg gtg aac att agc acc aag tcc tgg ctc gtc acg tgg	1104
	Ala Asp His Thr Val Asn Ile Ser Thr Lys Ser Trp Leu Val Thr Trp	
	355 360 365	
50	tgg atg tcg aac ctg aac ttt cag atc gag cac cac ctc ttc ccc acg	1152
	Trp Met Ser Asn Leu Asn Phe Gln Ile Glu His His Leu Phe Pro Thr	
	370 375 380	
55	gcg ccg cag ttc cgc ttc aag gaa atc agt cct cgc gtc gag gcc ctc	1200
	Ala Pro Gln Phe Arg Phe Lys Glu Ile Ser Pro Arg Val Glu Ala Leu	
	385 390 395 400	
60	ttc aag cgc cac aac ctc ccg tac tac gac ctg ccc tac acg agc gcg	1248
	Phe Lys Arg His Asn Leu Pro Tyr Tyr Asp Leu Pro Tyr Thr Ser Ala	
	405 410 415	
65	gtc tcg acc acc ttt gcc aat ctt tat tcc gtc ggc cac tcg gtc ggc	1296
	Val Ser Thr Thr Phe Ala Asn Leu Tyr Ser Val Gly His Ser Val Gly	
	420 425 430	
70	gcc gac acc aag aag cag gac tga	1320
	Ala Asp Thr Lys Lys Gln Asp *	
	435	
75	<210> 4	
80	<211> 439	
85	<212> PRT	
90	<213> <i>Thraustochytrium</i> sp.	

# ES 2 343 106 T3

<400> 4

	Met	Gly	Lys	Gly	Ser	Glu	Gly	Arg	Ser	Ala	Ala	Arg	Glu	Met	Thr	Ala	
	1				5					10					15		
5	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				
10	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	
15					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				
20	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	
						150					155					160	
25	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	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				
30	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	
						230					235					240	
35	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			
40	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	
						310					315					320	
45	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			
50	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	
							375					380					
55	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		
60	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												

65 <210> 5  
 <211> 1371  
 <212> ADN

# ES 2 343 106 T3

<213> *Thraustochytrium* sp.

<220>

<221> CDS

5 <222> (1)...(1380)

<400> 5

10	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	
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	
20	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	
25	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	
30	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	
35	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	
40	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	
45	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	
	115 120 125	
50	gtg ttc gcg gtc gcc gcg ctc gcg tac atg gcc atg aat gag ctc atc	432
	Val Phe Ala Val Ala Ala Leu Ala Tyr Met Gly Met Asn Glu Leu Ile	
	130 135 140	
55	cct gga aac gtg tac gcg cag tac gcg tgg ggc gtg gtg ttc ggt gtc	480
	Pro Gly Asn Val Tyr Ala Gln Tyr Ala Trp Gly Val Val Phe Gly Val	
	145 150 155 160	
60	ttc cag gcg ctg cca ttg ctg cac gtg atg cac gac tcg tcg cac gcg	528
	Phe Gln Ala Leu Pro Leu Leu His Val Met His Asp Ser Ser His Ala	
	165 170 175	
65	gca tgc tcg agc agc cca gcg atg tgg cag atc atc ggt cgt ggt gtg	576
	Ala Cys Ser Ser Ser Pro Ala Met Trp Gln Ile Ile Gly Arg Gly Val	
	180 185 190	

# ES 2 343 106 T3

	atg gac tgg ttc gct ggc gcc agc atg gtg tcg tgg ttg aac cag cac	624
	Met Asp Trp Phe Ala Gly Ala Ser Met Val Ser Trp Leu Asn Gln His	
	195 200 205	
5	gtt gtg ggc cac cac atc tac acg aac gtc gcg ggc gcg gac ccg gat	672
	Val Val Gly His His Ile Tyr Thr Asn Val Ala Gly Ala Asp Pro Asp	
	210 215 220	
10	ctc ccg gtc gac ttt gag agc gac gtg cgc cgc atc gtg cac cgc cag	720
	Leu Pro Val Asp Phe Glu Ser Asp Val Arg Arg Ile Val His Arg Gln	
	225 230 235 240	
15	gtg ctg ctg ccg atc tac aag ttc cag cac atc tac ctg cca ccg ctc	768
	Val Leu Leu Pro Ile Tyr Lys Phe Gln His Ile Tyr Leu Pro Pro Leu	
	245 250 255	
20	tac ggc gtg ctg ggc ctc aag ttc cgc atc cag gac gtg ttc gag acg	816
	Tyr Gly Val Leu Gly Leu Lys Phe Arg Ile Gln Asp Val Phe Glu Thr	
	260 265 270	
25	ttc gtg tcg ctc acg aac ggc ccg gtg cgt gtg aac ccg cac ccg gtg	864
	Phe Val Ser Leu Thr Asn Gly Pro Val Arg Val Asn Pro His Pro Val	
	275 280 285	
30	tcg gac tgg gtg caa atg atc ttc gcc aag gcg ttc tgg acg ttc tac	912
	Ser Asp Trp Val Gln Met Ile Phe Ala Lys Ala Phe Trp Thr Phe Tyr	
	290 295 300	
35	cgc atc tac atc ccg ttg gcg tgg ctc aag atc acg ccg tcg acg ttc	960
	Arg Ile Tyr Ile Pro Leu Ala Trp Leu Lys Ile Thr Pro Ser Thr Phe	
	305 310 315 320	
40	tgg ggc gtg ttt ttc ctc gcc gag ttc acc aca ggt tgg tac ctc gcg	1008
	Trp Gly Val Phe Phe Leu Ala Glu Phe Thr Thr Gly Trp Tyr Leu Ala	
	325 330 335	
45	ttc aac ttc cag gtg agc cac gtc tcg acc gag tgc gag tac ccg tgc	1056
	Phe Asn Phe Gln Val Ser His Val Ser Thr Glu Cys Glu Tyr Pro Cys	
	340 345 350	
50	ggt gat gcg ccg tcg gcc gag gtc ggt gac gag tgg gcg atc tcg cag	1104
	Gly Asp Ala Pro Ser Ala Glu Val Gly Asp Glu Trp Ala Ile Ser Gln	
	355 360 365	
55	gtc aag tcg tcg gtg gac tac gcg cac ggc tcg ccg ctc gcg gcg ttc	1152
	Val Lys Ser Ser Val Asp Tyr Ala His Gly Ser Pro Leu Ala Ala Phe	
	370 375 380	
60	ctc tgc ggc gcg ctc aac tac cag gtg acc cac cac ttg tac ccg ggc	1200
	Leu Cys Gly Ala Leu Asn Tyr Gln Val Thr His His Leu Tyr Pro Gly	
	385 390 395 400	
65	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	
70	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	



# ES 2 343 106 T3

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

5 aag ccc gtg gag atc cac atg ggt taa 1371  
Lys Pro Val Glu Ile His Met Gly \* \* \*  
450 455

10 <210> 6  
<211> 456  
<212> PRT  
<213> *Thraustochytrium* sp.

15 <400> 6

	Met	Thr	Glu	Lys	Ala	Ser	Asp	Glu	Phe	Thr	Trp	Gln	Glu	Val	Ala	Lys
	1				5					10					15	
20	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
25		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		
30	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
35		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	
40	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
45		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		
50	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
55				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		
60	Phe	Asn	Phe	Gln	Val	Ser	His	Val	Ser	Thr	Glu	Cys	Glu	Tyr	Pro	Cys
				340					345					350		
	Gly	Asp	Ala	Pro	Ser	Ala	Glu	Val	Gly	Asp	Glu	Trp	Ala	Ile	Ser	Gln
				355				360					365			

65

# ES 2 343 106 T3

Val Lys Ser Ser Val Asp Tyr Ala His Gly Ser Pro Leu Ala Ala Phe  
 370 375 380  
 Leu Cys Gly Ala Leu Asn Tyr Gln Val Thr His His Leu Tyr Pro Gly  
 385 390 395 400  
 5 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  
 10 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  
 15 <210> 7  
 <211> 1380  
 <212> ADN  
 20 <213> *Thraustochytrium* sp.  
 <220>  
 <221> CDS  
 <222> (1)...(1380)  
 25 <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  
 30 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  
 35 aag gtc tac gac atc tcc aag tgg gac tcg cac ccg ggt ggc tcc gtg 144  
 Lys Val Tyr Asp Ile Ser Lys Trp Asp Ser His Pro Gly Gly Ser Val  
 35 40 45  
 40 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  
 45 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  
 50 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  
 55 100 105 110  
 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  
 60 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  
 65

# ES 2 343 106 T3

	ctc	tgc	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	
5	gtc	gcc	ggc	gtg	att	atg	ggg	ctc	ttc	tac	cag	cag	tcc	gga	tgg	ctg	528
	Val	Ala	Gly	Val	Ile	Met	Gly	Leu	Phe	Tyr	Gln	Gln	Ser	Gly	Trp	Leu	
					165					170					175		
10	gcg	cac	gac	ttc	ttg	cac	aac	cag	gtg	tgc	gag	aac	cgc	acg	ctc	ggc	576
	Ala	His	Asp	Phe	Leu	His	Asn	Gln	Val	Cys	Glu	Asn	Arg	Thr	Leu	Gly	
				180					185					190			
15	aac	ctt	atc	ggc	tgc	ctc	gtg	ggc	aac	gcc	tgg	cag	ggc	ttc	agc	gtg	624
	Asn	Leu	Ile	Gly	Cys	Leu	Val	Gly	Asn	Ala	Trp	Gln	Gly	Phe	Ser	Val	
			195					200					205				
20	cag	tgg	tgg	aag	aac	aag	cac	aac	ctg	cac	cac	gcg	gtg	ccg	aac	ctg	672
	Gln	Trp	Trp	Lys	Asn	Lys	His	Asn	Leu	His	His	Ala	Val	Pro	Asn	Leu	
	210						215					220					
25	cac	agc	gcc	aag	gac	gag	ggc	ttc	atc	ggc	gac	ccg	gac	atc	gac	acc	720
	His	Ser	Ala	Lys	Asp	Glu	Gly	Phe	Ile	Gly	Asp	Pro	Asp	Ile	Asp	Thr	
	225					230					235					240	
30	atg	ccg	ctg	ctg	gcg	tgg	tct	aag	gag	atg	gcg	cgc	aag	gcg	ttc	gag	768
	Met	Pro	Leu	Leu	Ala	Trp	Ser	Lys	Glu	Met	Ala	Arg	Lys	Ala	Phe	Glu	
					245					250					255		
35	tgc	gcg	cac	ggc	ccg	ttc	ttc	atc	cgc	aac	cag	gcg	ttc	cta	tac	ttc	816
	Ser	Ala	His	Gly	Pro	Phe	Phe	Ile	Arg	Asn	Gln	Ala	Phe	Leu	Tyr	Phe	
				260					265					270			
40	ccg	ctg	ctg	ctg	ctc	gcg	cgc	ctg	agc	tgg	ctc	gcg	cag	tgc	ttc	ttc	864
	Pro	Leu	Leu	Leu	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Ala	Gln	Ser	Phe	Phe	
			275					280					285				
45	tac	gtg	ttc	acc	gag	ttc	tgc	ttc	ggc	atc	ttc	gac	aag	gtc	gag	ttc	912
	Tyr	Val	Phe	Thr	Glu	Phe	Ser	Phe	Gly	Ile	Phe	Asp	Lys	Val	Glu	Phe	
	290						295					300					
50	gac	gga	ccg	gag	aag	gcg	ggt	ctg	atc	gtg	cac	tac	atc	tgg	cag	ctc	960
	Asp	Gly	Pro	Glu	Lys	Ala	Gly	Leu	Ile	Val	His	Tyr	Ile	Trp	Gln	Leu	
	305					310					315					320	
55	gcg	atc	ccg	tac	ttc	tgc	aac	atg	agc	ctg	ttt	gag	ggc	gtg	gca	tac	1008
	Ala	Ile	Pro	Tyr	Phe	Cys	Asn	Met	Ser	Leu	Phe	Glu	Gly	Val	Ala	Tyr	
					325					330					335		
60	ttc	ctc	atg	ggc	cag	gcg	tcc	tgc	ggc	ttg	ctc	ctg	gcg	ctg	gtg	ttc	1056
	Phe	Leu	Met	Gly	Gln	Ala	Ser	Cys	Gly	Leu	Leu	Leu	Ala	Leu	Val	Phe	
				340					345					350			
65	agt	att	ggc	cac	aac	ggc	atg	tgc	gtg	tac	gag	cgc	gaa	acc	aag	ccg	1104
	Ser	Ile	Gly	His	Asn	Gly	Met	Ser	Val	Tyr	Glu	Arg	Glu	Thr	Lys	Pro	
			355					360					365				
70	gac	ttc	tgg	cag	ctg	cag	gtg	acc	acg	acg	cgc	aac	atc	cgc	gcg	tgc	1152
	Asp	Phe	Trp	Gln	Leu	Gln	Val	Thr	Thr	Thr	Arg	Asn	Ile	Arg	Ala	Ser	
			370				375					380					

# ES 2 343 106 T3

```

gta ttc atg gac tgg ttc acc ggt ggc ttg aac tac cag atc gac cat 1200
Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Asp His
385          390          395          400

5   cac ctg ttc ccg ctc gtg ccg cgc cac aac ttg cca aag gtc aac gtg 1248
His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val
          405          410          415

10  ctc atc aag tcg cta tgc aag gag ttc gac atc ccg ttc cac gag acc 1296
Leu Ile Lys Ser Leu Cys Lys Glu Phe Asp Ile Pro Phe His Glu Thr
          420          425          430

15  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

20  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
<212> PRT
25  <213> Thraustochytrium sp.

<400> 8

30  Met Val Asp Leu Lys Pro Gly Val Lys Arg Leu Val Ser Trp Lys Glu
    1      5      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

```

# ES 2 343 106 T3

	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				
5	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	
10	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				
15	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
20					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				
25	Ser	Lys	Glu	Phe	Ile	Thr	Glu	Phe	Pro	Ala	Met					
		450					455									

30

35

40

45

50

55

60

65