

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 623 610**

51 Int. Cl.:

C12N 15/53 (2006.01)

C12N 15/82 (2006.01)

C12P 7/64 (2006.01)

C12N 9/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.03.2010 PCT/US2010/028009**

87 Fecha y número de publicación internacional: **23.09.2010 WO10108114**

96 Fecha de presentación y número de la solicitud europea: **19.03.2010 E 10754185 (6)**

97 Fecha y número de publicación de la concesión europea: **15.03.2017 EP 2408797**

54 Título: **Moléculas de ácido nucleico de ácido graso poliinsaturado sintasas, y polipéptidos, composiciones y métodos de preparación y usos de los mismos**

30 Prioridad:

19.03.2009 US 161742 P

19.01.2010 US 296460 P

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

11.07.2017

73 Titular/es:

DSM IP ASSETS B.V. (100.0%)

Het Overloon 1

6411 TE Heerlen , NL

72 Inventor/es:

APT, KIRK, E.;

RICHTER, LESLIE;

SIMPSON, DAVID y

ZIRKLE, ROSS

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 623 610 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Moléculas de ácido nucleico de ácido graso poliinsaturado sintasas, y polipéptidos, composiciones y métodos de preparación y usos de los mismos

5

Antecedentes de la invención

Campo de la invención

10 La presente invención se refiere a moléculas de ácido nucleico y polipéptidos de ácido graso poliinsaturado sintasas (PUFA) aislados implicados en la producción de los PUFA, incluyendo los PUFA enriquecidos con ácido docosahexaenoico (DHA), ácido eicosapentaenoico (EPA), o una combinación de los mismos. La presente invención se refiere a vectores y células hospedadoras que comprenden las moléculas de ácido nucleico, polipéptidos

15 codificados por las moléculas de ácido nucleico, composiciones que comprenden las moléculas de ácido nucleico o polipéptidos, y métodos de preparación y usos de los mismos.

Antecedentes de la invención

20 Los traustóquitridos son microorganismos del orden Traustóquitriales, que incluye miembros del género *Thraustochytrium* y el género *Schizochytrium*, y se han reconocido como una fuente alternativa de los PUFA. Véase, por ejemplo, documento de Patente de Estados Unidos N.º 5.130.242. Recientemente se ha mostrado que los sistemas similares a los de la poliquétido sintasa (PKS) en bacterias marinas y traustóquitridos son capaces de sintetizar ácidos grasos poliinsaturados (PUFA) a partir de acetil-CoA y malonil-CoA. En el presente documento estos sistemas similares a los de la PKS sintasa también se denominan sistemas de PUFA sintasa. Los sistemas de

25 PUFA sintasa en las bacterias marinas *Shewanella* y *Vibrio marinus* se describen en el documento de Patente de Estados Unidos N.º 6.140.486. Un sistema de PUFA sintasa en un traustóquitrido del género *Schizochytrium* se describe en el documento de Patente de Estados Unidos N.º 6.566.583. Los sistemas de PUFA sintasa en traustóquitridos del género *Schizochytrium* y del género *Thraustochytrium* también se describen en el documento de Patente de Estados Unidos N.º 7.247.461. El documento de Patente de Estados Unidos N.º 7.211.418 describe un sistema de PUFA sintasa en un traustóquitrido del género *Thraustochytrium* y la producción de ácido eicosapentaenoico (C20: 5, omega-3) (EPA) y otros PUFA que usan el sistema. El documento de Patente de Estados Unidos N.º 7.217.856 describe los sistemas de PUFA sintasa en *Shewanella olleyana* y *Shewanella japonica*. El documento WO 2005/097982 describe un sistema de PUFA sintasa en la cepa SAM2179. Los documentos de Patente de Estados Unidos N.ºs 7.208.590 y 7.368.552 describen genes y proteínas de PUFA

35 sintasa de *Thraustochytrium aureum*.

En la bibliografía tradicionalmente se ha descrito que los sistemas de PKS entran en uno de tres tipos básicos, denominados por lo general Tipo I (modular o repetitivo), Tipo II, y Tipo III. El sistema de PKS modular de Tipo I también se ha denominado sistema de PKS "modular", y el sistema de PKS repetitivo de Tipo I también se ha

40 denominado sistema de PKS de "Tipo I". El sistema de Tipo II se caracteriza por proteínas separables, cada una de las cuales realiza una reacción enzimática distinta. Las enzimas funcionan de manera sincronizada para producir el producto final y cada enzima individual del sistema por lo general participa varias veces en la producción del producto final. Este tipo de sistema funciona de manera análoga a los sistemas de ácido graso sintasa (FAS) encontrados en plantas y bacterias. Los sistemas de PKS repetitivos de Tipo I son similares al sistema de Tipo II en

45 que las enzimas se usan de una manera repetitiva para producir el producto final. El sistema repetitivo de Tipo I se diferencia del sistema de Tipo II en que las actividades enzimáticas, en lugar de estar asociadas con proteínas separables, se producen como dominios de proteínas de mayor tamaño. Este sistema es análogo a los sistemas de FAS de Tipo I encontrados en animales y hongos.

50 A diferencia de los sistemas de Tipo II, cada dominio enzimático de los sistemas de PKS modulares de Tipo I se usa solamente una vez en la producción del producto final. Los dominios se encuentran en muchas proteínas grandes y el producto de cada reacción se transmite a otro dominio en la proteína PKS.

Más recientemente se han descubierto sistemas de Tipo III y pertenecen a la familia de la chalcona sintasa vegetal de enzimas de condensación. Las PKS de Tipo III son distintas de los sistemas de PKS de Tipo I y de Tipo II y utilizan sustratos de CoA libre en reacciones de condensación repetitivas para producir normalmente un producto final heterocíclico.

60 En la ruta convencional o estándar para la síntesis de PUFA, los ácidos grasos saturados con una cadena de longitud media (productos de un sistema de ácido graso sintasa (FAS)) se modifican mediante una serie de reacciones de elongación y desaturación. Los sustratos para la reacción de elongación son acil-CoA graso (la cadena de ácido graso a alargar) y malonil-CoA (la fuente de los dos carbonos añadidos durante la reacción de elongación). El producto de la reacción de elongasa es una acil-CoA graso que tiene dos carbonos adicionales en la cadena lineal. Las desaturasas crean dobles enlaces cis en la cadena de ácido graso existente previamente

65 mediante extracción de dos hidrógenos en una reacción dependiente de oxígeno. Los sustratos para las desaturasas son cualquiera de acil-CoA (en algunos animales) o el ácido graso que se esterifica a la cadena principal de glicerol

de un fosfolípido (por ejemplo, fosfatidilcolina).

Los ácidos grasos se clasifican basándose en la longitud y características de saturación de la cadena del carbono. Los ácidos grasos se denominan ácidos grasos de cadena corta, cadena media o cadena larga basándose en el número de carbonos presentes en la cadena, se denominan ácidos grasos saturados cuando no hay presencia de dobles enlaces entre los átomos de carbono, y se denominan ácidos grasos insaturados cuando hay presencia de dobles enlaces. Los ácidos grasos insaturados de cadena larga son monoinsaturados cuando está presente solo un doble enlace y son poliinsaturados cuando hay presencia de más de un doble enlace.

Los PUFA se clasifican basándose en la posición del primer doble enlace desde el extremo metilo del ácido graso: los ácidos grasos omega-3 (n-3) contienen un primer doble enlace en el tercer carbono, mientras que los ácidos grasos omega-6 (n-6) contienen un primer doble enlace en el sexto carbono. Por ejemplo, el ácido docosahexaenoico ("DHA") es un PUFA omega-3 con una longitud de la cadena de 22 carbonos y 6 dobles enlaces, a menudo denominado "22: 6 n-3". Otros PUFA omega-3 incluyen ácido eicosapentaenoico ("EPA"), denominado "20: 5 n-3", y ácido docosapentaenoico omega-3 ("DPA n-3"), denominado "22: 5 n-3". DHA y EPA se han denominado ácidos grasos "esenciales". Los PUFA omega-6 incluyen ácido araquidónico ("ARA"), denominado "20: 4 n-6", y ácido docosapentaenoico omega-6 ("DPA n-6"), denominado "22: 5 n-6".

Los ácidos grasos omega-3 son moléculas biológicamente importantes que influyen en la fisiología celular debido a su presencia en membranas celulares, regulan la producción y la expresión genética de compuestos biológicamente activos, y sirven como sustratos biosintéticos. Roche, H. M., *Proc. Nutr. Soc.* 58: 397-401 (1999). Por ejemplo, el DHA representa aproximadamente un 15 %-20 % de lípidos en la corteza cerebral humana, y un 30 %-60 % de lípidos en la retina, se concentra en los testículos y el esperma, y es un componente importante de la leche de mama. Bergé, J.P., y Barnathan, G. *Adv. Biochem. Eng. Biotechnol.* 96: 49-125 (2005). El DHA representa hasta un 97 % de los ácidos grasos omega-3 en el cerebro y hasta un 93 % de los ácidos grasos omega-3 en la retina. Además, el DHA es esencial para el desarrollo tanto fetal como infantil, así como para el mantenimiento de las funciones cognitivas en adultos. Del mismo modo, dado que los ácidos grasos omega-3 no se sintetizan *de novo* en el cuerpo humano, estos ácidos grasos se deben obtener a partir de fuentes nutricionales.

El aceite de semillas de lino y los aceites de pescado se consideran buenas fuentes dietéticas de ácidos grasos omega-3. El aceite de lino no contiene EPA, DHA, DPA, o ARA, pero es un lugar contiene ácido linolénico (C18: 3 n-3), un componente básico que permite que el cuerpo fabrique EPA. Sin embargo, existen evidencias de que la tasa de conversión metabólica puede ser lenta y variable, en particular entre las personas con problemas de salud. Los aceites de pescado varían considerablemente en el tipo y nivel de composición de ácidos grasos dependiendo de la especie en particular y sus dietas. Por ejemplo, los peces criados mediante acuicultura tienden a tener un nivel más bajo de ácidos grasos omega-3 que los que están en el medio silvestre. Además, los aceites de pescado conllevan el riesgo de contener contaminantes ambientales y pueden estar asociados con problemas de estabilidad y un olor o sabor a pescado.

Los aceites producidos a partir de traustóquitridos a menudo tienen perfiles de ácido graso poliinsaturado más sencillos que los correspondientes aceites de pescado o microalgas. Lewis, T.E., *Mar. Biotechnol.* 1: 580-587 (1999). Se ha informado que las cepas de especies de traustóquitridos producen ácidos grasos omega-3 como un alto porcentaje de los ácidos grasos totales producidos por los organismos. Documento de Patente de Estados Unidos N.º 5.130.242; Huang, J. *et al.*, *J. Am. Oil. Chem. Soc.* 78: 605-610 (2001); Huang, J. *et al.*, *Mar. Biotechnol.* 5: 450-457 (2003). Sin embargo, los traustóquitridos aislados varían en la identidad y cantidades de los PUFA producidos, de modo que algunas cepas descritas previamente pueden tener perfiles de PUFA indeseables.

Se han realizado esfuerzos para producir los PUFA y plantas de cultivo de semillas oleaginosas mediante modificación de los ácidos grasos producidos de forma endógena. La modificación genética de estas plantas con diversos genes individuales para elongasas y desaturasas de ácidos grasos ha producido hojas o semillas que contienen niveles que se pueden medir de los PUFA tales como EPA, pero que también contienen niveles significativos de PUFA mixtos de cadena más corta y menos insaturados (Qi *et al.*, *Nature Biotech.* 22: 739 (2004); Publicación de PCT N.º WO 04/071467; Abbadi *et al.*, *Plant Cell* 16: 1 (2004)); Napier y Sayanova, *Proc. Nutrition Society* 64: 387-393 (2005); Robert *et al.*, *Functional Plant Biology* 32: 473-479 (2005); y Solicitud de Publicación de Estados Unidos N.º 2004/0172682).

El documento US2008/044871 desvela sistemas de PUFA polipéptido sintasa y usos de los mismos.

Como tal, existe una necesidad continua para el aislamiento de moléculas de ácido nucleico y polipéptidos asociados con perfiles de PUFA deseables y métodos para producir perfiles de PUFA deseables a través del uso de tales moléculas de ácido nucleico y polipéptidos.

Breve resumen de la invención

La presente invención se refiere a una molécula de ácido nucleico aislada seleccionada entre el grupo que consiste en: (a) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica

a la SEQ ID NO: 5, en la que la secuencia de polinucleótidos codifica un polipéptido que tiene actividad de ácido graso poliinsaturado (PUFA) sintasa que consiste en actividad de β -hidroxiacil-ACP deshidrasa similar a FabA (DH) y actividad de enoil-ACP reductasa (ER); (b) una molécula de ácido nucleico que comprende un polinucleótido que tiene la secuencia que se establece en la SEQ ID NO: 5; (c) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifican un polipéptido, en la que el polipéptido comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER; y (d) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifican un polipéptido que comprende la secuencia de aminoácidos como se establece en la SEQ ID NO: 6.

La presente invención se refiere a una molécula de ácido nucleico aislada que comprende una secuencia de polinucleótidos que es totalmente complementaria a cualquiera de las secuencias de polinucleótidos que se han descrito anteriormente.

La presente invención se refiere a una molécula de ácido nucleico recombinante que comprende cualquiera de las moléculas de ácido nucleico descritas anteriormente o combinaciones de las mismas y una secuencia de control de la transcripción. En algunas realizaciones, la molécula de ácido nucleico recombinante es un vector recombinante.

La presente invención se refiere a una célula hospedadora que expresa cualquiera de las moléculas de ácido nucleicos descritas anteriormente, cualquiera de las moléculas de ácido nucleico recombinante descritas anteriormente, y combinaciones de las mismas. En algunas realizaciones, la célula hospedadora se selecciona entre el grupo que consiste en una célula vegetal, una célula microbiana, y una célula animal. En algunas realizaciones, la célula microbiana es una bacteria. En algunas realizaciones, la bacteria es *E. coli*. En algunas realizaciones, la bacteria es una bacteria marina. En algunas realizaciones, la célula microbiana es un traustóquitrido. En algunas realizaciones, el traustóquitrido es un *Schizochytrium*. En algunas realizaciones, el traustóquitrido es un *Ulkenia*.

La presente invención se refiere a un método para producir al menos un PUFA, que comprende: expresar un gen de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir PUFA, en el que el gen de PUFA sintasa comprende cualquiera la molécula de ácido nucleico aislada como se ha descrito anteriormente, cualquiera de las moléculas de ácido nucleico recombinante descritas anteriormente, o combinaciones de las mismas, y en el que se produce al menos un PUFA. En un aspecto de esta realización, la célula hospedadora se selecciona entre el grupo que consiste en una célula vegetal, una célula animal aislada, y una célula microbiana. En otro aspecto de esta realización, el al menos un PUFA comprende ácido docosahexaenoico (DHA) o ácido eicosapentaenoico (EPA).

La presente invención se refiere a un método para producir lípidos enriquecidos con DHA, EPA, o una combinación de los mismos, que comprende: expresar un gen de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir lípidos, en el que el gen de PUFA sintasa comprende cualquiera de la molécula de ácido nucleico aislada como se ha descrito anteriormente, cualquiera de la molécula de ácido nucleico recombinante descritas anteriormente, o combinaciones de las mismas en la célula hospedadora, y en el que se producen lípidos enriquecidos con DHA, EPA, o una combinación de los mismos. La presente invención se refiere a un método para preparar un vector recombinante que comprende insertar una cualquiera de la molécula de ácido nucleico aislada como se ha descrito anteriormente en un vector.

La presente invención se refiere a un método para preparar una célula hospedadora recombinante que comprende introducir un vector recombinante como se ha descrito anteriormente en una célula hospedadora. En algunas realizaciones, la célula hospedadora se selecciona entre el grupo que consiste en una célula vegetal, una célula animal aislada, y una célula microbiana.

La presente invención se refiere a un polipéptido aislado codificado por cualquiera de las secuencias de polinucleótidos que se han descrito anteriormente.

La presente invención también proporciona un polipéptido aislado seleccionado entre el grupo que consiste en: (a) un polipéptido que comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de ácido graso poliinsaturado (PUFA) sintasa que consiste en actividad de β -hidroxiacil-ACP deshidrasa similar a FabA (DH) y actividad de enoil-ACP reductasa (ER); y (b) un polipéptido que comprende una secuencia de aminoácidos como se establece en la SEQ ID NO: 6.

En algunas realizaciones, las secuencias de aminoácidos son al menos un 90 % idénticas o al menos un 95 % idénticas a la SEQ ID NO: 6 respectivamente. En algunas realizaciones, los polipéptidos comprenden la secuencia de aminoácidos de SEQ ID NO: 6.

En algunas realizaciones, cualquiera de los polipéptidos aislados de la invención puede ser un polipéptido de fusión.

La presente invención se refiere a una composición que comprende cualquiera de los polipéptidos descritos anteriormente un vehículo biológicamente aceptable.

La presente invención se refiere a un método para aumentar la producción de DHA en un organismo que tiene actividad de ácido graso poliinsaturado (PUFA) sintasa, que comprende: expresar la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas, en el organismo en condiciones eficaces para producir DHA, y en el que la producción de DHA en organismo aumenta.

La presente invención se refiere a un método para aislar lípidos a partir de una célula hospedadora, que comprende: (a) expresar un gen de ácido graso poliinsaturado (PUFA) sintasa en la célula hospedadora en condiciones eficaces para producir lípidos, en el que el sistema de poliquétido sintasa (PKS) comprende la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas, en la célula hospedadora, y (b) aislar lípidos a partir de la célula hospedadora.

La presente invención se define mediante las reivindicaciones. La materia objeto fuera del alcance de las reivindicaciones se proporciona solamente para información.

Breve descripción de los dibujos/figuras

FIG. 1 muestra la arquitectura genética de las PUFA sintasas de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, de la invención.

FIG. 2 muestra la arquitectura genética de las PUFA sintasas de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* de la invención.

FIG. 3 muestra la arquitectura de dominio de las PUFA sintasas de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, y las PUFA sintasas de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, de la invención y sintasas de la cepa 20888 de la ATCC de *Schizochytrium sp.*, la cepa 20892 de la ATCC de *Thraustochytrium sp.*, *Thraustochytrium aureum*, y SAM2179.

FIG. 4 muestra un alineamiento de una secuencia de aminoácidos de Pfa1p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 2) y una secuencia de aminoácidos de Pfa1p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 69) de la invención con las secuencias de OrfA de la cepa 20888 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 54) y de la cepa 20892 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 56) y la secuencia ORF A de *Thraustochytrium aureum* (SEQ ID NO: 55).

FIG. 5 muestra un alineamiento de una secuencia de aminoácidos de Pfa2p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 4) y una secuencia de aminoácidos de Pfa2p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 71) de la invención con las secuencias de OrfB de la cepa 20888 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 57) y de la cepa 20892 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 58) y la secuencia ORF B de *Thraustochytrium aureum* (SEQ ID NO: 59).

FIG. 6 muestra un alineamiento de una secuencia de aminoácidos de Pfa3p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 6) y una secuencia de aminoácidos de Pfa3p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 73) de la invención con las secuencias de OrfC de la cepa 20888 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 61) y de la cepa 20892 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 60).

FIG. 7 muestra la secuencia de polinucleótidos de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 1).

FIG. 8 muestra la secuencia de aminoácidos de Pfa1p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 2).

FIG. 9 muestra la secuencia de polinucleótidos de *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 3).

FIG. 10 muestra la secuencia de aminoácidos de Pfa2p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 4).

FIG. 11 muestra la secuencia de polinucleótidos de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5).

FIG. 12 muestra la secuencia de aminoácidos de Pfa3p de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 6).

FIG. 13 muestra la secuencia de polinucleótidos de *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 68).

FIG. 14 muestra una secuencia de polinucleótidos de *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 120) que se ha optimizado con codón para su expresión en *Schizochytrium*.

FIG. 15 muestra la secuencia de aminoácidos de Pfa1p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 69).

FIG. 16 muestra la secuencia de polinucleótidos de *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 70).

FIG. 17 muestra una secuencia de polinucleótidos de *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 121) que se ha optimizado con codón para su expresión en *Schizochytrium*.

FIG. 18 muestra la secuencia de aminoácidos de Pfa2p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 71).

FIG. 19 muestra la secuencia de polinucleótidos de *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 72).

FIG. 20 muestra una secuencia de polinucleótidos de *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. (SEQ ID NO: 122) que se ha optimizado con codón para su expresión en *Schizochytrium*.

FIG. 21 muestra la secuencia de aminoácidos de Pfa3p de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. (SEQ ID NO: 73).

FIG. 22 muestra una tabla de uso de codón para *Schizochytrium*.

Descripción detallada de la invención

La presente invención se refiere a moléculas de ácido nucleico y polipéptidos de ácido graso poliinsaturado sintasas (PUFA) aislados implicados en la producción de los PUFA, incluyendo los PUFA enriquecidos con ácido docosahexaenoico (DHA), ácido eicosapentaenoico (EPA), o una combinación de los mismos. La presente invención se refiere a vectores y células hospedadoras que comprenden las moléculas de ácido nucleico, polipéptidos codificados por las moléculas de ácido nucleico, composiciones que comprenden las moléculas de ácido nucleico o polipéptidos, y métodos de preparación y usos de los mismos.

PUFA Sintetas

Como se usa en el presente documento, la expresión "PUFA sintasa" se refiere a una enzima que está implicada en la producción de ácidos grasos poliinsaturados. Véase, por ejemplo, Metz *et al.*, Science 293: 290-293 (2001).

La presente invención se dirige en parte a una subunidad de PUFA sintasa denominada Pfa3p (SEQ ID NO: 6), así como al gen que codifica las subunidades denominadas *PFA3* (SEQ ID NO: 5). Véanse, FIGs. 1-3 y 7-21. Las PUFA sintasas en otros traustochitridos también se han designado como ORF 1, ORF 2, y ORF 3, respectivamente, o como OrfA, OrfB, y OrfC, respectivamente. Véase, por ejemplo, *Schizochytrium* sp., (ATCC 20888) y *Thraustochytrium* sp. (ATCC 20892) en los documentos de patente de Estados Unidos N.º 7.247.461 y 7.256.022, que hacen referencia a genes *orfA*, *orfB*, y *orfC* y a las correspondientes proteínas OrfA, OrfB, y OrfC, y *Thraustochytrium aureum* (ATCC 34304) en el documento de patente de Estados Unidos N.º 7.368.552, que hacen referencia a genes y proteínas ORF A, ORF B, y ORF C. Véase también, la cepa SAM2179 en el documento WO/2005/097982, que hace referencia a genes y proteínas ORF 1, ORF 2, y ORF 3.

Moléculas de Ácido Nucleico

La presente invención se refiere a moléculas de ácido nucleico aislados que comprenden secuencias de polinucleótidos para genes y dominios de PUFA sintasa obtenidas a partir de un microorganismo aislado que es el objeto de la Solicitud de Estados Unidos N.º 12/407.687 en trámite junto con la presente, presentada el 19 de marzo de 2009, incorporada en el presente documento por referencia en su totalidad. El microorganismo se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 7 de enero de 2009, y se le dio el N.º PTA-9695 de Registro en la ATCC, y también se denomina cepa PTA-9695 de la ATCC de *Schizochytrium* sp.. Cuando se expresan, estos genes producen perfiles de ácido graso únicos, caracterizados en parte por niveles elevados de ácidos grasos omega-3, en particular niveles elevados de DHA.

La presente invención se refiere a moléculas de ácido nucleico aisladas que comprenden secuencias de polinucleótidos para genes y dominios de PUFA sintasa obtenidas a partir de un microorganismo aislado que es el objeto de la Solicitud de Estados Unidos N.º 61/296.456, en trámite junto con la presente, presentada el 19 de enero de 2010. El microorganismo se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 14 de julio de 2009, y se le dio el N.º PTA-10212, de Registro en la ATCC, y también se denomina cepa PTA-10212 de la ATCC de *Thraustochytrium* sp.. Cuando se expresan, estos genes producen perfiles de ácido graso únicos, caracterizados en parte por niveles elevados de ácidos grasos omega-3, en particular niveles elevados de DHA, EPA, o una combinación de los mismos.

Como se usa en el presente documento, un "polinucleótido" puede comprender un enlace fosfodiéster convencional o un enlace no convencional (*por ejemplo*, un enlace amida, tal como se encuentra en ácidos peptidonucleicos (PNA)). Un polinucleótido puede contener la secuencia de nucleótidos de la secuencia de ADNc de longitud completa, incluyendo las secuencias sin traducir en las posiciones 5' y 3', las secuencias de codificación, así como fragmentos, epítopo, dominios y variantes de la secuencia de ácidos nucleicos. El polinucleótido puede estar formado por cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN sin modificar o ARN o ADN modificado. Por ejemplo, los polinucleótidos pueden estar formados por ADN mono y bicatenario, ADN que sea una mezcla de regiones mono y bicatenarias, ARN mono y bicatenario, y ARN que sea una mezcla de mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser mono y bicatenarios, más habitualmente, bicatenarias o una mezcla de regiones mono y bicatenarias. Además, los polinucleótidos pueden estar formados por regiones tricatenarias que comprenden ARN o ADN o tanto ARN como ADN. Los polinucleótidos pueden contener ribonucleósidos (adenosina, guanosina, uridina, o citidina; "moléculas de ARN") o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitimidina, o desoxicitidina; "moléculas de ADN"), o cualquier análogo fosfoéster de los mismos, tal como fosforotioatos y tioésteres. Los polinucleótidos también pueden

contener una o más bases o cadenas principales de ADN o ARN modificadas para estabilidad o por otras razones. Las bases "modificadas" incluyen, por ejemplo, bases tritiladas y bases inusuales tales como inosina. En ADN y ARN se pueden hacer una diversidad de modificaciones; por lo tanto, "polinucleótido" incluye formas modificadas química, enzimática, o metabólicamente. La expresión molécula de ácido nucleico se refiere solo a la estructura primaria y secundaria de la molécula, y no la limita a ninguna forma terciaria en particular. Por lo tanto, esta expresión incluye ADN bicatenario que se encuentra, entre otros, en moléculas de ADN lineal o circular (por ejemplo, fragmentos de restricción), plásmidos y cromosomas. Al discutir la estructura de moléculas de ADN de bicatenarias en particular, las secuencias se pueden describir en el presente documento de acuerdo con la convención normal de dar solo la secuencia en la dirección 5' a 3' a lo largo de la hebra de ADN no transcrita (es decir, la hebra que tiene una secuencia homóloga a la del ARNm).

La expresión molécula de ácido nucleico "aislado" se refiere a una molécula de ácido nucleico, ADN o ARN, que se ha retirado de su entorno nativo. Los ejemplos adicionales de moléculas de ácido nucleico aislado incluyen moléculas de ácido nucleico que comprenden polinucleótidos recombinantes mantenidos en células hospedadoras heterólogas o polinucleótidos purificados (parcial o sustancialmente) en solución. Las moléculas de ARN aislado incluyen transcritos de ARN *in vivo* o *in vitro* de polinucleótidos de la presente invención. Las moléculas de ácido nucleico aislado de acuerdo con la presente invención incluyen adicionalmente las moléculas de este tipo producidas por vía sintética. Además, una molécula de ácido nucleico o polinucleótido puede incluir un elemento regulador tal como un promotor, sitio de unión a ribosoma, o un terminador de la transcripción.

Un "gen" se refiere a un conjunto de nucleótidos que codifican un polipéptido, e incluye ADNc y ácidos nucleicos de ADN genómico. "Gen" también se refiere a un fragmento de ácido nucleico que expresa una proteína específica, incluyendo secuencias de intervención (intrones) entre segmentos codificante sin individuales (exones), así como secuencias reguladoras que preceden (secuencias no codificantes en la posición 5') y que siguen (secuencias no codificantes en la posición 3') a la secuencia codificante. "Gen nativo" se refiere a un gen tal como se encuentra en la naturaleza con sus propias secuencias reguladoras.

La presente invención se refiere a una molécula de ácido nucleico aislada que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la secuencia de polinucleótidos de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5), en la que el polinucleótido codifica un polipéptido que tiene actividad de PUFA sintasa que consiste en actividad de β -hidroxiacil-ACP deshidrasa similar a FabA (DH) y actividad de enoil-ACP reductasa (ER).

Las actividades de PUFA sintasa están asociadas con uno o más dominios en cada polipéptido de sintasa, en las que los dominios se pueden identificar mediante sus motivos estructurales o funcionales conservados basándose en su homología con motivos conocidos y también se pueden identificar basándose en sus actividades bioquímicas específicas. Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.217.856, incorporado por referencia en el presente documento en su totalidad. Los ejemplos de dominios de PUFA sintasa incluyen: el dominio de beta-cetoacil-ACP sintasa (KS), dominio de malonil-CoA:ACP aciltransferasa (MAT), dominios de proteína de vehículo de acilo (ACP), dominio de cetorreductasa (KR), y dominio de beta-hidroxiacil-ACP deshidrasa (DH) en Pfa1p; el dominio de KS, el dominio de factor de longitud de cadena (CLF), dominio de aciltransferasa (AT), y dominio de enoil-ACP reductasa (ER) en Pfa2p; y los dominios DH y el dominio de ER en Pfa3p.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad biológica (función) de beta-cetoacil-ACP sintasa (KS) es capaz de realizar la etapa inicial del ciclo de reacción de elongación de ácido graso. La expresión "beta-cetoacil-ACP sintasa" se ha usado indistintamente con las expresiones "3-ceto acil-ACP sintasa", "beta-cetoacil-ACP sintasa", y "ceto-acil ACP sintasa". En otros sistemas, se ha mostrado que el grupo acilo para elongación está relacionado con un resto de cisteína en el sitio activo de KS mediante un enlace tioéster, y el acil-KS experimenta condensación con malonil-ACP para formar -cetoacil-ACP, CO₂, y KS sin unir ("libre"). En los sistemas de este tipo, se ha mostrado que KS posee una especificidad de sustrato mayor que otros polipéptidos del ciclo de reacción. Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de KS mediante homología con secuencias conocidas de KS.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad de malonil-CoA:ACP aciltransferasa (MAT) es capaz de transferir el resto malonilo de malonil-CoA a ACP. La expresión "malonil-CoA:ACP aciltransferasa" se ha usado indistintamente con "malonil aciltransferasa". Además del motivo de sitio activo (GxSxG), se ha mostrado que las MAT poseen un motivo ampliado (aminoácidos R y Q en posiciones fundamentales). Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de MAT mediante homología con secuencias conocidas de MAT y mediante su estructura de motivo ampliado.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad de proteína transportadora de acilo (ACP) es capaz de funcionar como un vehículo para crecimiento de cadenas de acilo graso a través de un enlace tioéster a un cofactor unido de manera covalente. Las ACP por lo general tienen una longitud de aproximadamente 80 a aproximadamente 100 aminoácidos y se ha mostrado que se convierte en formas apo inactivas a formas holo funcionales mediante transferencia del resto fosfopanteteinilo de CoA a un resto de serina altamente conservado de la ACP. También se ha mostrado por los grupos acilo según en a las ACP mediante un

enlace tioéster en el extremo libre del resto fosfopanteteínico. La presencia de variaciones de un motivo de sitio activo (LGIDS*) también se ha reconocido como una firma de las ACP. La funcionalidad de la serina de sitio activo (S*) se ha demostrado en una PUFA sintasa bacteriana (Jiang *et al.*, J. Am. Chem. Soc. 130: 6336-7 (2008)). Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de ACP mediante etiquetado con panteteína radiactiva y mediante homología de secuencias con ACP conocidas.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad de deshidrasa o deshidratasa (DH) es capaz de catalizar una reacción de deshidratación. La referencia a la actividad de DH por lo general se refiere a la actividad biológica de beta-hidroxiacil-ACP deshidrasa similar a FabA. La actividad biológica de beta-hidroxiacil-ACP deshidrasa similar a FabA retira HOH de una beta-cetoacil-ACP e inicialmente produce un doble enlace trans en la cadena de carbono. La expresión "beta-hidroxiacil-ACP deshidrasa similar a FabA" se ha usado indistintamente con las expresiones "beta-hidroxi acil-ACP deshidrasa similar a FabA", "beta-hidroxiacil-ACP deshidrasa", y "deshidrasa". Previamente se ha demostrado que los dominios de DH de los sistemas de PUFA sintasa muestran homología con enzimas de DH bacterianas asociadas con sistemas de FAS (en lugar de los dominios de DH de otros sistemas de PKS). Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.217.856, incorporado por referencia en el presente documento en su totalidad. Un subconjunto de DH bacterianas, las DH similares a FabA, poseen actividad de isomerasa cis-trans (Heath *et al.*, J. Biol. Chem., 271, 27795 (1996)). Basándose en la homología de las proteínas de DH similares a FabA, uno o todos los dominios de DH del sistema de PUFA sintasa pueden ser responsables de la inserción de dobles enlaces cis en los productos de PUFA sintasa. Un polipéptido o dominio también puede tener actividad de DH no similar a FabA, o actividad beta-hidroxiacil-ACP deshidrasa no similar a FabA (DH). De forma más específica, un motivo de sitio activo conservado de aproximadamente 13 aminoácidos de longitud se ha identificado previamente en dominios de DH de PUFA sintasa: LxxHxxxGxxxxP (la posición L también puede ser una I en emotivo). Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.217.856, y Donadio S, Katz L., Gene 111 (1): 51-60 (1992), cada uno de los cuales se incorpora por referencia en el presente documento en su totalidad. Este motivo conservado se encuentra en una región similar de todas las secuencias conocidas de PUFA sintasa y podría ser responsable de una deshidratación similar a no FabA.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad de beta-cetoacil-ACP reductasa (KR) es capaz de catalizar la reducción dependiente de piridina-nucleótido de formas 3-cetoacilo de ACP. La expresión "beta-cetoacil-ACP reductasa" se ha usado indistintamente con las expresiones "cetorreductasa", "3-cetoacil-ACP reductasa", y "ceto-acil ACP reductasa". En otros sistemas se ha determinado que la función de KR implica la primera etapa reductora en el ciclo de elongación de biosíntesis de ácidos grasos *de novo*. Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de KR mediante homología de secuencias con las KR de PUFA sintasa conocidas.

Previamente se ha definido que un polipéptido o un dominio de un polipéptido que tiene actividad del factor de longitud de cadena (CLF) tiene una o más de las actividades o características siguientes: (1) puede determinar el número de ciclos de elongación y por tanto la longitud de la cadena del producto final, (2) tiene homología con KS, pero carece de la cisteína de sitio activo de KS, (3) se puede heterodimerizar con KS, (4) puede proporcionar el grupo acilo inicial a elongar, o (5) puede decarboxilar el malonato (como malonil-ACP), formando de este modo un grupo acetato que se puede transferir al sitio activo de KS y que puede actuar como la molécula de 'cebado' que experimenta la reacción de elongación (condensación) inicial. Un dominio de CLF se encuentra en todos los y los sistemas de PUFA sintasa identificados en la actualidad y en cada caso se encuentra como parte de una proteína de múltiples dominios. Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de CLF mediante homología de secuencias con los CLF de PUFA sintasa conocidos.

Previamente se ha definido que un polipéptido o un dominio de un polipéptido que tiene actividad de aciltransferasa (AT) tiene una o más de las actividades o características siguientes: (1) puede transferir el grupo acilo graso desde el dominio o dominios de ACP a agua (es decir, una tioesterasa), liberando el grupo acilo graso como un ácido graso libre, (2) puede transferir un grupo acilo graso a un aceptor tal como CoA, (3) puede transferir el grupo acilo entre los diversos dominios de ACP, o (4) puede transferir el grupo acilo graso a una molécula aceptora lipófila (por ejemplo, a ácido lisofosfórico). Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de AT mediante homología de secuencias con las AT de PUFA sintasa conocidas.

Previamente se ha mostrado que un polipéptido o un dominio de un polipéptido que tiene actividad biológica de enoil-ACP reductasa (ER) es capaz de reducir el doble enlace trans (introducido por la actividad de DH) en el acilo graso-ACP, dando como resultado la saturación de los carbonos asociados. Previamente se ha mostrado que el dominio de ER en los sistemas de PUFA sintasa tiene homología con una familia de enzimas de ER (Heath *et al.*, Nature 406: 145-146 (2000), incorporado por referencia en el presente documento en su totalidad), y se ha mostrado que un homólogo de ER funciona como una enoil-ACP reductasa *in vitro* (Bumpus *et al.*, J. Am. Chem. Soc., 130: 11614-11616 (2008), incorporado por referencia en el presente documento en su totalidad). La expresión "enoil-ACP reductasa" se ha usado indistintamente con "enoil reductasa", "enoil ACP-reductasa", y "enoil acil-ACP reductasa". Los polipéptidos (o dominios de polipéptidos) se pueden identificar fácilmente como pertenecientes a la familia de ER mediante homología de secuencias con las ER de PUFA sintasa conocidas.

La molécula de ácido nucleico comprende una secuencia de polinucleótidos al menos un 80 % idéntica a una secuencia de polinucleótidos dentro de *PFA3* (SEQ ID NO: 5) que codifica un dominio de DH (tal como SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 110, o SEQ ID NO: 112), y un dominio de ER (SEQ ID NO: 41 o SEQ ID NO: 114). La presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 5, en la que la secuencia de polinucleótidos codifica un polipéptido que tiene actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER.

En algunas realizaciones, la presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 39, en la que la secuencia de polinucleótidos codifica un polipéptido que comprende actividad de DH.

En algunas realizaciones, la presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 110, en la que la secuencia de polinucleótidos codifica un polipéptido que comprende actividad de DH.

En algunas realizaciones, la presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 112, en la que la secuencia de polinucleótidos codifica un polipéptido que comprende actividad de DH.

En algunas realizaciones, la presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 41 o SEQ ID NO: 114, en la que la secuencia de polinucleótidos codifica un polipéptido que comprende actividad de ER.

La presente invención se refiere a una molécula de ácido nucleico aisladas que comprende secuencias de polinucleótidos que codifican polipéptidos, en la que los polipéptidos comprenden secuencias de aminoácidos que son al menos un 80 % idénticas a las secuencias de aminoácidos de *Pfa3p* (SEQ ID NO: 6), en la que the polinucleótidos codifican polipéptidos que tienen actividades de PUFA sintasa que consisten en actividad de DH y actividad de ER.

En algunas realizaciones, la presente invención se refiere una molécula de ácido nucleicos que comprende una secuencia de polinucleótidos que codifican un polipéptido, en la que el polipéptido comprende una secuencia de aminoácidos al menos un 80 % idéntica a una secuencia de aminoácidos dentro de *Pfa3p* (SEQ ID NO: 6) que comprende uno o más dominios de PUFA sintasa tales como un dominio de DH (tal como SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 111, o SEQ ID NO: 113), un dominio de ER (SEQ ID NO: 42 o SEQ ID NO: 115), y combinaciones de los mismos.

En algunas realizaciones, la presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifican un polipéptido, en la que el polipéptido comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 36 o SEQ ID NO: 109, y en la que el polipéptido comprende actividad de ER.

La presente invención se refiere a una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifican un polipéptido, en la que el polipéptido comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER.

En algunas realizaciones, las moléculas de ácido nucleico comprenden secuencias de polinucleótidos idénticas en al menos aproximadamente un 80 %, 85 %, o 90 % a las secuencias de polinucleótidos informadas en el presente documento, o al menos idénticas en aproximadamente un 95 %, 96 %, 97 %, 98 %, 99 %, o 100 % a las secuencias de polinucleótidos informadas en el presente documento. La expresión "porcentaje de identidad", como se conoce en la técnica, es una relación entre dos o más secuencias de aminoácidos o dos o más secuencias de polinucleótidos, tal como se determina mediante comparación de las secuencias. En la técnica, "identidad" también se refiere al grado de relación de las secuencias entre secuencias de aminoácidos o polinucleótidos, según pueda ser el caso, tal como se determina mediante el emparejamiento entre hebras de tales secuencias.

Por una molécula de ácido nucleico que tiene una secuencia de polinucleótidos "idénticas" en al menos, por ejemplo, un 95 % a una secuencia de polinucleótidos de referencia de la presente invención, se pretende hacer referencia a que la secuencia de polinucleótidos de la molécula de ácido nucleico es idéntica a la secuencia de referencia excepto por que la secuencia de polinucleótidos puede incluir hasta cinco diferencias de nucleótidos por cada 100 nucleótidos de la secuencia de polinucleótidos de referencia. En otras palabras, para obtener una molécula de ácido nucleico que tenga una secuencia de polinucleótidos al menos un 95 % idéntica a una secuencia de polinucleótidos de referencia, hasta un 5 % de los nucleótidos en la secuencia de referencia se pueden suprimir o sustituir con otro nucleótido, o un número de nucleótidos hasta un 5 % de los nucleótidos en la secuencia de referencia se pueden insertar en la secuencia de referencia.

Como una cuestión práctica, cualquier secuencia de polinucleótidos o secuencia de aminoácidos en particular es al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % idéntica a una secuencia de polinucleótidos o secuencia de aminoácidos de la presente invención se puede determinar de manera convencional usando programas informáticos conocidos. Un método para determinar el mejor emparejamiento global entre una secuencia de búsqueda (una secuencia de la presente invención) y una secuencia objeto se puede determinar usando el alineamiento de secuencias y cálculo de puntuaciones de identidad. Los alineamientos se realizaron usando el programa informático AlignX, que es un componente del paquete Vector NTI Suite 10.0 de Invitrogen (www.invitrogen.com). Los alineamientos se realizaron usando un alineamiento ClustalW (Thompson, J.D., *et al.*, Nucl. Acids Res. 22: 4673-4680 (1994)) para alineamientos de secuencias tanto de aminoácidos como de polinucleótidos. Las matrices de puntuación por defecto, Blosum62mt2 y swgapdnamt, se usaron para alineamientos de secuencias de aminoácidos y polinucleótidos, respectivamente. Para secuencias de aminoácidos, la penalización de apertura de hueco por defecto es 10 y la penalización por extensión del hueco es 0,1. Para secuencias de polinucleótidos, la penalización de apertura de hueco por defecto es 15 y la penalización por extensión del hueco es 6,66.

La presente invención se refiere a una molécula de ácido nucleico aislada que comprende una secuencia de polinucleótidos que codifican un polipéptido que comprende actividad de PUFA sintasa que consiste en actividad de ER y actividad de DH, en la que el polinucleótido se hibrida en condiciones rigurosas al complemento de cualquiera de las secuencias de polinucleótidos que se han descrito anteriormente.

Una molécula de ácido nucleico se puede "hibridar" a otra molécula de ácido nucleico, tal como un ADNc, ADN genómico, o ARN, cuando una forma monocatenaria de la molécula de ácido nucleico se puede hibridar a la otra molécula de ácido nucleico en las condiciones apropiadas de temperatura y fuerza iónica de la solución. Las condiciones de hibridación y lavado se conocen bien y se ejemplifican. Véase, por ejemplo, Sambrook J. y Russell D. 2001. Molecular cloning: A laboratory manual, 3ª edición. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. Las condiciones de temperatura y fuerza iónica determinan la "rigurosidad" de la hibridación. Las condiciones de rigurosidad se pueden ajustar para identificar sistemáticamente fragmentos moderadamente similares, tales como secuencias homólogas de organismos relacionados de manera alejada, a fragmentos altamente similares, tales como genes quintuplican enzimas funcionales de organismos muy relacionados. Los lavados después de la hibridación determinan las condiciones de rigurosidad. Un conjunto de condiciones usa una serie de grabados que parte de 6X SSC, SDS al 0,5 % a temperatura ambiente durante 15 min, a continuación se repite con 2X SSC, SDS al 0,5 % a 45 °C durante 30 min, y a continuación se repite dos veces con 0,2X SSC, SDS al 0,5 % a 50 °C durante 30 min. Para condiciones más rigurosas, se realizan lavados a temperaturas más elevadas en que los lavados son idénticos a los mencionados anteriormente excepto para temperatura de los dos lavados finales de 30 min en 0,2X SSC, SDS al 0,5 % que aumentan a 60 °C. Otro conjunto de condiciones altamente rigurosas usados lavados finales en 0,1X SSC, 0,1 % SDS a 65 °C. Un conjunto adicional de condiciones altamente rigurosas se definen mediante hibridación a 0,1X SSC, 0,1 % SDS, 65 °C y se lava con 2X SSC, SDS al 0,1 % seguido de 0,1X SSC, SDS al 0,1 %.

La presente invención se refiere a una molécula de ácido nucleico aislada que comprende una secuencia de polinucleótidos que es totalmente complementaria a cualquiera de las secuencias de polinucleótidos que se han descrito anteriormente. El término "complementario" se usa para describir la relación entre bases de nucleótidos. Por ejemplo, con respecto al ADN, la adenosina es complementaria a la timina y la citosina es complementaria a la guanina.

En ciertas realizaciones, el polinucleótido o ácido nucleico es ADN. En el caso del ADN, una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifica un polipéptido puede incluir normalmente un promotor y/o otros elementos de control de la transcripción o de la traducción asociados de manera operativa con una o más regiones codificantes. Una asociación operable se produce cuando la región codificante para un producto genético, por ejemplo, un polipéptido, se asocia con una o más secuencias reguladoras de una manera tal que se coloca la expresión del producto genético bajo la influencia o control de la secuencia o secuencias reguladoras. Dos fragmentos de ADN (tales como una región codificante de polipéptido y un promotor asociado a la misma) se "asocian de manera operativa" si la inducción de la función del promotor da como resultado la transcripción de ARNm que codifica el producto genético deseado y si la naturaleza de la unión entre los dos fragmentos de ADN no interfiere con la capacidad de las secuencias reguladoras de la expresión para dirigir la expresión del producto genético o interferir con la capacidad del molde de ADN a transcribir. Por lo tanto, una región promotora se podría asociar de manera operativa con una secuencia de polinucleótidos que codifican un polipéptido si el promotor fuera capaz de realizar la transcripción de una secuencia de polinucleótidos. El promotor puede ser un promotor específico de célula que dirija la transcripción sustancial del ADN solamente en células predeterminadas. En general, una región codificante se sitúa en la posición 3' con respecto a un promotor. Los promotores se pueden obtener en su totalidad a partir de un gen nativo, o pueden estar formados por diferentes elementos obtenidos a partir de diferentes promotores encontrados en la naturaleza, o incluso pueden comprender segmentos de ADN sintético. Los expertos en la materia entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos de células, o en diferentes estadios del desarrollo, o como respuesta a diferentes condiciones ambientales o psicológicas. Los promotores que hacen que un gen se exprese en la mayoría de los tipos celulares la mayoría de las veces comúnmente se denominan "promotores constitutivos". Además se reconoce que dado que en la mayoría

de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, los fragmentos de ADN de diferentes longitudes pueden tener una actividad promotora idéntica. Un promotor por lo general se une en su extremo 3' terminal mediante el sitio de iniciación de la transcripción y se extiende cadena arriba (dirección 5') para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Dentro del promotor se encontrará sitio de inicio de la transcripción (definido de forma conveniente por ejemplo, mediante formación de mapas con nucleasa S1), así como dominios de unión a proteína (secuencias consenso) responsables de la unión de ARN polimerasa.

Las regiones reguladoras adecuadas incluyen regiones de ácido nucleico situadas cadena arriba (secuencias no codificantes en la posición 5'), dentro, o cadena abajo (secuencias no codificantes en la posición 3') de una región codificante, y que influyen en la transcripción, procesamiento o estabilidad del ARN, o traducción de la región codificante asociada. Las regiones reguladoras pueden incluir promotores, secuencias líder de traducción, sitios de procesamiento de ARN, sitios de unión a efector y estructuras de tallo-lazo. Otros elementos de control de la transcripción, además de un promotor, por ejemplo potenciadores, operadores, represores y señales de terminación de la transcripción, se pueden asociar de forma operativa con el polinucleótido para dirigir la transcripción específica de la célula. Los límites de la región codificante se determinan mediante un codón de inicio en el extremo 5' terminal (amino) y un codón de terminación de la traducción en el extremo 3' terminal (carboxilo). Una región codificante puede incluir, pero no se limita a, regiones procariotas, ADNc de ARNm, moléculas de ADN genómico, moléculas de ADN sintético, o moléculas de ARN. Si la región codificante está destinada a la expresión en una célula eucariótica, una señal de poliadenilación y secuencia de terminación de la transcripción se situarán normalmente en la posición 3' con respecto a la región codificante.

En ciertos aspectos de la invención, como cebadores de PCR se pueden usar secuencias de polinucleótidos que tengan al menos 20 bases, al menos 30 bases, al menos 50 bases y que se hibridan con una secuencia de polinucleótidos de la presente invención. Por lo general, en técnicas de amplificación de tipo PCR, los cebadores tienen secuencias diferentes y no son complementarios entre sí. Dependiendo de las condiciones de ensayo deseadas, las secuencias de los cebadores se deberían diseñar para proporcionar una replicación tanto eficaz como fiel del ácido nucleico diana. Los métodos de diseño de cebadores de PCR son comunes y se conocen bien en la técnica. Por lo general se pueden usar dos segmentos cortos de las secuencias presentes en protocolos de reacción en cadena de polimerasa (PCR) para amplificar fragmentos de ácido nucleico más largos que codifican genes homólogos a partir de ADN o ARN. La reacción en cadena de la polimerasa también se puede realizar en una biblioteca de fragmentos de ácido nucleico clonados en los que la secuencia de un cebador se obtiene a partir de los presentes fragmentos de ácido nucleico y la secuencia del otro cebador aprovecha la presencia de los tramos de ácido poliadenílico con respecto al extremo en la posición 3' del precursor de ARNm que codifica genes microbianos. Como alternativa, la segunda secuencia de cebadores se puede basar en secuencias obtenidas a partir del vector de clonación.

Además, se pueden diseñar cebadores específicos y usar para amplificar una parte o una longitud completa de las presentes secuencias. Los productos de amplificación resultantes se pueden etiquetar directamente durante las reacciones de amplificación o se pueden etiquetar después de reacciones de amplificación, y se pueden usar como sondas para aislar fragmentos de ADN de longitud completa en condiciones de rigurosidad apropiada.

Por lo tanto, las moléculas de ácido nucleico de la presente invención se pueden usar para aislar genes que codifican proteínas homólogas de la misma especie o de otras especies bacterianas. El aislamiento de genes homólogos usando protocolos dependientes de secuencia se conoce bien en la técnica. Los ejemplos de protocolos dependientes de secuencia incluyen, pero no se limitan a, métodos de hibridación de ácidos nucleicos y métodos de amplificación de ADN y ARN como se usa a modo de ejemplo mediante diversos usos de tecnologías de amplificación de ácidos nucleicos (por ejemplo, reacción en cadena de la polimerasa, Mullis *et al.*, documento de Patente de Estados Unidos N.º 4.683.202; reacción en cadena de ligasa (LCR) (Tabor, S. *et al.*, Proc. Acad. Sci. USA 82: 1074 (1985)); o amplificación por desplazamiento de la hebra (SDA; Walker, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 89: 392 (1992)).

En algunas realizaciones, las moléculas de ácido nucleico aisladas de la presente invención se usan para aislar moléculas de ácido nucleico homólogas de otros organismos con el fin de identificar PUFA sintasas que produzcan perfiles de PUFA similares o mejorados. En algunas realizaciones, las moléculas de ácido nucleico aisladas de la presente invención se usan para aislar moléculas de ácido nucleico homólogas de otros organismos que están involucrados en la producción de altas cantidades de DHA.

Las moléculas de ácido nucleico de la presente invención también comprenden secuencias de polinucleótidos que codifican un gen de PUFA sintasa, un dominio de un gen de PUFA sintasa, un fragmento del gen de PUFA sintasa fusionado en marco a una secuencia marcadora que permite la detección del polipéptido de la presente invención. Las secuencias marcadoras incluyen marcadores auxotróficos o dominantes conocidos por alguien con una experiencia habitual en la materia tal como ZEO (zeocina), NEO (G418), higromicina, arsenito, HPH, NAT, y similares.

La presente invención también incluye variantes del gen de PUFA sintasa. Las variantes pueden contener alteraciones en las regiones codificantes, regiones no codificantes, o ambas. Los ejemplos son variantes de secuencias de polinucleótidos que contienen alteraciones que producen sustituciones silenciosas, adiciones, o delecciones, pero no alteran las propiedades o actividades del polipéptido codificado. En ciertas realizaciones, las variantes de secuencias de polinucleótidos se producen por sustituciones silenciosas debido a la degeneración del código genético. En otras realizaciones, se pueden producir variantes de secuencias de polinucleótidos por diversas razones, por ejemplo, para optimizar la expresión del codón para un huésped particular (por ejemplo, el cambio de codones en el ARNm de traustokuirido a los preferidos por otros organismos tales como *E. coli* o *Saccharomyces cerevisiae*).

En la presente invención también se proporcionan variantes alélicas, ortólogos, y/o especies homólogas. Los procedimientos conocidos en la técnica se pueden usar para obtener genes de longitud completa, variantes alélicas, variantes de corte y empalme, porciones de codificación de longitud completa, ortólogos, y/o homólogos de especies de los genes que se describen en el presente documento usando información de las secuencias desveladas en el presente documento. Por ejemplo, las variantes alélicas y/o los homólogos de especies se pueden aislar e identificar preparando sondas o cebadores adecuados a partir de las secuencias proporcionadas en el presente documento e identificando sistemáticamente una fuente de ácido nucleico adecuada para variantes alélicas y/o el homólogo deseado.

La presente invención se refiere a una molécula de ácido nucleico recombinante que comprende cualquiera de las moléculas de ácido nucleico de descritas anteriormente o combinaciones de las mismas y una secuencia de control de la transcripción. En algunas realizaciones, la molécula de ácido nucleico recombinante es un vector recombinante.

La presente invención se refiere a un método para preparar un vector recombinante que comprende insertar una o más moléculas de ácido nucleico aisladas como se describe en el presente documento en un vector.

Los vectores de la presente invención pueden ser, por ejemplo, un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en forma de un plásmido, una partícula viral, un fago, etc.

Las secuencias de polinucleótidos de la invención se pueden incluir en una cualquiera de una diversidad de vectores de expresión para expresar un polipéptido. Los vectores de este tipo incluyen secuencias de ADN o ARN cromosómico, no cromosómico, y sintético, por ejemplo, derivados de SV40; plásmidos bacterianos; y plásmidos de levadura. Sin embargo, se puede usar cualquier otro vector apropiado conocido para alguien con una experiencia habitual en la materia.

La secuencia de ADN apropiada se puede insertar en el vector mediante una variedad de procedimientos. En general, la secuencia de ADN se inserta en un sitio o sitios de endonucleasa de restricción apropiados mediante procedimientos conocidos en la técnica. Dichos procedimientos y otros se consideran dentro del alcance de los expertos en la materia.

La presente invención también incluye construcciones recombinantes que comprenden una o más de las secuencias de polinucleótidos que se han descrito anteriormente. Las construcciones comprenden un vector, tal como un plásmido o vector viral, en el que se ha insertado una o más secuencias de la invención, en una orientación directa o inversa. En un aspecto de la presente realización, la construcción comprende además secuencias reguladoras, que incluyen, por ejemplo, un promotor, asociado en forma operativa a la secuencia. Los expertos en la materia conocen grandes números de vectores y promotores adecuados, y están disponibles en el mercado.

Polipéptidos

La presente invención se refiere a polipéptidos aislados que comprenden secuencias de aminoácidos para proteínas y dominios de PUFA sintasa obtenidos a partir de los microorganismos aislados depositados como Registro en la ATCC N.^{os} PTA-9695 y PTA-10212.

Como se usa en el presente documento, el término "polipéptido" pretende incluir un "polipéptido" en singular así como "polipéptidos" en plural y se refiere a una molécula formada por monómeros (aminoácidos) unidos de manera lineal mediante enlaces amida (también conocidos como enlaces peptídicos). El término "polipéptido" se refiere a cualquier cadena o cadenas de dos o más aminoácidos y no se hace referencia a una longitud específica del producto. Por lo tanto, péptidos, dipéptidos, tripéptidos, oligopéptidos, "proteína", "cadena de aminoácidos", o cualquier otro término usado para hacer referencia a una cadena o cadenas de dos o más aminoácidos se incluyen dentro de la definición de "polipéptido", y el término "polipéptido" se puede usar en lugar de o indistintamente con cualquiera de estos términos. El término "polipéptido" también pretende hacer referencia a los productos de modificaciones después de la expresión del polipéptido, que incluyen, pero no se limitan a, glicosilación, acetilación, fosforilación, amidación, derivatización mediante grupos de prevención/bloqueo conocidos, escisión proteolítica, o modificación mediante cualquier aminoácido de origen natural.

Los polipéptidos, como se describe en el presente documento, pueden incluir fragmentos, variante o moléculas derivadas de los mismos sin limitación. Los términos "fragmento", "variante", "derivado" y "análogo", cuando se refieren a un polipéptido, incluyen cualquier polipéptido que retenga al menos una cierta actividad biológica. Los fragmentos de polipéptidos pueden incluir fragmentos proteolíticos, fragmentos de delección, y fragmentos que alcancen más fácilmente el sitio de acción cuando se administran a un animal. Los fragmentos de polipéptidos incluyen adicionalmente cualquier porción del polipéptido que comprende un epítipo antigénico o inmunogénico del polipéptido nativo, incluyendo epítipos lineales así como tridimensionales. Los fragmentos de polipéptidos pueden comprender regiones variantes, incluyendo fragmentos como se ha descrito anteriormente, y también polipéptidos con secuencias de aminoácidos alteradas debido a sustituciones, delecciones o inserciones de aminoácidos. Las variantes se pueden producir de manera natural, tal como una variante alélica. Por una "variante alélica" se hace referencia a formas alternas de un gen que ocupa un locus dado en un cromosoma de un organismo. Se pueden producir variantes de origen no natural usando técnicas de mutagénesis conocidas en la técnica. Los fragmentos de polipéptidos de la invención pueden comprender sustituciones, delecciones o adiciones de aminoácidos conservativas o no conservativas. En el presente documento los polipéptidos variantes también se pueden denominar "análogos de polipéptido". Los fragmentos de polipéptido de la presente invención también pueden incluir moléculas derivadas. Como se usa en el presente documento un "derivado" de un polipéptido o un fragmento de polipéptido se refiere a un polipéptido objeto que tiene uno más restos derivatizados químicamente por reacción de un grupo lateral funcional. Como "derivados" también se incluyen los péptidos que contienen uno o más derivados de aminoácidos de origen natural de los veinte aminoácidos convencionales. Por ejemplo, 4-hidroxiprolina se puede sustituir por prolina; 5-hidroxilisina se puede sustituir por lisina; 3-metilhistidina se puede sustituir por histidina; homoserina se puede sustituir por serina; y ornitina se puede sustituir por lisina.

Los polipéptidos de la invención se pueden codificar mediante cualquiera de las moléculas de ácido nucleico de la invención.

La presente invención se refiere a polipéptidos aislados que comprenden secuencias de aminoácidos que son al menos un 80 % idénticas a las secuencias de aminoácidos de Pfa3p (SEQ ID NO: 6), y combinaciones de los mismos, en las que el polipéptido tiene una actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER.

La presente invención se refiere a polipéptidos que comprenden secuencias de aminoácidos que son al menos un 80 % idénticas a una secuencia de aminoácidos dentro de Pfa3p (SEQ ID NO: 6) que comprende dominios de PUFA sintasa.

La presente invención se refiere a un polipéptido que comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER.

En algunas realizaciones, los polipéptidos comprenden secuencias de aminoácidos al menos aproximadamente un 80 %, 85 %, o un 90 % idénticas a las secuencias de aminoácidos informadas en el presente documento, o al menos aproximadamente un 95 %, 96 %, 97 %, 98 %, 99 %, o un 100 % idénticas a las secuencias de aminoácidos informadas en el presente documento.

Por un polipéptido que tiene una secuencia de aminoácidos al menos, por ejemplo, un 95 % "idéntica" a una secuencia de aminoácidos de búsqueda de la presente invención, se hace referencia a que la secuencia de aminoácidos del polipéptido objeto es idéntica a la secuencia de búsqueda excepto en que la secuencia de polipéptidos objeto puede incluir hasta cinco alteraciones de aminoácidos por cada 100 aminoácidos de la secuencia de aminoácidos de búsqueda. En otras palabras, para obtener un polipéptido que tenga una secuencia de aminoácidos al menos un 95 % idéntica a una secuencia de aminoácidos de búsqueda, hasta un 5 % de los restos de aminoácidos en la secuencia objeto se puede insertar, suprimir, (indels) o sustituir con otro aminoácido. Estas alteraciones de la secuencia de referencia se pueden producir en las posiciones amino o carboxi terminales de la secuencia de aminoácidos de referencia o en cualquier parte entre esas posiciones terminales, intercaladas ya sea de forma individual entre restos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia.

Como una cuestión práctica, cualquier polipéptido en particular que tenga una secuencia de aminoácidos que sea al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % idéntica a, por ejemplo, la secuencia de aminoácidos de la presente invención se puede determinar de manera convencional usando programas informáticos conocidos. Como se ha analizado anteriormente, un método para determinar el mejor emparejamiento global entre una secuencia de búsqueda (una secuencia de la presente invención) y una secuencia objeto se puede determinar usando el alineamiento de secuencias y cálculo de puntuaciones de identidad. Los alineamientos se realizaron usando el programa informático AlignX, que es un componente del paquete Vector NTI Suite 10.0 de Invitrogen (www.invitrogen.com). Los alineamientos se realizaron usando un alineamiento ClustalW (J. Thompson *et al.*, Nucleic Acids Res. 22 (22): 4673-4680 (1994)). Se usó la matriz Blosum62mt2 de puntuación por defecto. La penalización de apertura de hueco por defecto es 10 y la penalización por extensión del hueco es 0,1.

En aspectos adicionales de la invención, las moléculas de ácido nucleico que tienen secuencias de polinucleótidos al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % idénticas a las secuencias de polinucleótidos desveladas en el presente documento, codifican un polipéptido que tiene una o más actividades de PUFA sintasa. Los polipéptidos que tienen una o más actividades de PUFA sintasa presentan una o más actividades similares a, pero no necesariamente idénticas a, una o más actividades de una PUFA sintasa de la presente invención.

Por supuesto, debido a la degeneración del código genético, alguien con experiencia habitual en la materia reconocerá inmediatamente que una gran parte de las moléculas de ácido nucleico que tienen una secuencia de polinucleótidos al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, o un 99 % idénticas a las secuencias de polinucleótidos que se describen en el presente documento codificarán polipéptidos "que tengan actividad funcional de PUFA sintasa". De hecho, dado que todas las variantes degeneradas de cualquiera de estas secuencias de polinucleótidos codifican el mismo polipéptido, en muchos casos, el experto en la materia puede predecir, basándose en el conocimiento de sustituciones conservativas así como de dominios funcionales conservados, qué polipéptidos presentarán actividad. En ciertos aspectos de la invención, los polipéptidos y polinucleótidos de la presente invención se proporcionan en una forma aislada, por ejemplo, purificada hasta homogeneidad. Como alternativa, los polipéptidos y polinucleótidos de la invención se pueden producir de forma sintética mediante sintetizadores convencionales.

Como se sabe en la técnica, la "similitud" entre todos polipéptidos se determina comparando la secuencia de aminoácidos y sustitutos de aminoácidos conservados de la misma del polipéptido con respecto a la secuencia de un segundo polipéptido.

En algunas realizaciones, un polipéptido de la invención es un polipéptido de fusión.

Como se usa en el presente documento, "polipéptido de fusión" se refiere a un polipéptido que comprende un primer polipéptido conectado de manera lineal, a través de enlaces peptídicos, a un segundo polipéptido. El primer polipéptido y el segundo polipéptido pueden ser idénticos o diferentes, y se pueden conectar directamente, o se pueden conectar a través de un conector peptídico. Como se usa en el presente documento, los términos "unido", "fusionado", o "fusión" se usan indistintamente. Estos términos se refieren a la unión en conjunto de dos o más elementos o componentes mediante cualquier medio que incluye medios de conjugación química o medios recombinantes. Una "fusión en marco" se refiere a la unión de dos o más marcos de lectura abiertos para formar un marco de lectura abiertos más largo continuo, de una manera que mantenga el marco de lectura correcto de los marcos de lectura abiertos originales. Por lo tanto, la proteína de fusión recombinante resultante es una sola proteína que contiene dos o más segmentos que corresponden a polipéptidos codificados por los marcos de lectura abiertos originales (cuyos segmentos normalmente no están unidos a la naturaleza). Aunque el marco de lectura por lo tanto se hace continuo a través de los segmentos fusionados, los segmentos pueden estar separados física o espacialmente, por ejemplo, mediante secuencia conectora en marco. Una secuencia "conectora" es una serie de uno o más aminoácidos que separan dos regiones codificantes de polipéptidos en una proteína de fusión.

También se describe una composición que comprende uno o más polipéptidos de la invención y un vehículo biológicamente aceptable.

En algunas composiciones, la composición incluye un "excipiente" biológicamente aceptable, en la que el excipiente es un componente, o mezcla de componentes, que se usa en una composición de la presente invención para dar características deseables a la composición, y también incluyen vehículos. "Biológicamente aceptable" se refiere a un compuesto, material, composición, sal, y/o forma de dosificación que, dentro del alcance del criterio médico sensato, es adecuado para su contacto con los tejidos de células vivas sin una toxicidad, irritación, respuesta inflamatoria excesivas, otras complicaciones problemáticas con respecto a la duración deseada del contacto o proporcionar con una relación razonable de beneficio/riesgo. Se pueden usar diversos excipientes. En algunas realizaciones, el excipiente puede ser, pero no se limita a, un agente alcalino, un estabilizante, un antioxidante, un agente de adhesión, un agente de separación, un agente de revestimiento, un componente de fase exterior, un componente de liberación controlada, un disolvente, un tensioactivo, un humectante, un agente de tamponamiento, una carga, un emoliente, o combinaciones de los mismos. Los excipientes, además de los que se discuten en el presente documento, pueden incluir excipientes que se enumeran en, aunque no se limitan a to, Remington: The Science and Practice of Pharmacy, 21ª ed. (2005). La inclusión de un excipiente en una clasificación en particular en el presente documento (por ejemplo, "disolvente") pretende ilustrar más que limitar el papel del excipiente. Un excipiente en particular puede entrar dentro de múltiples clasificaciones.

También se describe un fragmento, variante, derivado, o análogo de cualquiera de los polipéptidos que se desvelan en el presente documento.

El polipéptido de la presente invención puede ser un polipéptido recombinante, un polipéptido natural, o un polipéptido sintético.

Células Hospedadoras

La presente invención se refiere a una célula hospedadora que expresa cualquiera de las moléculas de ácido nucleico y moléculas de ácido nucleico recombinante descritas anteriormente así como combinaciones de los mismos.

El término "expresión", como se usa en el presente documento, se refiere a un proceso mediante el cual un gen produce un agente bioquímico, por ejemplo, un ARN o polipéptido. El proceso incluye cualquier manifestación de la presencia funcional del gen dentro de la célula que incluye, pero no se limita, supresión genética así como tanto expresión transitoria como expresión estable. Incluye, pero no se limita, transcripción del gen en ARN mensajero (ARNm), ARN de transferencia (ARNt), ARN horquillado pequeño (ARNsh), ARN de interferencia pequeño (ARNsi), o cualquier otro producto de ARN, y la traducción de tal ARNm en polipéptido(s). Si el producto final deseado es bioquímico, la expresión incluye la creación de ese agente bioquímico y cualquier precursor.

Para producir uno o más ácidos grasos poliinsaturados deseados, una célula hospedadora se puede modificar genéticamente para introducir un sistema de PUFA sintasa de la presente invención en la célula hospedadora.

Cuando se modifican genéricamente organismos para expresar un sistema de PUFA sintasa de acuerdo con la presente invención, algunos organismos operadores pueden expresar de forma endógena proteínas auxiliares que se requieren en conjunto con un sistema de PUFA sintasa para producir los PUFA. Sin embargo, puede ser necesario transformar algunos organismos con moléculas de ácido nucleico que codifican una o más proteína o proteínas auxiliares para permitir o para aumentar la producción de los PUFA por el organismo, incluso si el organismo produce de manera endógena una proteína auxiliar homóloga. Algunas proteínas auxiliares heterólogas pueden funcionar de manera más eficaz o más eficiente con las proteínas de PUFA sintasa transformadas que quien la proteína o proteínas auxiliares endógenos de las células hospedadoras.

Las proteínas auxiliares se definen en el presente documento como proteínas de las que no se considera que forman parte del sistema de PUFA sintasa núcleo (es decir, no forman parte del propio complejo enzimático de PUFA sintasa) pero que pueden ser necesarias para la producción de PUFA o una producción eficaz de PUFA usando el complejo enzimático de PUFA sintasa de núcleo de la presente invención. Por ejemplo, para producir los PUFA, un sistema de PUFA sintasa debe funcionar con una proteína auxiliar que transfiera un resto de 4'-fosfopanteteinilo desde la coenzima A al dominio o dominios de proteína transportadora de acilo (ACP). Por lo tanto, se puede considerar que un sistema de PUFA sintasa incluye al menos un dominio de 4'-fosfopanteteinil transferasa (PPTasa), o se puede considerar que un dominio de ese tipo es un dominio o proteína auxiliar para el sistema de PUFA sintasa. Las características estructurales y funcionales de las PPTasas se han descrito con detalle, por ejemplo, en las Publicaciones de Solicitud de Estados Unidos N.º 2002/0194641; 2004/0235127; y 2005/0100995.

Un dominio o proteínas que tenga actividad (función) biológica de 4'-fosfopanteteinil transferasa (PPTasa) se caracteriza como la enzima que transfiere un resto de 4'-fosfopanteteinilo desde la Coenzima A a la proteína transportadora de acilo (ACP). Esta transferencia a un resto de serina invariable de la ACP activa la forma apo inactiva en la forma holo. En las síntesis tanto de poliquétido como de ácido graso, el grupo fosfopanteteína forma tioésteres con las cadenas de acilo en crecimiento. Las PPTasas son una familia de enzimas que se han caracterizado bien en síntesis de ácido graso, síntesis de poliquétido, y síntesis de péptidos no ribosómicos. Se conocen las secuencias de muchas PPTasas, se han determinado estructuras cristalinas (por ejemplo, Reuter K., *et al.*, EMBO J. 18 (23): 6823-31 (1999)), y el análisis de mutación ha identificado restos de aminoácidos importantes para la actividad (Mofid M.R., *et al.*, Biochemistry 43 (14): 4128-36 (2004)).

Una PPTasa heteróloga que previamente se ha demostrado que reconoce dominios de ACP de *Schizochytrium* como sustratos es la proteína Het I de la cepa PCC 7120 de *Nostoc sp.* (denominada anteriormente cepa PCC 7120 de *Anabaena sp.*). Het I está presente en un grupo de genes en *Nostoc* conocido por ser responsable de la síntesis de ácidos hidroxi-grasos de cadena larga que son un componente de una capa glico-lipídica presente en heterocistos de ese organismo (Black y Wolk, J. Bacteriol. 176: 2282-2292 (1994); Campbell *et al.*, Arch. Microbiol. 167: 251-258 (1997)). Es probable que Het I active los dominios de ACP de una proteína, Hgl E, presente en ese grupo. Se han descrito secuencias y construcciones que contienen Het I, por ejemplo, en la Publicación de Solicitud de Estados Unidos N.º 2007/02441 92.

Otra PPTasa heteróloga que previamente se ha demostrado que reconoce los dominios de ACP de *Schizochytrium* es Sfp, obtenida a partir de *Bacillus subtilis*. Sfp se ha caracterizado bien y se usa ampliamente debido a su capacidad para reconocer una amplia gama de sustratos. Basándose en la información de secuencias publicadas (Nakana, *et al.*, Molecular and General Genetics 232: 313-321 (1992)), previamente un vector de expresión se produjo para Sfp por clonación de la región codificante, junto con secuencias de ADN de flanqueo cadena arriba y cadena abajo definidas, en un vector de clonación pACYC-184. Esta construcción codifica una PPTasa funcional como se demuestra mediante su capacidad para ser coexpresada con los Orf de *Schizochytrium* en *E. coli* que, en condiciones apropiadas, da como resultado la acumulación de DHA en esas células (véase, la Publicación de Solicitud de Estados Unidos N.º 2004/02351 27).

Las células hospedadoras pueden incluir células microbianas; células animales; células vegetales; y células de insecto. Los ejemplos representativos de los predadores apropiados incluyen células bacterianas; bacterias termófilas o mesófilas; bacterias marinas; traustóquitridos; células fúngicas, tales como levadura; células vegetales; células de insecto; y células animales aisladas. Las células hospedadoras pueden ser células no transfectadas o células que ya están transfectadas con al menos otra molécula de ácido nucleico recombinante. Las células hospedadoras también pueden incluir células transgénicas que se han diseñado para expresar una PUFA sintasa. Se considera que la selección de un hospedador apropiado está dentro del alcance de los expertos en la materia a partir de las enseñanzas en el presente documento.

Las células hospedadoras incluyen cualquier microorganismo del orden *Thraustochytriales*, tal como microorganismos de un género que incluye, pero no se limita a: *Thraustochytrium Labyrinthoides*, *Japonochytrium*, y *Schizochytrium*. Las especies dentro de estos géneros incluyen, pero no se limitan a: cualquier especie de *Schizochytrium*, incluyendo *Schizochytrium aggregatum*, *Schizochytrium limacinum*, *Schizochytrium minutum*; cualquier especie de *Thraustochytrium* (incluyendo las especies anteriores de *Ulkenia* tales como *U. visurgensis*, *U. amoeboida*, *U. sarkariana*, *U. profunda*, *U. radiata*, *U. minuta* y *Ulkenia* sp. BP-5601), e incluyen *Thraustochytrium striatum*, *Thraustochytrium aureum*, *Thraustochytrium roseum*; y cualquier especie de *Japonochytrium*. Las cepas de *Thraustochytriales* incluyen, pero no se limitan a: *Schizochytrium* sp., (S31) (ATCC 20888); *Schizochytrium* sp., (S8) (ATCC 20889); *Schizochytrium* sp., (LC-RM) (ATCC 18915); *Schizochytrium* sp., (SR21); *Schizochytrium aggregatum* (Goldstein et Belsky) (ATCC 28209); *Schizochytrium limacinum* (Honda et Yokochi) (IFO 32693); *Thraustochytrium* sp. (23B) (ATCC 20891); *Thraustochytrium striatum* (Schneider) (ATCC 24473); *Thraustochytrium aureum* (Goldstein) (ATCC 34304); *Thraustochytrium roseum* (Goldstein) (ATCC 28210); y *Japonochytrium* sp. (L1) (ATCC 28207). Otros ejemplos de microorganismos hospedadores adecuados para modificación genética incluyen, pero no se limitan a, levaduras que incluyen *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis*, u otras levaduras tales como *Candida*, *Kluyveromyces*, u otros hongos, por ejemplo, hongos filamentosos tales como *Aspergillus*, *Neurospora*, *Penicillium*, etc. Las células bacterianas también se pueden usar como hospedadores. Esto incluye *Escherichia coli*, que puede ser útil en procesos de fermentación. Como alternativa, un hospedador tal como una especie de *Lactobacillus* o una especie de *Bacillus* se puede usar como un hospedador.

Las células hospedadoras vegetales incluyen, pero no se limitan a cualquier planta superior, incluyendo plantas tanto dicotiledóneas como monocotiledóneas, y tantas consumibles, incluyendo plantas de cultivo y plantas usadas por sus aceites. Las plantas de este tipo pueden incluir, por ejemplo: canola, soja, colza, lino, maíz, cártamos, girasoles, y tabaco. Otras plantas incluyen las plantas que se conocen porque producen compuestos usados como agentes farmacéuticos, agentes saborizantes, agentes nutracéuticos, ingredientes alimentarios funcionales, agentes cosméticamente activos, o plantas que se modifican genéticamente para producir estos compuestos/agentes. Por lo tanto, se puede seleccionar cualquier especie vegetal o célula vegetal. Los ejemplos de plantas y células vegetales, y plantas cultivadas u obtenidas a partir de las mismas, incluyen, pero no se limitan a, plantas y células vegetales obtenidas a partir de canola (*Brassica rapa* L.); variedades de cultivo de canola NQC02CNX12 (ATCC PTA-6011), NQC02CNX21 (ATCC PTA-6644), y NQC02CNX25 (ATCC PTA-6012) así como variedades de cultivo, variedades de cultivo de cría, y partes de plantas obtenidas a partir de variedades de cultivo de canola NQC02CNX12, NQC02CNX21, y NQC02CNX25 (véanse, los documentos de Patente de Estados Unidos N.º 7.355.100, 7.456.340, y 7.348.473, respectivamente); soja (*Glycine max*); colza (*Brassica* spp.); lino/linaza (*Linum usitatissimum*); maíz (maíz) (*Zea mays*); cártamo (*Carthamus tinctorius*); girasol (*Helianthus annuus*); tabaco (*Nicotiana tabacum*); *Arabidopsis thaliana*, de Brasil (*Betholettia excelsa*); semilla de ricino (*Ricinus communis*); coco (*Cocos nucifera*); cilantro (*Coriandrum sativum*); algodón (*Gossypium* spp.); cacahuete (*Arachis hypogaea*); yoyoba (*Simmondsia chinensis*); mostaza (*Brassica* spp. y *Sinapis alba*); aceite de palma (*Elaeis guineensis*); oliva (*Olea europaea*); arroz (*Oryza sativa*); calabacín (*Cucurbita maxima*); cebada (*Hordeum vulgare*); trigo (*Triticum aestivum*); y lentejas de agua (*Lemnaceae* sp.). Las líneas de plantas a partir de estas y otras plantas se pueden producir, seleccionar, u optimizar para un rasgo deseado tal como hubo asociado con, pero que no se limita a, rendimiento de la semilla, resistencia a la fijación en la siembra, aparición, resistencia o tolerancia enfermedades, madurez, nivel de integridad de la planta en estación tardía, altura de la planta, resistencia a la destrucción, facilidad de transformación de la planta, contenido de aceite, o perfil del aceite. Las líneas de plantas se pueden seleccionar a través de cultivo de plantas tales como cultivo de pedigrí, cultivo de selección recurrente, cultivo inter cruzado y retro cruzado, así como métodos tales como reproducción asistida por marcador y labranza. Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.348.473.

Las células animales incluyen cualquier célula animal aislada.

La presente invención se refiere a una célula hospedadora que expresa una o más moléculas de ácido nucleico o moléculas de ácido nucleico recombinante, incluyendo vectores, de la invención.

La presente invención se refiere a un método para preparar una célula hospedadora recombinante que comprende introducir un vector recombinante en una célula hospedadora.

Las células hospedadoras se pueden modificar por ingeniería genética (transducir o transformar o transfectar) con los vectores de la presente invención que pueden ser, por ejemplo, un vector de clonación o un vector de expresión. El vector se puede encontrar, por ejemplo, en forma de un plásmido, una partícula viral, un

fago, etc. El vector que contiene una secuencia de polinucleótidos como se describe en el presente documento, así como un promotor o secuencia de control apropiados, se puede usar para transformar un hospedador apropiado para permitir la expresión del polipéptido codificado por la secuencia de polinucleótidos. La modificación genética en células hospedadoras también por incluir la optimización de genes para uso de codón de hospedador preferente u óptimo.

Las células hospedadoras modificadas por ingeniería se pueden cultivar en medios nutrientes convencionales modificados y para apropiado para la activación de promotores, selección de transformantes, o amplificación de los genes de la presente invención. Las condiciones de cultivo, tales como temperatura, pH, y similares, son las que se han usado previamente con la célula hospedadora seleccionada para expresión, y sean evidentes para una persona con una experiencia habitual en la materia.

En algunas realizaciones, la presente invención se refiere a la modificación genética de una planta o parte de una planta para expresar un sistema de PUFA sintasa que se describe en el presente documento, que incluye al menos el complejo enzimático de PUFA sintasa de núcleo. Una "parte de una planta" o "parte de planta", se define en el presente documento incluye cualquier parte de una planta, tal como, pero no se limita a, semillas (inmaduras o maduras), aceites, polen, embriones, flores, frutos, brotes, hojas, raíces, yemas, explantes, etc. En algunas realizaciones, la planta o parte de una planta modificada genéticamente produce uno o más PUFA, tales como EPA, DHA, DPA (n-3 o n-6), ARA, GLA, SDA, otros PUFA, y combinaciones de los mismos. No se sabe si las plantas contienen de manera endógena un sistema de PUFA sintasa; por lo tanto, los sistemas de PUFA sintasa de la presente invención se pueden usar para modificar por ingeniería plantas con capacidades públicas de producción de ácido graso. En una realización más, la planta o parte de una planta además se modifica genéticamente para expresar al menos una proteína auxiliar de PUFA sintasa, (por ejemplo, una PPTasa). En algunas realizaciones, la planta es una planta con semillas oleaginosas, en la que las semillas oleaginosas, y/o y la aceite en las semillas oleaginosas, contienen los PUFA producidos por el sistema de PUFA sintasa. En algunas realizaciones, las plantas modificadas genéticamente, partes de plantas, semillas oleaginosas, y/o aceites en las semillas oleaginosas contienen una cantidad detectable de al menos un PUFA que es el producto del sistema de PUFA sintasa. En otras realizaciones, tales plantas, partes de plantas, semillas oleaginosas, y/o aceites en las semillas oleaginosas pueden estar sustancialmente libres de productos intermedios o secundarios que no son los productos primarios de PUFA del sistema de PUFA sintasa introducido y que no se produce una naturaleza por el sistema de FAS endógeno en las plantas de tipo silvestre. Aunque las plantas de tipo silvestre producen algunos PUFA de cadena corta o media de tipo silvestre, tales como los PUFA de 18 carbonos a través del sistema de FAS, se producirán PUFA nuevos o adicionales en la planta, partes de plantas, semillas oleaginosas, y/o aceites en las semillas oleaginosas como resultado de la modificación genética con un sistema de PUFA sintasa que se describe en el presente documento.

La modificación genética de una planta se puede realizar usando desarrollo de cepas y/o técnicas de genética molecular clásicos. Véase, Publicación de Solicitud de Estados Unidos N.º 2007/0244192. En la técnica se conocen métodos para producir una planta transgénica, en la que una molécula de ácido nucleico recombinante que codifica una secuencia de aminoácidos deseada se incorpora en el genoma de la planta. Por ejemplo, se pueden usar vectores virales para producir plantas transgénicas, tal como mediante transformación de una planta monocotiledónea con un vector viral usando los métodos descritos en los documentos de Patente de Estados Unidos N.ºs 5.569.597; 5.589.367; y 5.316.931. En la técnica también se conocen métodos para la ingeniería genética o modificación de plantas mediante transformación, incluyendo protocolos de transformación biológica y física. Véase, por ejemplo, B.L. Miki *et al.*, Procedures for Introducing Foreign DNA in Plants, en METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY 67-88 (Glick, B. R. y Thompson, J. E. eds., CRC Press, Inc., Boca Raton, 1993). Además, hay disponibilidad de vectores y métodos de cultivo *in vitro* para transformación de células o tejidos vegetales y regeneración de plantas. Véase, por ejemplo, M. Y. Gruber *et al.*, Vectors for Plant Transformation, en METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY 89-119 (Glick, B. R. y Thompson, J. E. eds., CRC Press, Inc., Boca Raton, 1993).

Un método ampliamente utilizado para introducir un vector de expresión en plantas se basa en el sistema de transformación natural de *Agrobacterium*. Véase, por ejemplo, Horsch *et al.*, Science 227: 1229 (1985) y el documento de Patente de Estados Unidos N.º 6.051.757. *A. tumefaciens* y *A. rhizogenes* son bacterias del suelo patógenas vegetales que transforman genéticamente las células vegetales. Los plásmidos Ti y Ri de *A. tumefaciens* y *A. rhizogenes*, respectivamente, por genes responsables de la transformación genética de la planta. Véase, por ejemplo, Kado, C. I., Crit. Rev. Plant. Sci. 10: 1 (1991). Las descripciones de sistemas de vector de *Agrobacterium* y métodos para transferencia genética mediada por *Agrobacterium* se proporcionan en numerosas referencias, incluyendo Gruber *et al.*, mencionado anteriormente; Miki *et al.*, mencionado anteriormente; Moloney *et al.*, Plant Cell Reports 8: 238 (1989); los documentos de Patente de Estados Unidos N.ºs 5.177.010; 5.104.310; 5.149.645; 5.469.976; 5.464.763; 4.940.838; 4.693.976; 5.591.616; 5.231.019; 5.463.174; 4.762.785; 5.004.863; y 5.159.135; y las Solicitudes de Patente Europea N.ºs 0131624, 120516, 159418, 176112, 116718, 290799, 320500, 604662, 627752, 0267159, y 0292435.

Otros métodos de transformación de plantas incluyen transformación mediada por microproyector, en la que el ADN se conserva en la superficie de los microproyectiles. El vector de expresión se introduce en tejidos vegetales con un dispositivo biolístico que acelera los microproyectiles a velocidades suficientes para penetrar las paredes y

membranas de las células vegetales. Véase, por ejemplo, Sanford *et al.*, Part. Sci. Technol. 5: 27 (1987), Sanford, J. C., Trends Biotech. 6: 299 (1988), Sanford, J. C., Physiol. Plant 79: 206 (1990), Klein *et al.*, Biotechnology 10: 268 (1992), y los documentos de Patente de Estados Unidos N.ºs 5.015.580 y 5.322.783. También se describen técnicas para acelerar el material genético revestido sobre micropartículas dirigidas en células, por ejemplo, en los documentos de Patente de Estados Unidos N.ºs 4.945.050 y 5.141.141. Otro método para administración física de ADN a plantas es la sonicación de células diana. Véase, por ejemplo, Zhang *et al.*, Bio/Technology 9: 996 (1991). Como alternativa, para introducir vectores de expresión en plantas se ha usado fusión de liposomas o esferoplastos. Véase, por ejemplo, Deshayes *et al.*, EMBO J., 4: 2731 (1985), Christou *et al.*, Proc Natl. Acad. Sci. USA 84: 3962 (1987). La absorción directa de ADN en protoplastos usando precipitación de CaCl_2 , inyección de ADN, alcohol polivinílico o poli-L-ornitina también se informó. Véase, por ejemplo, Hain *et al.*, Mol. Gen. Genet. 199: 161 (1985) y Draper *et al.*, Plant Cell Physiol. 23: 451 (1982). También se ha descrito la electroporación de protoplastos y células y tejidos completos. Véase, por ejemplo, Donn *et al.*, en Abstracts of VIIth International Congress on Plant Cell and Tissue Culture IAPTC, A2-38, p. 53 (1990); D'Halluin *et al.*, Plant Cell 4: 1495-1505 (1992); Spencer *et al.*, Plant Mol. Biol. 24: 51-61 (1994); Publicaciones de Solicitud Internacional N.ºs WO 87/06614, WO 92/09696, y WO 93/21335; y en los documentos de Patente de Estados Unidos N.ºs 5.472.869 y 5.384.253. Otra tecnología de transformación incluye tecnología de partículas, véanse, por ejemplo, los documentos de Patente de Estados Unidos N.ºs 5.302.523 y 5.464.765.

Los cloroplastos o plástidos también se pueden transformar directamente. Como tal, se pueden producir plantas recombinantes en las que solamente se ha modificado el ADN de cloroplasto o plástido con cualquiera de las moléculas de ácido nucleico y moléculas de ácido nucleico recombinante descritas anteriormente así como combinaciones de las mismas. En la técnica se conocen promotores que funcionan en cloroplastos o plástidos. Véase, por ejemplo, Hanley-Bowden *et al.*, Trends in Biochemical Sciences 12: 67-70 (1987). Se han descrito métodos y composiciones para obtener células que contienen cloroplastos en las que se ha insertado ADN heterólogo, por ejemplo, en los documentos de Patente de Estados Unidos N.ºs 5.693.507 y 5.451.513.

También se puede usar cualquier otro método que proporcione una transformación eficaz.

En la técnica se conocen vectores adecuados para su uso en transformación de plantas. Véanse, por ejemplo, los documentos de Patente de Estados Unidos N.ºs 6.495.738; 7.271.315; 7, 348,473; 7.355.100; 7.456.340; y referencias desveladas en los mismos.

Los vectores de expresión pueden incluir al menos un marcador genético, unido de forma operativa a un elemento regulador (un promotor, por ejemplo) que permite que las células transformadas que contienen el marcador se recuperen mediante selección negativa, es decir, que inhiben el crecimiento de las células que lo contienen el gen marcador seleccionable, o mediante selección positiva, es decir, identificando sistemáticamente el producto codificado por el marcador genético. Muchos genes marcadores seleccionables usados comúnmente para transformación de plantas se conocen bien en las técnicas de transformación, e incluyen, por ejemplo, genes que codifican enzimas que detoxifican por vía metabólica un agente químico selectivo que puede ser un antibiótico o un herbicida, o genes que codifican una diana alterada que es insensible al inhibidor. Los marcadores seleccionables adecuados para su uso en transformación de plantas incluyen, pero no se limitan a, el gen de aminoglicósido fosfotransferasa del transposón Tn5 (Aph II) que codifican resistencia a los antibióticos kanamicina, neomicina, y G418, así como esos genes que codifican resistencia o tolerancia a herbicidas glifosato, higromicina, metotrexato, fosfotricina (bialophos), imidazolinonas, sulfonilureas y triazopirimidinas, tales como clorsulfurón, bromoxinilo, dalapón, y similares. Un gen marcador seleccionable usado comúnmente para transformación de plantas es el gen de neomicina fosfotransferasa II (nptII) bajo el control de señales reguladoras de plantas que confiere resistencia a la kanamicina. Véase, por ejemplo, Fraley *et al.*, Proc. Natl. Acad. Sci. U.S.A. 80: 4803 (1983). Otro gen de marcador seleccionable usado comúnmente es el gen de higromicina fosfotransferasa que confiere resistencia al antibiótico higromicina. Véase, por ejemplo, Vanden Elzen *et al.*, Plant Mol. Biol. 5: 299 (1985). Los genes de marcador seleccionable adicionales de origen bacteriano que confieren resistencia a antibióticos incluyen gentamicin acetil transferasa, estreptomycin fosfotransferasa, aminoglicósido-3'-adenil transferasa, y el determinante de resistencia a bleomicina. Véase, por ejemplo, Hayford *et al.*, Plant Physiol. 86: 1216 (1988), Jones *et al.*, Mol. Gen. Genet. 210: 86 (1987), Svab *et al.*, Plant Mol. Biol. 14: 197 (1990), Hille *et al.*, Plant Mol. Biol. 7: 171 (1986). Otros genes de marcador seleccionables confieren resistencia a herbicidas tales como glifosato, glufosinato, o bromoxinilo. Véase, por ejemplo, Comai *et al.*, Nature 317: 741-744 (1985), Gordon-Kamm *et al.*, Plant Cell 2: 603-618 (1990) y Stalker *et al.*, Science 242: 419-423 (1988). Otros genes de marcador seleccionables para transformación de plantas no son de origen bacteriano. Estos genes incluyen, por ejemplo, dihidrofolato reductasa de ratón, 5-enolpiruvilshikimato-3-fosfato sintasa vegetal y acetolactato sintasa vegetal. Véase, por ejemplo, Eichholtz *et al.*, Somatic Cell Mol. Genet. 13: 67 (1987), Shah *et al.*, Science 233: 478 (1986), Charest *et al.*, Plant Cell Rep. 8: 643 (1990).

Un gen indicador se puede usar con o sin un marcador seleccionable. Los genes indicadores son genes que por lo general no están presentes en el organismo tejido del receptor y por lo general codifican proteínas que dan como resultado algún cambio fenotípico o propiedad enzimática. Véase, por ejemplo, K. Weising *et al.*, Ann. Rev. Genetics 22: 421 (1988). Los genes indicadores incluyen, pero no se limitan a genes de beta-glucuronidasa (GUS), beta-galactosidasa, cloranfenicol acetiltransferasa, proteína fluorescente de color verde, y luciferasa. Véase, por ejemplo, Jefferson, R. A., Plant Mol. Biol. Rep. 5: 387 (1987), Teeri *et al.*, EMBO J. 8: 343 (1989), Koncz *et al.*, Proc. Natl.

Acad. Sci U.S.A. 84: 131 (1987), DeBlock *et al.*, EMBO J. 3: 1681 (1984), y Chalfie *et al.*, Science 263: 802 (1994). Un ensayo para detectar la expresión de genes indicadores se puede realizar en un momento adecuado después de haber introducido el gen en células receptoras. Uno de tales ensayos implica el uso de la beta-glucuronidasa codificante genética (GUS) del locus uidA de *E. coli* como se describe en Jefferson *et al.*, Biochem. Soc. Trans. 15: 17-19 (1987).

Los elementos reguladores del promotor de una diversidad de fuentes se pueden usar de manera eficaz en células vegetales para expresar genes extraños. Por ejemplo, se pueden usar elementos reguladores del promotor de origen bacteriano, tales como el promotor de octopina sintasa, el promotor de nopalina sintasa, el promotor de manopina sintasa, así como promotores de origen viral, tales como el virus del mosaico de la coliflor (35S y 19S), 35T (que es un promotor 35S reestructurado mediante ingeniería, véase la Publicación de Solicitud Internacional WO 97/13402). Los elementos reguladores del promotor de vegetales incluyen, pero no se limitan a, ribulosa-1,6-bisfosfato (RUBP) subunidad pequeña de carboxilasa (ssu), promotor de beta-conglicina, promotor de beta-faseolina, promotor de ADH, promotores de choque térmico y promotores específicos de tejidos. Para mejorar la eficacia de la transcripción o la integración de ADN también se pueden usar regiones de unión a matriz, regiones de unión de armazón, intrones, potenciadores, y secuencias de poliadenilación. Los elementos de este tipo se pueden incluir para obtener un rendimiento óptimo del ADN transformado en la planta. Los elementos habituales incluyen, pero no se limitan a, Adh-intrón 1, Adh-intrón 6, la secuencia líder de la proteína de la cubierta del virus del mosaico de la alfalfa, la secuencia líder de la proteína de la cubierta del virus de la raíz de maíz, así como otras secuencias disponibles para un experto en la materia. También se pueden usar elementos reguladores del promotor constitutivo para dirigir la expresión genética continua. Los promotores constitutivos incluyen, pero no se limitan a, promotores de virus de plantas tales como el promotor 35S de CaMV (Odell *et al.*, Nature 313: 810-812 (1985)), y promotores de tales genes tales como actina de arroz (McElroy *et al.*, Plant Cell 2: 163-171 (1990)), ubiquitina (Christensen *et al.*, Plant Mol. Biol. 12: 619-632 (1989) y Christensen *et al.*, Plant Mol. Biol. 18: 675-689 (1992)), pEMU (Last *et al.*, Theor. Appl. Genet. 81: 581-588 (1991)), MAS (Velten *et al.*, EMBO J. 3: 2723-2730 (1984)), histona H3 de maíz (Lepetit *et al.*, Mol. Gen. Genetics 231: 276-285 (1992) y Atanassova *et al.*, Plant Journal 2(3): 291-300 (1992)), y el promotor de ALS, fragmento en la posición 5' de Xba1/NcoI con respecto al gen estructural ALS3 de *Brassica napus* (o una secuencia de nucleótidos similar a la del fragmento Xba1/NcoI) (Publicación de Solicitud Internacional N.º WO 96/30530). Para la expresión genética en tipos específicos de células o tejidos también se pueden usar elementos reguladores del promotor específico de tejidos, tales como hojas o semillas (por ejemplo, zeína, oleosina, napina, ACP, globulina, y similares). Los promotores específicos de tejidos o preferentes para tejidos incluyen, pero no se limitan a, un promotor preferente de raíz, tal como del gen de faseolina (Murai *et al.*, Science 23: 476-482 (1983) y Sengupta-Gopalan *et al.*, Proc. Natl. Acad. Sci. U.S.A. 82: 3320-3324 (1985)); un promotor específico de la hoja e inducido por la luz, tal como de repollo o rubisco (Simpson *et al.*, EMBO J. 4 (11): 2723-2729 (1985) y Timko *et al.*, Nature 318: 579-582 (1985)); un promotor específico de la antera tal como de LAT52 (Twell *et al.*, Mol. Gen. Genetics 217: 240-245 (1989)); un promotor específico del polen tal como de Zm13 (Guerrero *et al.*, Mol. Gen. Genetics 244: 161-168 (1993)); o un promotor preferente de microesporas tal como de apg (Twell *et al.*, Sex. Plant Reprod. 6: 217-224 (1993)). Los elementos reguladores del promotor también pueden estar activos durante una cierta etapa del desarrollo de las plantas, así como tejidos y órganos de las plantas, que incluyen, pero no se limitan a, elementos reguladores del promotor específicos de polen, específicos de embrión, específicos de la seda de maíz, específicos de la fibra de algodón, específicos de la raíz, y específicos de endospermo de semilla. Se puede usar un elemento regulador promotor inducible, que es responsable de la expresión de genes a una respuesta a una señal específica, tales como: estímulo físico (genes de choque térmico); luz (RUBP carboxilasa); hormona (Em); metabolitos; agentes químicos; y estrés. Los promotores inducibles incluyen, pero no se limitan a, un promotor del sistema ACEI que responde al cobre (Mett *et al.*, PNAS 90: 4567-4571 (1993)); del gen In2 del maíz que responde a fitoprotectores herbicidas de bencenosulfonamida (Hershey *et al.*, Mol. Gen. Genetics 227: 229-237 (1991) y Gatz *et al.*, Mol. Gen. Genetics 243: 32-38 (1994)), del represor Tet de Tn10 (Gatz *et al.*, Mol. Gen. Genetics 227: 229-237 (1991)); y de un gen de hormona esteroidea, cuya actividad de transcripción está inducida por una hormona glucocorticosteroidea (Schena *et al.*, Proc. Natl. Acad. Sci. U.S.A. 88: 0421 (1991)).

Para dirigir un polipéptido a cualquiera de un orgánulo intracelular o compartimento subcelular o para secreción del apoplasto también se pueden usar secuencias señal. Véase, por ejemplo, Becker *et al.*, Plant Mol. Biol. 20: 49 (1992), Knox, C., *et al.*, Plant Mol. Biol. 9: 3-17 (1987), Lerner *et al.*, Plant Physiol. 91: 124-129 (1989), Fontes *et al.*, Plant Cell 3: 483-496 (1991), Matsuoka *et al.*, Proc. Natl. Acad. Sci. 88: 834 (1991), Gould *et al.*, J. Cell. Biol. 108: 1657 (1989), Creissen *et al.*, Plant J. 2: 129 (1991), Kalderon, *et al.*, Cell 39: 499-509 (1984), y Steifel *et al.*, Plant Cell 2: 785-793 (1990). Las secuencias de dirección de este tipo proporcionan la proteína expresada deseada a transferir a la estructura celular en la que funciona de la manera más eficaz o a áreas de la célula en las que se concentran los procesos celulares necesarios para las funciones fenotípicas deseadas.

En algunas realizaciones, se usan secuencias señal para dirigir proteínas de la invención a un compartimento subcelular, por ejemplo, al plástido o cloroplasto. Los productos genéticos, incluyendo productos genéticos heterólogos, se pueden dirigir al plástido o cloroplasto fusionando el producto genético con una secuencia señal que se extiende durante la importación de cloroplasto proporcionando la proteína madura. Véase, por ejemplo, Comai *et al.*, J. Biol. Chem. 263: 15104-15109 (1988) y van den Broeck *et al.*, Nature 313: 358-363 (1985). El ADN que codifican secuencias señal apropiadas se puede aislar de los ADNc que codifican la proteína RUBISCO, la proteína CAB, la enzima EPSP sintasa, la proteína GS2, o de cualquier proteína dirigida a cloroplasto de origen natural que

contenga una secuencia señal (también denominado péptido de tránsito de cloroplasto (CTP)) que dirige la proteína dirigida a cloroplasto. Las proteínas dirigidas a cloroplasto de este tipo se conocen bien en la técnica. Las proteínas dirigidas a cloroplasto se sintetizan como proteínas precursoras más grandes que contienen un CTP amino-terminal, que dirige el precursor a la maquinaria de importación de cloroplasto. Por lo general, los CTP se escinden con endoproteasas específicas situadas dentro del orgánulo del cloroplasto, liberando de ese modo la proteína madura dirigida, incluyendo proteínas activas tales como enzimas, del precursor en el medio del cloroplasto. Los ejemplos de secuencias que codifican péptidos adecuados para dirigir un gen o producto genético al cloroplasto o plástido o de la célula vegetal incluyen el CTP de EPSPS de petunia, el CTP2 e intrón de EPSPS de *Arabidopsis*, y otras secuencias conocidas en la técnica. Los ejemplos específicos de CTP incluyen, pero no se limitan a, el péptido de tránsito de ats1A de subunidad pequeña de ribulosa bisfosfato de *Arabidopsis thaliana*, un péptido de tránsito de EPSPS de *Arabidopsis thaliana*, y un péptido de tránsito de subunidad pequeña de ribulosa bisfosfato carboxilasa de maíz Zea. Un péptido de tránsito optimizado se describe, por ejemplo, en Van den Broeck *et al.*, Nature 313: 358-363 (1985). Las secuencias de señal procariotas y eucariotas se desvelan, por ejemplo, en Michaelis *et al.*, Ann. Rev. Microbiol. 36: 425 (1982). Los ejemplos adicionales de péptidos de tránsito que se pueden usar en la invención incluyen péptidos de tránsito de cloroplasto descritos en Von Heijne *et al.*, Plant Mol. Biol. Rep. 9: 104-126 (1991); Mazur *et al.*, Plant Physiol. 85: 1110 (1987); Vorst *et al.*, Gene 65: 59 (1988); Chen y Jagendorf, J. Biol. Chem. 268: 2363-2367 (1993); un péptido de tránsito del gen rbcS de *Nicotiana plumbaginifolia* (Poulsen *et al.*, Mol. Gen. Genet. 205: 193-200 (1986)); y un péptido de tránsito obtenido a partir de acil-ACP tioesterasa de *Brassica napus* (Loader *et al.*, Plant Mol. Biol. 23: 769-778 (1993); Loader *et al.*, Plant Physiol. 110: 336-336 (1995).

Las plantas modificadas genéticamente se prevé modificar adicionalmente para producir delección o inactiva una ácido graso sintasa endógena, para reducir la competición endógena con el sistema de PUFA sintasa exógena para malonil CoA, para aumentar el nivel de malonil CoA en el organismo, y combinaciones de los mismos. Véase, por ejemplo, la Publicación de Solicitud de Estados Unidos N.º 2007/0245431.

Una planta modificada genéticamente se puede cultivar en un medio de fermentación o puede crecer en un medio de adecuado tal como el suelo. Un medio de crecimiento adecuado para plantas superiores f incluye cualquier medio de crecimiento para plantas, tal como, pero no se limita a, suelo, arena, cualquier otro medio de partículas que soporten el crecimiento de la raíz (por ejemplo, vermiculita, perlita, etc.) o cultivo hidropónico así como luz, agua, y suplementos nutricionales adecuados que optimicen el crecimiento de la planta superior. Los PUFA se pueden recuperar de las plantas modificadas genéticamente a través de procesos de purificación que extraen los compuestos de la planta. Los PUFA se pueden recuperar cosechando la planta así como cosechando la aceite de la planta (por ejemplo, de las semillas oleaginosas). La planta también se puede consumir en su estado natural o procesar adicionalmente en productos consumibles. En algunas realizaciones, la presente invención se refiere a una planta modificada genéticamente, en la que la planta produce al menos un PUFA como resultado de la modificación genética, y en la que el perfil de ácido graso total en la planta, o la parte de la planta que acumula los PUFA, comprende la cantidad detectable del PUFA producido como resultado de la modificación genética de la planta. En algunas realizaciones, la planta es una planta de semillas oleaginosas. En algunas realizaciones, la planta de semillas oleaginosas produce los PUFA en sus semillas maduras o contiene los PUFA en el aceite de sus semillas.

Para expresar proteína recombinante también se pueden usar diversos sistemas de cultivo de células de mamífero. Los vectores de expresión comprenderán un origen de replicación, un promotor y potenciador adecuados, y también cualquier sitio de unión a ribosoma necesario, sitio de poliadenilación, sitios donantes y receptores de corte y empalme, secuencias de terminación de la transcripción, y secuencias no transcritas de flanco en la posición 5'.

Métodos que implican expresión heteróloga

La presente invención se refiere a un método para producir al menos un PUFA que comprende la expresión de un sistema de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir PUFA, en el que el sistema de PUFA sintasa comprende cualquiera de las moléculas de ácido nucleico aisladas y las moléculas de ácido nucleico recombinante descritas en el presente documento así como combinaciones de las mismas, en el que se produce al menos un PUFA. En algunas realizaciones, el al menos un PUFA incluye DHA, EPA, o una combinación de los mismos. En algunas realizaciones, la célula hospedadora es una célula vegetal, una célula animal aislada, o una célula microbiana. En algunas realizaciones la célula hospedadora es un traustóquitrdo.

La presente invención se refiere a un método para producir lípidos enriquecidos con DHA, EPA, o una combinación de los mismos, que comprende expresar un gen de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir lípidos, en el que el gen de PUFA sintasa comprende cualquiera de las moléculas de ácido nucleico aisladas y las moléculas de ácido nucleico recombinante descritas en el presente documento así como combinaciones de las mismas en la célula hospedadora, en el que se producen lípidos enriquecidos con DHA, EPA, o una combinación de los mismos.

La invención se refiere a un método para aislar lípidos a partir de una célula hospedadora, que comprende expresar un gen de PUFA sintasa en la célula hospedadora en condiciones eficaces para producir lípidos, y aislar lípidos a partir de la célula hospedadora, en el que el sistema de PUFA sintasa en la célula hospedadora comprende cualquiera de las moléculas de ácido nucleico aisladas y las moléculas de ácido nucleico recombinante descritas en

el presente documento así como combinaciones de las mismas.

En algunas realizaciones, una o más fracciones lipídicas que contienen los PUFA se aíslan de las células hospedadoras. En algunas realizaciones, la una o más fracciones aisladas de la célula hospedadora incluyen la fracción total de ácido graso, la fracción de ésteres de esteroles, la fracción de triglicérido, la fracción de ácido graso libre, la fracción de esteroles, la fracción de diglicerol, la fracción de fosfolípido, o combinación de los mismos. En algunas realizaciones, los PUFA se aíslan de las células hospedadoras, en las que los PUFA están enriquecidos con ácidos grasos omega-3 ácidos, grasos omega-6, o combinaciones de los mismos basándose en la composición del sistema de PUFA sintasa introducido en una célula hospedadora. En algunas realizaciones, dos PUFA están enriquecidos con DHA, EPA, DPA n-6, ARA, o combinaciones de los mismos basándose en la composición del sistema de PUFA sintasa introducido en una célula hospedadora. En algunas realizaciones, los PUFA están enriquecidos con DHA, EPA, o una combinación de los mismos. En algunas realizaciones, el perfil de PUFA de los PUFA aislados de una célula hospedadora incluye concentraciones elevadas de DHA y concentraciones más bajas de EPA, ARA, DPA n-6, o combinaciones de los mismos. En algunas realizaciones, el perfil de PUFA de los PUFA aislados de una célula hospedadora incluye concentraciones elevadas de DHA y EPA, y concentraciones más bajas de ARA, DPA n-6, o combinaciones de los mismos. En algunas realizaciones, el perfil de PUFA de los PUFA aislados de una célula hospedadora incluye concentraciones elevadas de EPA y concentraciones más bajas de DHA, ARA, DPA n-6, o combinaciones de los mismos.

La invención se refiere a un método para reemplazar una actividad de PUFA sintasa inactiva o con delección, introducir una nueva actividad de PUFA sintasa, o mejorar una actividad de PUFA sintasa existente en un organismo que tiene actividad de PUFA sintasa, que comprende la expresión de cualquiera de las moléculas de ácido nucleico aisladas y las moléculas de ácido nucleico recombinante descritas en el presente documento así como combinaciones de las mismas en el organismo en condiciones eficaces para expresar la actividad de PUFA sintasa. En algunas realizaciones, la molécula de ácido nucleico comprende una o más secuencias de polinucleótido de PUFA sintasa de *PFA1*, *PFA2*, o *PFA3* descritas en el presente documento que codifican uno o más dominios de PUFA sintasa. En algunas realizaciones, los perfiles de PUFA de los organismos se alteran mediante la introducción de la una o más moléculas de ácido nucleico de la invención. En algunas realizaciones, los perfiles de PUFA alterados incluyen un aumento de ácidos grasos omega-3 y una disminución de ácidos grasos omega-6. En algunas realizaciones, los perfiles de PUFA alterados incluyen un aumento de ácidos grasos omega-6 y una disminución de ácidos grasos omega-3. En algunas realizaciones, los ácidos grasos tanto omega-3 como omega-6 aumentan. En algunas realizaciones, la cantidad de DHA aumenta mientras que las cantidades de uno o más de EPA, ARA, DPA n-6, o combinaciones de los mismos se mantienen o disminuyen. En algunas realizaciones, las cantidades de EPA y DHA aumentan mientras que las cantidades de ARA, DPA n-6, o una combinación de los mismos se mantienen o disminuyen. En algunas realizaciones, la cantidad de EPA aumenta mientras que las cantidades de uno o más de EPA, ARA, DPA n-6, o combinaciones de los mismos se mantienen o disminuyen. En algunas realizaciones, la molécula de ácido nucleico comprende la secuencia de polinucleótidos de *PFA3* o uno o más dominios en la misma. En algunas realizaciones, la molécula de ácido nucleico comprende la secuencia de polinucleótidos de *PFA3* o uno o más dominios en la misma y la cantidad de ácidos grasos omega-3 en el organismo aumenta mientras que la cantidad de ácidos grasos omega-6 disminuye. En algunas realizaciones, la molécula de ácido nucleico comprende la secuencia de polinucleótidos de *PFA2* o uno o más dominios en la misma y la cantidad de DHA en el organismo aumenta mientras que la cantidad de EPA disminuye.

La invención se refiere a métodos para aumentar la producción de DHA, EPA, o una combinación de los mismos en un organismo que tiene actividad de PUFA sintasa, que comprende la expresión de cualquiera de las moléculas de ácido nucleico aisladas y moléculas de ácido nucleico recombinante descritas en el presente documento así como combinaciones de las mismas en el organismo en condiciones eficaces para producir DHA, EPA, o una combinación de los mismos, en los que la actividad de PUFA sintasa reemplaza una actividad inactiva o con delección, introduce una nueva actividad, o mejor a una actividad existente en el organismo, y en los que la producción de DHA, EPA, o una combinación de los mismos en el organismo aumenta.

Por lo general habiendo descrito la presente invención, se puede obtener una comprensión adicional por referencia a los ejemplos que se proporcionan en el presente documento. Estos ejemplos son solo para fines de ilustración y no pretenden ser limitantes.

EJEMPLO 1

Se diseñaron cebadores degenerados para los dominios de KS y DH de PUFA sintasa para aislar las secuencias correspondientes a partir del microorganismo aislado depositado con el N.º PTA-9695 de Registro en la ATCC, también conocido como cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Se diseñaron cebadores degenerados para la región de KS de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (es decir, la región que contiene el dominio de KS) basándose en las secuencias publicadas de *PFA1* (denominadas previamente *orfA* u ORF 1) para *Shewanella japonica*, la cepa 20888 de la ATCC de *Schizochytrium sp.*, *Thraustochytrium aureum* (ATCC 34304), y ATCC 20892 de *Thraustochytrium sp.* 23B:

prDS173 (directo): GATCTACTGCAAGCGCGGNGGNTTYAT (SEQ ID NO: 62), y
prDS174 (inverso): GGCGCAGGCGGCRTCNACNAC (SEQ ID NO: 63).

Se diseñaron cebadores degenerados para la región de DH de PFA3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (denominado previamente *orfC* u ORF 3) basándose en las secuencias publicadas para *Moritella marina*; la cepa 20888 de la ATCC de *Schizochytrium sp.*; SCRC-2738 de *Shewanella sp.*; *Photobacter profundum*; y23B ATCC 20892 de *Thraustochytrium sp.*:

JGM190 (directo): CAYTGGTAYTTYCCNTGYCAYTT (SEQ ID NO: 64); y
BLR242 (inverso): CCNGGCATNACNGGRTC (SEQ ID NO: 65).

Las condiciones de PCR con molde de ADN cromosómico fueron las que siguen a continuación: dNTPs 0,2 µM, 0,1 uM cada cebador, DMSO al 8 %, 200 ng de ADN cromosómico, 2,5 U de polimerasa de fusión Herculase® II (Stratagene), y 1X de tampón Herculase® (Stratagene) en un volumen total de 50 µl. El Protocolo de PCR incluía las siguientes etapas: (1) 98 °C durante 3 minutos; (2) 98 °C durante 30 segundos; (3) 50 °C durante 30 segundos; (4) 72 °C durante 2 minutos; (5) etapas 2-4 repetidas durante 40 ciclos; (6) 72 °C durante 5 minutos; y (7) se mantiene a 6 °C.

Para ambos pares de cebadores, la PCR proporcionó productos de ADN distintos con los tamaños esperados usando moldes cromosómicos a partir de la cepa de *Schizochytrium sp.*, con N.º PTA-9695 de Registro en la ATCC. Los respectivos productos de PCR se clonaron en el vector pJET1.2/blunt (Fermentas) de acuerdo con las instrucciones del fabricante, y la secuencia del inserto se determinó usando cebadores convencionales suministrados.

Las secuencias de ADN obtenidas a partir de los productos de PCR se compararon con secuencias conocidas disponibles a partir del GenBank de NCBI en una búsqueda de BLASTx convencional (parámetros de BLASTx: Filtro de baja complejidad activado; Matriz: BLOSUM62; Coste de hueco; Existence 11, Extension 1. Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, y David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25: 3389-3402.).

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos deducida obtenida a partir del ADN clonado que contenía el fragmento de KS de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, fueron: cepa 20888 de la ATCC de *Schizochytrium sp.*, "subunidad A de ácido graso poliinsaturado sintasa" (Identidad = 87 %; positivos = 92 %); *Shewanella oneidensis* MR-1 "beta-cetoacil sintasa de múltiples dominios" (Identidad = 49 %; positivos = 64 %); y *Shewanella sp. MR-4* "beta-cetoacil sintasa" (Identidad = 49 %; positivos = 64 %).

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos deducida obtenida a partir del ADN clonado que contenía el fragmento de DH de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, fueron: cepa 20888 de la ATCC de *Schizochytrium sp.*, "subunidad C de ácido graso poliinsaturado sintasa" (Identidad = 61 %; positivos = 71 %); cepa 700345 de la ATCC de *Shewanella pealeana* "Beta-hidroxilacil-(proteína transportadora de acilo) deshidratasa FabA/FabZ" (Identidad = 35 %; positivos = 50 %); y *Shewanella sediminis* HAW-EB3 "PfaC de ácido graso poliinsaturado sintasa omega-3" (Identidad = 34 %; positivos = 50 %).

EJEMPLO 2

Los genes de PUFA sintasa se identificaron a partir de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

El ADN genómico se preparó a partir del microorganismo mediante procedimientos convencionales. Véase, por ejemplo, Sambrook J. y Russell D. 2001. Molecular cloning: A laboratory manual, 3ª edición. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. En resumen: (1) se sedimentaron 500 µl de células a partir del cultivo de logaritmo medio. Las células se volvieron a centrifugar, y todas las trazas de líquido se retiraron del sedimento celular con una punta de calibre pequeño; (2) los sedimentos se volvieron a suspender con 200 µl de tampón de lisis (Tris 20 mM a pH 8,0, 125 µg/ml de Proteínas K, NaCl 50 mM, EDTA 10 mM a pH 8,0, SDS al 0,5 %); (3) las células se lisaron a 50 °C durante 1 hora; (4) la mezcla de lisis se pipeteó en tubos de 2 ml de gel en fase de bloqueo (PLG -Eppendorf); (5) se añadió un volumen igual de P:C:I y se permitió la mezcla durante 1,5 horas; (6) los tubos se centrifugaron a 12k x g durante 5 minutos; (7) la fase acuosa se retiró del gel mencionado anteriormente dentro del tubo de PLG y se añadió un volumen igual de cloroformo a la fase acuosa, y se mezcló durante 30 minutos; (8) los tubos centrifugaron a 14k durante aproximadamente 5 minutos; (9) la fase superior (fase acuosa) se retiró mediante pipeteo del cloroformo, y se colocó en un nuevo tubo; (10) se añadieron 0,1 volúmenes de NaOAC 3 M se y se mezcló (se invirtió varias veces); (11) se añadieron 2 volúmenes de EtOH al 100 % y se mezcló (se invirtió varias veces) con precipitante de ADN genómico formándose en esta etapa; (12) los tubos se centrifugaron a 4 °C en una microcentrifugadora a 14k durante aproximadamente 15 minutos; (13) el líquido se vertió suavemente con el ADN genómico restante en el fondo del tubo; (14) del sedimento se lavó con 0,5 ml de EtOH al

70 %; (15) los tubos centrifugaron a 4 °C en una microcentrifugadora a 14k durante aproximadamente 5 minutos; (16) el EtOH se vertió suavemente y el sedimento de ADN genómico se secó; y (17) se añadió un volumen adecuado de H₂O y RNasa directamente al sedimento de ADN genómico.

El ADN genómico aislado se usó para generar bibliotecas recombinantes que consisten en fragmentos grandes (de aproximadamente 40 kB) de acuerdo con las instrucciones del fabricante en el pWEB-TNC™ cósmido (Epicentre). Las bibliotecas de cósmidos se identificaron sistemáticamente mediante procedimientos de hibridación de colonias convencionales usando sondas etiquetadas radiactivamente con ³²P (Sambrook J. y Russell D. 2001. Molecular cloning: A laboratory manual, 3ª edición. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). Las sondas contenían ADN homólogo a las secuencias de PUFA sintasa publicadas de otros organismos como se describe en el Ejemplo 1. Estas sondas se generaron mediante una digestión de restricción de ADN de los fragmentos clonados a partir de los respectivos clones de pJET1.2/blunt descritos anteriormente y se etiquetaron con métodos convencionales. En todos los casos, la fuerte hibridación de las sondas individuales para ciertos cósmidos indicaban clones que contenían ADN homólogo a los genes de PUFA sintasa.

El clon pDS115 cósmido demostraba una fuerte hibridación de la sonda a la región de KS y se seleccionó para secuenciación de ADN del gen de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* El clon pDS115 cósmido, que contenía los genes de *PFA1* y *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 27 de enero de 2009, y se le dio el N.º PTA-9737 de Registro en la ATCC. Se diseñaron cebadores de secuenciación para la secuencia de ADN de la región de KS determinada en el Ejemplo 1 usando métodos convencionales. Para determinar la secuencia de ADN de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se realizaron rondas sucesivas de secuenciación de ADN, que implicaban el posterior diseño de cebador de secuenciación con métodos convencionales, para "acompañar" al clon cósmido.

En los sistemas de PUFA sintasa de traustóquitrado publicados previamente, los genes de *PFA1* y *PFA2* de PUFA sintasa se agruparon en conjunto y se colocaron para que se transcribieran de manera divergente. Este también es el caso para *PFA1* y *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* A través del "acompañamiento" de la secuencia de ADN del clon cósmido pDS115, se encontró que el inicio conceptual de *PFA2* tenía 493 nucleótidos desde el inicio de *PFA1* y se transcribía de manera divergente. Cada par de bases de nucleótidos de los genes de PUFA sintasa de *PFA1* y *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se cubrieron con al menos dos reacciones de secuenciación de ADN separadas con alta calidad con al menos una puntuación Phred mínima agregada de 40 (nivel de confianza de un 99,99 %).

El clon pBS4 cósmido demostraba una fuerte hibridación de la sonda a la región de DH y se seleccionó para secuenciación de ADN del gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* El clon pBS4 cósmido, que contiene el gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 27 de enero de 2009, y se le dio el N.º de Registro en la ATCC PTA-9736. Se diseñaron cebadores de secuenciación usando métodos convencionales para la secuencia de ADN de la región de DH determinada en el Ejemplo 1. Para determinar la secuencia de ADN de el gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se realizaron rondas sucesivas de secuenciación de ADN, que implicaban el posterior diseño de cebador de secuenciación con métodos convencionales, para "acompañar" al clon cósmido. Cada par de bases de nucleótidos de el gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* se cubrió con al menos dos reacciones de secuenciación de ADN separadas de alta calidad con al menos una puntuación Phred mínima agregada de 40 (nivel de confianza de un 99,99 %).

La Tabla 1 muestra identidades para las secuencias de polinucleótidos de *PFA1* (SEQ ID NO: 1), *PFA2* (SEQ ID NO: 3), y *PFA3* (SEQ ID NO: 5) de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en comparación con secuencias publicadas previamente. Las identidades se determinaron con la matriz de puntuación "swgapdnamt" dentro del programa "AlignX" del programa VectorNTI, un patrón para alineamiento de ADN.

Tabla 1: Porcentaje de Identidad para Secuencias de Polinucleótidos de *PFA1*, *PFA2*, y *PFA3*

Fuente de Secuencias de <i>PFA1</i> , <i>PFA2</i> , y <i>PFA3</i> Publicadas	% de Identidad de <i>PFA1</i> (<i>orfA</i>) a <i>PFA1</i> publicadas (SEQ ID NO: 1)	% de Identidad de <i>PFA2</i> (<i>orfB</i>) a <i>PFA2</i> publicadas (SEQ ID NO: 3)	% de Identidad de <i>PFA3</i> (<i>orfC</i>) a <i>PFA3</i> publicadas (SEQ ID NO: 5)
ATCC 20888 de <i>Schizochytrium sp.</i>	70	66	75
ATCC 34304 de <i>Thraustochytrium aureum</i>	65	62	no publicado
23B ATCC 20892 de <i>Thraustochytrium sp.</i>	56	55	67

La Tabla 2 muestra identidades para las secuencias de aminoácidos de Pfa1p (SEQ ID NO: 2), Pfa2p (SEQ ID NO: 4), y Pfa3p (SEQ ID NO: 6) de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en comparación con secuencias de aminoácidos de PUFA sintasa publicadas previamente. Las identidades se determinaron a través del uso de la matriz de puntuación "blosum62mt2" dentro del programa "AlignX" del programa VectorNTI, un patrón para alineamiento de proteínas.

Tabla 2: Porcentaje de Identidad para Secuencias de Aminoácidos de Pfa1p, Pfa2p, y Pfa3p

Fuente de Secuencias de Pfa1p, Pfa2p, y Pfa3p Publicadas	% de Identidad de Pfa1p (OrfA) a Pfa1p publicadas (SEQ ID NO: 2)	% de Identidad de Pfa2p (OrfB) a Pfa2p publicadas (SEQ ID NO: 4)	% de Identidad de Pfa3p (OrfC) a Pfa3p publicadas (SEQ ID NO: 6)
ATCC 20888 de <i>Schizochytrium sp.</i>	60	53	70
ATCC 34304 de <i>Thraustochytrium aureum</i>	60	54	no publicado
23B ATCC 20892 de <i>Thraustochytrium sp.</i>	52	52	70

EJEMPLO 3

El análisis de dominios se realizó para anotar las coordenadas de la secuencia para los dominios de PUFA sintasa y sitios activos de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, respectivamente. Los dominios identificaron basándose en homología con respecto a dominios conocidos de PUFA sintasa, ácido graso sintasa, y policéptido sintasa.

La Tabla 3 muestra los dominios y los sitios activos asociados con *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Tabla 3: Análisis de Dominio de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
KS	7-1401 de SEQ ID NO: 1 (SEQ ID NO: 7)	3-467 de SEQ ID NO: 2 (SEQ ID NO: 8)	Activo - DXAC* (SEQ ID NO: 43)	607-609 de SEQ ID NO: 1	C203 de SEQ ID NO: 2
			Terminal - GFGG (SEQ ID NO: 44)	1363-1374 de SEQ ID NO: 1 (SEQ ID NO: 45)	455-458 de SEQ ID NO: 2
MAT	1798-2700 de SEQ ID NO: 1 (SEQ ID NO: 9)	600-900 de SEQ ID NO: 2 (SEQ ID NO: 10)	GHS*LG Activo (SEQ ID NO: 46)	2095-2097 de SEQ ID NO: 1	S699 de SEQ ID NO: 2
ACP	3298-5400 de SEQ ID NO: 1 (SEQ ID NO: 11)	1100-1800 de SEQ ID NO: 2 (SEQ ID NO: 12)	Dominio de ACP1	3325-3600 de SEQ ID NO: 1 (SEQ ID NO: 13)	1109-1200 de SEQ ID NO: 2 (SEQ ID NO: 14)
			ACP1 Activo LGIDS* (SEQ ID NO: 47)	3454-3456 de SEQ ID NO: 1	S1152 de SEQ ID NO: 2

ES 2 623 610 T3

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
			Dominio de ACP2	3667-3942 de SEQ ID NO: 1 (SEQ ID NO: 15)	1223-1314 de SEQ ID NO: 2 (SEQ ID NO: 16)
			ACP2 Activo LGIDS* (SEQ ID NO: 47)	3796-3798 de SEQ ID NO: 1	S1266 de SEQ ID NO: 2
			Dominio de ACP3	4015-4290 de SEQ ID NO: 1 (SEQ ID NO: 17)	1339-1430 de SEQ ID NO: 2 (SEQ ID NO: 18)
			ACP3 Activo LGIDS* (SEQ ID NO: 47)	4144-4146 de SEQ ID NO: 1	S1382 de SEQ ID NO: 2
			Dominio de ACP4	4363-4638 de SEQ ID NO: 1 (SEQ ID NO: 19)	1455-1546 de SEQ ID NO: 2 (SEQ ID NO: 20)
			ACP4 Activo LGIDS* (SEQ ID NO: 47)	4492-4494 de SEQ ID NO: 1	S1498 de SEQ ID NO: 2
			Dominio de ACP5	4711-4986 de SEQ ID NO: 1 (SEQ ID NO: 21)	1571-1662 de SEQ ID NO: 2 (SEQ ID NO: 22)
			ACP5 Activo LGIDS* (SEQ ID NO: 47)	4840-4842 de SEQ ID NO: 1	S1614 de SEQ ID NO: 2
			Dominio de ACP6	5053-5328 de SEQ ID NO: 1 (SEQ ID NO: 23)	1685-1776 de SEQ ID NO: 2 (SEQ ID NO: 24)
			ACP6 Activo LGIDS* (SEQ ID NO: 47)	5182-5184 de SEQ ID NO: 1	S1728 de SEQ ID NO: 2
KR	5623-7800 de SEQ ID NO: 1 (SEQ ID NO: 25)	1875-2600 de SEQ ID NO: 2 (SEQ ID NO: 26)	"región núcleo"	5998-6900 de SEQ ID NO: 1 (SEQ ID NO: 48)	2000-2300 de SEQ ID NO: 2 (SEQ ID NO: 49)

Dominio	Posición de ADN	Posición de AA	Símbolos	Posición de ADN	Posición de AA
Motivo de DH	7027-7065 de SEQ ID NO: 1	2343-2355 de SEQ ID NO: 2	LxxHxxxGxxxxP	7027-7065 de SEQ ID NO: 1	2343-2355 de SEQ ID NO: 2
	(SEQ ID NO: 27)	(SEQ ID NO: 28)	(SEQ ID NO: 50)	(SEQ ID NO: 27)	(SEQ ID NO: 28)

El primer dominio en Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, es un dominio de KS. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de KS de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 7, que corresponde a las posiciones 7-1401 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene el dominio de KS de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 8, que corresponde a las posiciones 3-467 de SEQ ID NO: 2. El dominio de KS contiene un motivo de sitio activo: DXAC* (SEQ ID NO: 43), con un *sitio de unión a acilo que corresponde a C203 de SEQ ID NO: 2. Además, un motivo característico está presente en el extremo del dominio de KS: GFGG (SEQ ID NO: 44), que corresponde a las posiciones 455-458 de SEQ ID NO: 2 y a las posiciones 453-456 de SEQ ID NO: 8.

El segundo dominio en Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, es un dominio MAT. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio MAT de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 9, que corresponde a las posiciones 1798-2700 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene el dominio MAT de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 10, que corresponde a las posiciones 600-900 de SEQ ID NO: 2. El dominio MAT contiene un motivo de sitio activo: GHS*²XG (SEQ ID NO: 46), con un *sitio de unión a acilo que corresponde a S699 de SEQ ID NO: 2.

Los dominios tercero a octavo de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, son seis dominios de ACP en tándem, también denominados ACP1, ACP2, ACP3, ACP4, ACP5, y ACP6 en el presente documento. La secuencia de nucleótidos que contiene el primer dominio de ACP, ACP1, en el presente documento se representa como SEQ ID NO: 13 y está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 3325 a aproximadamente la posición 3600 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP1, representada en el presente documento como SEQ ID NO: 14, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1109 a aproximadamente la posición 1200 de SEQ ID NO: 2. La secuencia de nucleótidos que contiene ACP2, representada en el presente documento como SEQ ID NO: 15, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 3667 a aproximadamente la posición 3942 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP2, representada en el presente documento como SEQ ID NO: 16, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1223 a aproximadamente la posición 1314 de SEQ ID NO: 2. La secuencia de nucleótidos que contiene ACP3, representada en el presente documento como SEQ ID NO: 17, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4015 a aproximadamente la posición 4290 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP3, representada en el presente documento como SEQ ID NO: 18, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1339 a aproximadamente la posición 1430 de SEQ ID NO: 2. La secuencia de nucleótidos que contiene ACP4, representada en el presente documento como SEQ ID NO: 19, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4363 a aproximadamente la posición 4638 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP4, representada en el presente documento como SEQ ID NO: 20, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1455 a aproximadamente la posición 1546 de SEQ ID NO: 2. La secuencia de nucleótidos que contiene ACP5, representada en el presente documento como SEQ ID NO: 21, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4711 a aproximadamente la posición 4986 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP5, representada en el presente documento como SEQ ID NO: 22, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1571 a aproximadamente la posición 1662 de SEQ ID NO: 2. La secuencia de nucleótidos que contiene ACP6, representada en el presente documento como SEQ ID NO: 23, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 5053 a aproximadamente la posición 5328 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene ACP6, representada en el presente documento como SEQ ID NO: 24, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1685 a aproximadamente la posición 1776 de SEQ ID NO: 2. Los seis dominios de ACP en conjunto se extienden a una región de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, desde aproximadamente la posición 3298 a aproximadamente la posición 5400 de SEQ ID NO: 1, que corresponde a posiciones del aminoácido de aproximadamente 1100 a aproximadamente 1800 de SEQ ID NO: 2. La secuencia de nucleótidos para toda la región de ACP que contiene los seis dominios en el presente documento se representa como SEQ ID NO: 11; mientras que la secuencia de aminoácidos para toda la región de ACP que contiene los seis dominios en el presente documento se representa como SEQ ID NO: 12. Se encontró que el intervalo de repetición para los seis dominios de ACP dentro de SEQ ID NO: 11 se producía a aproximadamente cada 342 nucleótidos (el número real de aminoácidos

medido entre serinas de sitio activo adyacentes varía de 114 a 116 aminoácidos). Cada uno de los seis dominios de ACP contiene un motivo de unión a panteteína LGIDS* (SEQ ID NO: 47) en el que S* es la serina de sitio de unión a panteteína (S). La serina de sitio de unión a panteteína (S) está situada cerca del centro de cada secuencia del dominio de ACP. Las posiciones de los restos de serina de sitio activo (es decir, el sitio de unión a panteteína) para cada uno de los seis dominios de ACPD, con respecto a la secuencia de aminoácidos de SEQ ID NO: 2 son: ACP1 = S1152, ACP2 = S1266, ACP3 = S1382, ACP4 = S1498, ACP5 = S1614, y ACP6 = S1728.

El noveno dominio en Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, es un dominio KR. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio KR de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 25, que corresponde a las posiciones 5623-7800 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene el dominio KR de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 26, que corresponde a las posiciones 1875-2600 de SEQ ID NO: 2. Dentro del dominio KR se encuentra una región de núcleo (contenida dentro de la secuencia de nucleótidos de SEQ ID NO: 48, y la secuencia de aminoácidos de SEQ ID NO: 49) con homología con respecto a las aldehído-deshidrogenasas de cadena corta (KR es un miembro de esta familia). Esta región de núcleo se extiende desde aproximadamente la posición 5998 a aproximadamente 6900 de SEQ ID NO: 1, que corresponde las posiciones de los aminoácidos 2000-2300 de SEQ ID NO: 2.

El décimo dominio en Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, es un dominio de DH. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 27, que corresponde a las posiciones 7027-7065 de SEQ ID NO: 1. La secuencia de aminoácidos que contiene el dominio de DH de Pfa1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en el presente documento se representa como SEQ ID NO: 28, que corresponde a las posiciones 2343-2355 de SEQ ID NO: 2. El dominio de DH contiene un motivo de sitio activo conservado (Véase, Donadio, S. y Katz., L., *Gene* 111 (1): 51-60 (1992)): LxxHxxxGxxxP (SEQ ID NO: 50).

La Tabla 4 muestra los dominios y los sitios activos asociados con PFA2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Tabla 4: Análisis de Dominio de PFA2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Dominio	Posiciones de ADN	Posiciones de AA	Sitios	Posiciones de ADN	Posiciones de AA
KS	10-1350 de SEQ ID NO: 3 (SEQ ID NO: 29)	4-450 de SEQ ID NO: 4 (SEQ ID NO: 30)	DXAC* (SEQ ID NO: 43)	571-573 de SEQ ID NO: 3	C191 de SEQ ID NO: 4
			Terminal - GFGG (SEQ ID NO: 44)	1312-1323 de SEQ ID NO: 3 (SEQ ID NO: 51)	438-441 de SEQ ID NO: 4
CLF	1408-2700 de SEQ ID NO: 3 (SEQ ID NO: 31)	470-900 de SEQ ID NO: 4 (SEQ ID NO: 32)			
AT	2998-4200 de	1000-1400 de	GxS*xG	3421-3423 de	S1141 de SEQ ID
	SEQ ID NO: 3 (SEQ ID NO: 33)	SEQ ID NO: 4 (SEQ ID NO: 34)	(SEQ ID NO: 52)	SEQ ID NO: 3	NO: 4
ER	4498-5700 de SEQ ID NO: 3 (SEQ ID NO: 35)	1500-1900 de SEQ ID NO: 4 (SEQ ID NO: 36)			

El primer dominio en Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, es un dominio de KS. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de KS de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 29, que corresponde a

las posiciones 10-1350 de SEQ ID NO: 3. La secuencia de aminoácidos que contiene el dominio de KS de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 30, que corresponde a las posiciones 4-450 de SEQ ID NO: 4. El dominio de KS contiene un motivo de sitio activo: DXAC* (SEQ ID NO: 43), con un *sitio de unión a acilo que corresponde a C191 de SEQ ID NO: 4. Además, un motivo característico está presente en el extremo del dominio de KS: GFGG (SEQ ID NO: 44), que corresponde a las posiciones 438-441 de SEQ ID NO: 4 y a las posiciones 435-438 de SEQ ID NO: 30.

El tercer dominio en Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* es un dominio de CLF. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de CLF Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 31, que corresponde a las posiciones 1408-2700 de SEQ ID NO: 3. La secuencia de aminoácidos que contiene el dominio de CLF Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 32, que corresponde a las posiciones 470-900 de SEQ ID NO: 4.

El tercer dominio en Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* es un dominio de AT. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de AT de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 33, que corresponde a las posiciones 2998-4200 de SEQ ID NO: 3. La secuencia de aminoácidos que contiene el dominio de AT de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 34, que corresponde a las posiciones 1000-1400 de SEQ ID NO: 4. El dominio de AT contiene un motivo de sitio activo de GxS*xG (SEQ ID NO: 52) es característico de proteínas aciltransferasa (AT), con un resto de serina de sitio activo que corresponde a S 1141 de SEQ ID NO: 4.

El cuarto dominio de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* es un dominio de ER. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de ER de Pfa2 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 35, que corresponde a las posiciones 4498-5700 de SEQ ID NO: 3. La secuencia de aminoácidos que contiene el dominio de ER de Pfa2 en el presente documento se representa como SEQ ID NO: 36, que corresponde a las posiciones 1500-1900 de SEQ ID NO: 4.

La Tabla 5 muestra los dominios y los sitios activos asociados con PFA3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Tabla 5: Análisis de Dominio de PFA3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*

Dominio	Posiciones de ADN	Posiciones de AA	Sitios	Posiciones de ADN	Posiciones de AA
DH1	1-1350 de SEQ ID NO: 5 (SEQ ID NO: 37)	1-450 de SEQ ID NO: 6 (SEQ ID NO: 38)	FxxH*F (SEQ ID NO: 53)	931-933 de SEQ ID NO: 5	H310 de SEQ ID NO: 6
DH2	1501-2700 de SEQ ID NO: 5 (SEQ ID NO: 39)	501-900 de SEQ ID NO: 6 (SEQ ID NO: 40)	FxxH*F (SEQ ID NO: 53)	2401-2403 de SEQ ID NO: 5	H801 de SEQ ID NO: 6
ER	2848-4200 de SEQ ID NO: 5 (SEQ ID NO: 41)	950-1400 de SEQ ID NO: 6 (SEQ ID NO: 42)			

El primer y segundo dominios de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, son dominios de DH, denominados DH1 y DH2 en el presente documento, respectivamente. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH1 de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 37, que corresponde a las posiciones 1-1350 de SEQ ID NO: 5. La secuencia de aminoácidos que contiene el dominio de DH1 de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 38, que corresponde a las posiciones 1-450 de SEQ ID NO: 6. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH2 de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 39, que corresponde a las posiciones 1501-2700 de SEQ ID NO: 5. La secuencia de aminoácidos que contiene el dominio de DH2 de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 40, que corresponde a las posiciones 501-900 de SEQ ID NO: 6. Los dominios de DH contienen un motivo de sitio activo: FxxH*F (SEQ ID NO: 53). La secuencia de nucleótidos que contiene el motivo de sitio activo en DH1 corresponde a las posiciones 931-933 de SEQ ID NO: 5, mientras que la secuencia de nucleótidos que contiene el motivo de sitio activo en DH2 corresponde a las posiciones 2401-2403 de SEQ ID NO: 5. El sitio activo H* en el motivo FxxH*F se basa en datos de Leesong *et al.*, Structure 4: 253-64 (1996) y Kimber *et al.*, J Biol Chem. 279: 52593-602 (2004), con el sitio activo H* en DH1 que corresponde a H310 de SEQ ID NO: 6 y el sitio activo H* en DH2 que corresponde a H801 de SEQ ID NO: 6.

El tercer dominio de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* es un dominio de ER. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de ER de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 41, que corresponde a las posiciones 2848-4200 de SEQ ID NO: 5. La secuencia de aminoácidos que contiene el dominio de ER de Pfa3 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en el presente documento se representa como SEQ ID NO: 42, que corresponde a las posiciones 950-1400 de SEQ ID NO: 6.

EJEMPLO 4

Se diseñaron cebadores degenerados para los dominios de KS, ER, y DH de PUFA sintasa para aislar las secuencias correspondientes a partir del microorganismo aislado depositado con el N.º PTA-10212 de Registro en la ATCC, también conocido como cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Se diseñaron cebadores degenerados para la región de KS de PFA1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (es decir, la región que contiene el dominio de KS) basándose en las secuencias de PFA1 (denominado previamente *orfA* u ORF 1) publicadas para la cepa 20888 de la ATCC de *Schizochytrium sp.*, *Thraustochytrium aureum* (ATCC 34304), y cepa 23B ATCC 20892 de *Thraustochytrium sp.*:

prDS233 (directo): TGATATGGGAGGAATGAATTGTGTNGTNGAYGC (SEQ ID NO: 123)

prDS235 (inverso): TTCCATAACAAAATGATAATTAGCTCCNCCRAANCC (SEQ ID NO: 124).

Se diseñaron cebadores degenerados para la región de ER de PFA2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (es decir, la región que contiene el dominio de ER) las secuencias de PFA2 (denominado previamente *orfB* u ORF 2) publicadas para *Shewanella japonica*, la cepa 20888 de la ATCC de *Schizochytrium sp.*, *Thraustochytrium aureum* (ATCC 34304), y la cepa 23B ATCC 20892 de *Thraustochytrium sp.*:

prDS183 (directo): GGCGGCCACACCGAYAAVMGNCC (SEQ ID NO: 125)

prDS184 (inverso): CGGGGCCGCACCANAYYTGRTA (SEQ ID NO: 126).

Se diseñaron cebadores degenerados para la región de ER de PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (es decir, la región que contiene el dominio de ER) basándose en las secuencias de PFA3 (denominado previamente *orfC* u ORF 3) publicadas para *Shewanella japonica*, la cepa 20888 de la ATCC de *Schizochytrium sp.*, *Thraustochytrium aureum* (ATCC 34304), y la cepa 23B ATCC 20892 de *Thraustochytrium sp.*:

prDS181 (directo): TCCTTCGGNGCNGSNGG (SEQ ID NO: 127)

prDS184 (inverso): CGGGGCCGCACCANAYYTGRTA (SEQ ID NO: 126).

Los cebadores degenerados JGM190 (directo, SEQ ID NO: 64) y BLR242 (inverso, SEQ ID NO: 65), como se ha descrito anteriormente, se usaron para amplificar la región de DH de PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Las condiciones de PCR con molde de ADN cromosómico fueron las que siguen a continuación: dNTPs 0,2 µM, 0,1 uM cada cebador, DMSO al 6 %, 200 ng de ADN cromosómico, 2,5 U de polimerasa de fusión Herculase® II (Stratagene), y 1X de tampón Herculase® (Stratagene) en un volumen total de 50 µl. El Protocolo de PCR incluía las siguientes etapas: (1) 98 °C durante 3 minutos; (2) 98 °C durante 30 segundos; (3) 54 °C durante 45 segundos; (4) 72 °C durante 1 minuto; (5) etapas 2-4 repetidas durante 40 ciclos; (6) 72 °C durante 5 minutos; y (7) se mantiene a 6 °C.

Para todos los pares de cebadores, la PCR proporcionó distintos productos de ADN con los tamaños esperados usando moldes cromosómicos de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* Los respectivos productos de PCR se exploraron en el vector pJET1.2/blunt (Fermentas) de acuerdo con las instrucciones del fabricante, y la secuencia del inserto se determinó usando cebadores convencionales suministrados.

Las secuencias de ADN obtenidas a partir de los productos de PCR se compararon con secuencias conocidas disponibles a partir del NCBI GenBank como se describe en el Ejemplo 1.

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos obtenida a partir del ADN clonado que contenía el fragmento de KS de PFA1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* fueron: cepa 20888 de la ATCC de *Schizochytrium sp.*, "subunidad de ácido graso poliinsaturado sintasa A" (Identidad = 80 %; positivos = 90 %); KT99 de *Shewanella benthica* "PfaA de ácido graso poliinsaturado sintasa omega-3" (Identidad = 51 %; positivos = 67 %); PV-4 de *Shewanella loihica* "beta-cetoacil sintasa" (Identidad = 50 %; positivos = 67 %); ATCC 51908 *Shewanella woodyi* "PfaA de ácido graso poliinsaturado sintasa similar a poliéctido" (Identidad = 51 %; positivos = 66 %).

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos obtenida a partir del ADN clonado que contenía el fragmento de ER de PFA2 de la cepa PTA-10212 de

la ATCC de *Thraustochytrium* sp. fueron: cepa 20888 de la ATCC de *Schizochytrium* sp., "subunidad de ácido graso poliinsaturado sintasa B" (Identidad = 70 %; positivos = 85 %); la cepa 20888 de la ATCC de *Schizochytrium* sp., "subunidad de ácido graso poliinsaturado sintasa C" (Identidad = 66 %; positivos = 83 %); CCY9414 de *Nodularia spumigena* "2-nitropropano dioxigenasa" (Identidad = 57 %; positivos = 74 %); PE36 de *Moritella* sp. "PfaD de ácido graso poliinsaturado sintasa" (Identidad = 57 %; positivos = 71 %).

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos obtenida a partir del ADN clonado que contenía el fragmento de ER de PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. fueron: cepa 20888 de la ATCC de *Schizochytrium* sp., "subunidad de ácido graso poliinsaturado sintasa C" (Identidad = 80 %; positivos = 90 %); la cepa 20888 de la ATCC de *Schizochytrium* sp., "subunidad de ácido graso poliinsaturado sintasa B" (Identidad = 78 %; positivos = 89 %); PE36 de *Moritella* sp. "PfaD de ácido graso poliinsaturado sintasa" (Identidad = 56 %; positivos = 71 %); SB2B de *Shewanella amazonensis* "omega-3 PfaD de ácido graso poliinsaturado sintasa" (Identidad = 55 %; positivos = 73 %).

Al nivel del aminoácido, las secuencias con el nivel de homología más elevado con respecto a la secuencia de aminoácidos obtenida a partir del ADN clonado que contenía el fragmento de DH de PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. fueron: cepa 20888 de la ATCC de *Schizochytrium* sp., "subunidad de ácido graso poliinsaturado sintasa C" (Identidad = 63 %; positivos = 76 %); ATCC 700345 de *Shewanella pealeana* "Beta-hidroxilacil-(proteína transportadora de acilo) deshidratasa FabA/FabZ" (Identidad = 35 %; positivos = 53 %); WP3 de *Shewanella piezotolerans* "beta-cetoacil sintasa de múltiples dominios" (Identidad = 36 %; positivos = 52 %); KT99 de *Shewanella benthica* "PfaC de ácido graso poliinsaturado sintasa omega-3" (Identidad = 35 %; positivos = 51 %).

EJEMPLO 5

Los genes de PUFA sintasa se identificaron a partir de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp.

A partir de un criovial a -80 °C, se descongeló 1 ml de células a temperatura ambiente y se añadió a 50 ml de medio líquido de HSFM (a continuación) en un matraz sin tabique deflector de 250 ml. El matraz se incubó a 23 °C durante 3 días. Las células se recogieron y se utilizaron para construcción de biblioteca convencional de Cromosoma Artificial Bacteriano (BAC) (Lucigen Corporation, Middleton, WI USA).

Tabla 6: Medios de HSFM

Ingrediente	Concentración		Intervalos
Na ₂ SO ₄	g/l	31,0	0-50, 15-45, o 25-35
NaCl	g/l	0,625	0-25, 0,1-10, o 0,5-5
KCl	g/l	1,0	0-5, 0,25-3, o 0,5-2
MgSO ₄ ·7H ₂ O	g/l	5,0	0-10, 2-8, o 3-6
(NH ₄) ₂ SO ₄	g/l	0,44	0-10, 0,25-5, o 0,05-3
MSG·1H ₂ O	g/l	6,0	0-10, 4-8, o 5-7
CaCl ₂	g/l	0,29	0,1-5, 0,15-3, o 0,2-1
T 154 (extracto de levadura)	g/l	6,0	0-20, 0,1-10, o 1-7
KH ₂ PO ₄	g/l	0,8	0,1-10, 0,5-5, o 0,6-1,8
Después de autoclave (Metales)			
Ácido cítrico	mg/l	3,5	0,1-5000, 10-3000, o 3-2500
FeSO ₄ ·7H ₂ O	mg/l	10,30	0,1-100, 1-50, o 5-25
MnCl ₂ ·4H ₂ O	mg/l	3,10	0,1-100, 1-50, o 2-25
ZnSO ₄ ·7H ₂ O	mg/l	3,10	0,01-100, 1-50, o 2-25
CoCl ₂ ·6H ₂ O	mg/l	0,04	0-1, 0,001-0,1, o 0,01-0,1
Na ₂ MoO ₄ ·2H ₂ O	mg/l	0,04	0,001-1, 0,005-0,5, o 0,01-0,1
CuSO ₄ ·5H ₂ O	mg/l	2,07	0,1-100, 0,5-50, o 1-25
NiSO ₄ ·6H ₂ O	mg/l	2,07	0,1-100, 0,5-50, o 1-25

Después de autoclave (Vitaminas)

Tiamina	mg/l	9,75	0,1-100, 1-50, o 5-25
Vitamina B 12	mg/l	0,16	0,01-100, 0,05-5, o 0,1-1
Ca $\frac{1}{2}$ -pantotenato	mg/l	2,06	0,1-100, 0,1-50, o 1-10
Biotina	mg/l	3,21	0,1-100, 0,1-50, o 1-10

Después de autoclave (Carbono)

Glicerol	g/l	30,0	5-150, 10-100, o 20-50

Alimentación de nitrógeno:

Ingrediente	Concentración		
MSG·1H ₂ O	g/l	17	0-150, 10-100, o 15-50

Las condiciones habituales de cultivo podrían incluir las siguientes:

5	pH	aproximadamente 6,5 - aproximadamente 9,5, aproximadamente 6,5 - aproximadamente 8,0, o aproximadamente 6,8 - aproximadamente 7,8;		
	temperatura:	aproximadamente 15 - aproximadamente 30 grados Celsius, aproximadamente 18 - aproximadamente 28 grados Celsius, o aproximadamente 21 a aproximadamente 23 grados Celsius;		
10	oxígeno disuelto:	aproximadamente 0,1 - aproximadamente 100 % de saturación, aproximadamente 5 - aproximadamente 50 % de saturación, o aproximadamente 10 - aproximadamente 30 % de saturación; y/o		
15	glicerol controlado a:	aproximadamente 5 - aproximadamente 50 g/l, aproximadamente 10 - aproximadamente 40 g/l, o aproximadamente 15 - aproximadamente 35 g/l.		

Las bibliotecas de BAC recombinante, que consisten en grandes fragmentos (promedio de aproximadamente 120 kB) se manipularon de acuerdo con las instrucciones del fabricante en el vector pSMART® de BAC (Lucigen Corporation). Las bibliotecas de BAC se identificaron sistemáticamente mediante procedimientos convencionales de hibridación de colonias usando sondas etiquetadas radiactivamente con ³²P (Sambrook J. y Russell D. 2001. Molecular cloning: A laboratory manual, 3ª edición. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). Las sondas contenían ADN homólogo a las secuencias de PUFA sintasa publicadas de otros organismos como se describe en el Ejemplo 4. Estas sondas se generaron mediante una digestión de restricción de ADN de los fragmentos clonados a partir de los respectivos clones de pJET1.2/blunt descritos anteriormente y se etiquetaron con métodos convencionales. En todos los casos, la fuerte hibridación de las sondas individuales para ciertos BAC indicaban clones que contenían ADN homólogo a los genes de PUFA sintasa.

El clon pLR130 de BAC (también conocido como LuMaBAC 2M23) demostraba una fuerte hibridación de la sonda tanto en la región de KS como en la región de ER, lo que indica que contenía los genes de *PFA1* y *PFA2*, y se seleccionó para secuenciación de ADN de los genes *PFA1* y *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. El BAC se secuenció mediante procedimientos convencionales (Eurofins MWG Operon, Huntsville, AL). El clon pLR130 de BAC, que contenía los genes de *PFA1* y *PFA2*, se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 1 de diciembre de 2009, y se le dio el N.º de Registro en la ATCC PTA-10511.

En los sistemas de PUFA sintasa de traustochytrido publicados previamente, los genes de *PFA1* y *PFA2* de PUFA sintasa se agruparon en conjunto y se colocaron para que se transcribieran de manera divergente. Este también es el caso para *PFA1* y *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. Se encontró que el inicio conceptual de *PFA2* era de 693 nucleótidos desde el inicio de *PFA1* y se transcribió de manera divergente.

El clon pDS127 de BAC (también conocido como LuMaBAC 9K17) demostraba una fuerte hibridación de la sonda tanto en la región de DH como en la región de ER de *PFA3* y se seleccionó para secuenciación de ADN del gen de *PFA3*. El clon pDS127 de BAC, que contenía el gen de *PFA3*, se depositó bajo el Tratado de Budapest en la Colección Americana de Cultivos Tipo, Depósito de Patentes, 10801 University Boulevard, Manassas, VA 20110-2209, el 1 de diciembre de 2009, y se le dio el N.º de Registro en la ATCC PTA-10510. Se diseñaron cebadores de secuenciación usando métodos convencionales para la región de DH y la región de ER y la secuencia de ADN determinada en el Ejemplo 4. Para determinar la secuencia de ADN del gen de *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp., se realizaron rondas sucesivas de secuenciación de ADN, que implicaban el posterior diseño de cebador de secuenciación con métodos convencionales, para "acompañar" al clon de BAC. Cada par de bases de nucleótidos del gen *PFA3* se cubrió con al menos dos reacciones de secuenciación de ADN

separadas de alta calidad con al menos una puntuación Phred mínima agregada de 40 (nivel de confianza de un 99,99 %).

La Tabla 7 muestra identidades para las secuencias de polinucleótidos de *PFA1* (SEQ ID NO: 68), *PFA2* (SEQ ID NO: 70), y *PFA3* (SEQ ID NO: 72) de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en comparación con secuencias publicadas previamente y las secuencias de la cepa PTA-9695 de *Schizochytrium sp.* Las identidades se determinaron con la matriz de puntuación "swgapdnamt" dentro del programa "AlignX" del programa VectorNTI, un patrón para alineamiento de ADN.

Tabla 7: Porcentaje de Identidad para Secuencias de Polinucleótidos de *PFA1*, *PFA2*, y *PFA3*

Fuente de Comparación de Secuencias de <i>PFA1</i> , <i>PFA2</i> , y <i>PFA3</i>	% de Identidad de Comparación de <i>PFA1</i> (<i>orfA</i>) para <i>PFA1</i>	% de Identidad de Comparación de <i>PFA2</i> (<i>orfB</i>) para <i>PFA2</i>	% de Identidad de Comparación de <i>PFA3</i> (<i>orfC</i>) para <i>PFA3</i>
20888 de la ATCC de <i>Schizochytrium sp.</i>	55	54	59
34304 de la ATCC de <i>Thraustochytrium aureum</i>	55	53	no publicado
23B ATCC 20892 de <i>Thraustochytrium sp.</i>	55	57	62
PTA-9695 de <i>Schizochytrium sp.</i>	55	52	59

La Tabla 8 muestra identidades para las secuencias de aminoácidos de *Pfa1p* (SEQ ID NO: 69), *Pfa2p* (SEQ ID NO: 71), y *Pfa3p* (SEQ ID NO: 73) de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en comparación con secuencias de aminoácidos de PUFA sintasa publicadas previamente y las secuencias de la cepa PTA-9695 de *Schizochytrium sp.* Las identidades se determinaron a través del uso de la matriz de puntuación "blosum62mt2" dentro del programa "AlignX" del programa VectorNTI, un patrón para alineamiento de proteínas.

Tabla 8: Porcentaje de Identidad para Secuencias de Aminoácidos de *Pfa1p*, *Pfa2p*, y *Pfa3p*

Fuente de Comparación de Secuencias de <i>Pfa1p</i> , <i>Pfa2p</i> , y <i>Pfa3p</i>	% de Identidad de Comparación de <i>Pfa1p</i> (<i>OrfA</i>) para <i>Pfa1p</i>	% de Identidad de Comparación de <i>Pfa2p</i> (<i>OrfB</i>) para <i>Pfa2p</i>	% de Identidad de Comparación de <i>Pfa3p</i> (<i>OrfC</i>) para <i>Pfa3p</i>
20888 de la ATCC de <i>Schizochytrium sp.</i>	62	57	69
ATCC 34304 de <i>Thraustochytrium aureum</i>	58	54	no publicado
23B 20892 de la ATCC de <i>Thraustochytrium sp.</i>	54	54	71
PTA-9695 de <i>Schizochytrium sp.</i>	59	53	73

EJEMPLO 6

El análisis de dominios se realizó para anotar las coordenadas de la secuencia para los dominios de PUFA sintasa y sitios activos de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, respectivamente. Los dominios identificaron basándose en homología con respecto a dominios conocidos de PUFA sintasa, ácido graso sintasa, y policéptido sintasa

La Tabla 9 muestra los dominios y los sitios activos asociados con el *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Tabla 9: Análisis de Dominio de *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
KS	13-1362 de SEQ ID NO: 68 (SEQ ID NO: 74)	5-545 de SEQ ID NO: 69 (SEQ ID NO: 75)	Activo - DXAC* (SEQ ID NO: 43)	601-612 de SEQ ID NO: 68	C204 de SEQ ID NO: 69

ES 2 623 610 T3

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
			Terminal - GFGG (SEQ ID NO: 44)	1351-1362 de SEQ ID NO: 68 (SEQ ID NO: 45)	451-454 de SEQ ID NO: 69
MAT	1783-2703 de SEQ ID NO: 68 (SEQ ID NO: 76)	595-901 de SEQ ID NO: 69 (SEQ ID NO: 77)	GHS*LG Activo (SEQ ID NO: 46)	2083-2085 de SEQ ID NO: 68 (SEQ ID NO: 116)	S695 de SEQ ID NO: 69
ACP	3208-6510 de SEQ ID NO: 68 (SEQ ID NO: 78)	1070-2170 de SEQ ID NO: 69 (SEQ ID NO: 77)	Dominio de ACP1	3280-3534 de SEQ ID NO: 68 (SEQ ID NO: 80)	1094-1178 de SEQ ID NO: 69 (SEQ ID NO: 81)
			ACP1 Activo LGIDS* (SEQ ID NO: 47)	3403-3405 de SEQ ID NO: 68	S1135 de SEQ ID NO: 69
			Dominio de ACP2	3607-3861 de SEQ ID NO: 68 (SEQ ID NO: 82)	1203-1287 de SEQ ID NO: 69 (SEQ ID NO: 83)
			ACP2 Activo LGIDS* (SEQ ID NO: 47)	3730-3732 de SEQ ID NO: 68	S1244 de SEQ ID NO: 69
			Dominio de ACP3	3934-4185 de SEQ ID NO: 68 (SEQ ID NO: 84)	1312-1396 de SEQ ID NO: 69 (SEQ ID NO: 85)
			ACP3 Activo LGIDS* (SEQ ID NO: 47)	4057-4059 de SEQ ID NO: 68	S1353 de SEQ ID NO: 69
			Dominio de ACP4	4261-4515 de SEQ ID NO: 68 (SEQ ID NO: 86)	1421-1505 de SEQ ID NO: 69 (SEQ ID NO: 87)
			ACP4 Activo LGIDS* (SEQ ID NO: 47)	4384-4386 de SEQ ID NO: 68	S1462 de SEQ ID NO: 69
			Dominio de ACP5	4589-4842 de SEQ ID NO: 68 (SEQ ID NO: 88)	1530-1614 de SEQ ID NO: 69 (SEQ ID NO: 89)

ES 2 623 610 T3

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
			ACP5 Activo LGIDS* (SEQ ID NO: 47)	4711-4713 de SEQ ID NO: 68	S1571 de SEQ ID NO: 69
			Dominio de ACP6	4915-5169 de SEQ ID NO: 68 (SEQ ID NO: 90)	1639-1723 de SEQ ID NO: 69 (SEQ ID NO: 91)
			ACP6 Activo LGIDS* (SEQ ID NO: 47)	5038-5040 de SEQ ID NO: 68	S1680 de SEQ ID NO: 69
			Dominio de ACP7	5242-5496 de SEQ ID NO: 68 (SEQ ID NO: 92)	1748-1832 de SEQ ID NO: 69 (SEQ ID NO: 93)
			ACP7 Activo LGIDS* (SEQ ID NO: 47)	5365-5367 de SEQ ID NO: 68	S1789 de SEQ ID NO: 69
			Dominio de ACP8	5569-5823 de SEQ ID NO: 68 (SEQ ID NO: 94)	1857-1941 de SEQ ID NO: 69 (SEQ ID NO: 95)
			ACP8 Activo LGIDS* (SEQ ID NO: 47)	5692-5694 de SEQ ID NO: 68	S1898 de SEQ ID NO: 69
			Dominio de ACP9	5896-6150 de SEQ ID NO: 68 (SEQ ID NO: 96)	1966-2050 de SEQ ID NO: 69 (SEQ ID NO: 97)
			ACP9 Activo LGIDS* (SEQ ID NO: 47)	6019-6021 de SEQ ID NO: 68	S2007 de SEQ ID NO: 69
			Dominio de ACP10	6199-6453 de SEQ ID NO: 68 (SEQ ID NO: 98)	2067-2151 de SEQ ID NO: 69 (SEQ ID NO: 99)
			ACP10 Activo LGIDS* (SEQ ID NO: 47)	6322-6324 de SEQ ID NO: 68	S2108 de SEQ ID NO: 69

Dominio	Posición de ADN	Posición de AA	Sitios	Posición de ADN	Posición de AA
KR	6808-8958 de SEQ ID NO: 68 (SEQ ID NO: 100)	2270-2986 de SEQ ID NO: 69 (SEQ ID NO: 101)	"región núcleo"	7198-8100 de SEQ ID NO: 68 (SEQ ID NO: 116)	2400-2600 de SEQ ID NO: 69 (SEQ ID NO: 117)
Motivo de DH	8203-8241 de SEQ ID NO: 68 (SEQ ID NO: 118)	2735-2747 de SEQ ID NO: 69 (SEQ ID NO: 119)	LxxHxxxGxxxxP (SEQ ID NO: 50)	8203-8241 de SEQ ID NO: 68 (SEQ ID NO: 118)	2735-2747 de SEQ ID NO: 69 (SEQ ID NO: 119)

El primer dominio en Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de KS. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de KS de Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en el presente documento se representa como SEQ ID NO: 74, que corresponde a las posiciones 13-1362 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene el dominio de KS de Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en el presente documento se representa como SEQ ID NO: 75, que corresponde a las posiciones 5-454 de SEQ ID NO: 69. El dominio de KS contiene un motivo de sitio activo: DXAC* (SEQ ID NO: 43), con un *sitio de unión a acilo que corresponde a C204 de SEQ ID NO: 69. Además, un motivo característico está presente en el extremo del dominio de KS: GFGG (SEQ ID NO: 44), que corresponde a las posiciones 451-454 de SEQ ID NO: 69 y a las posiciones 447-450 de SEQ ID NO: 75.

El segundo dominio en Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio MAT. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio MAT de Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en el presente documento se representa como SEQ ID NO: 76, que corresponde a las posiciones 1783-2703 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene el dominio MAT de Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en el presente documento se representa como SEQ ID NO: 77, que corresponde a las posiciones 595-901 de SEQ ID NO: 69. El dominio MAT contiene un motivo de sitio activo: GHS**XG* (SEQ ID NO: 46), con un *sitio de unión a acilo que corresponde a S695 de SEQ ID NO: 69.

Los dominios tercero a decimosegundo de Pfa1p de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* son diez dominios de ACP en tándem, también denominados ACP1, ACP2, ACP3, ACP4, ACP5, ACP6, ACP7, ACP8, ACP9, y ACP10 en el presente documento. La secuencia de nucleótidos que contiene el primer dominio de ACP, ACP1, en el presente documento se representa como SEQ ID NO: 80 y está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 3280 a aproximadamente la posición 3534 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP1, representada en el presente documento como SEQ ID NO: 81, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1094 a aproximadamente la posición 1178 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP2, representada en el presente documento como SEQ ID NO: 82, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 3607 a aproximadamente la posición 3861 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP2, representada en el presente documento como SEQ ID NO: 83, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1203 a aproximadamente la posición 1287 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP3, representada en el presente documento como SEQ ID NO: 84, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 3934 a aproximadamente la posición 4185 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP3, representada en el presente documento como SEQ ID NO: 85, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1312 a aproximadamente la posición 1396 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP4, representada en el presente documento como SEQ ID NO: 86, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4261 a aproximadamente la posición 4515 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP4, representada en el presente documento como SEQ ID NO: 87, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1421 a aproximadamente la posición 1505 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP5, representada en el presente documento como SEQ ID NO: 88, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4589 a aproximadamente la posición 4842 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP5, representada en el presente documento como SEQ ID NO: 89, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1530 a aproximadamente la posición 1614 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP6, representada en el presente documento como SEQ ID NO: 90, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 4915 a aproximadamente la posición 5169 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP6, representada en el presente documento como SEQ ID NO: 91, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1639 a

aproximadamente la posición 1723 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP7, representada en el presente documento como SEQ ID NO: 92, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 5242 a aproximadamente la posición 5496 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP7, representada en el presente documento como SEQ ID NO: 93, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1748 a aproximadamente la posición 1832 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP8, representada en el presente documento como SEQ ID NO: 94, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 5569 a aproximadamente la posición 5832 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP8, representada en el presente documento como SEQ ID NO: 95, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1857 a aproximadamente la posición 1941 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP9, representada en el presente documento como SEQ ID NO: 96, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 5896 a aproximadamente la posición 6150 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP9, representada en el presente documento como SEQ ID NO: 97, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 1966 a aproximadamente la posición 2050 de SEQ ID NO: 69. La secuencia de nucleótidos que contiene ACP10, representada en el presente documento como SEQ ID NO: 98, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 6199 a aproximadamente la posición 6453 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene ACP10, representada en el presente documento como SEQ ID NO: 99, está contenida dentro de la secuencia de nucleótidos que se extiende desde aproximadamente la posición 2067 a aproximadamente la posición 2151 de SEQ ID NO: 69. Los diez dominios de ACP en conjunto se extienden a una región de Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* desde aproximadamente la posición 3208 a aproximadamente la posición 6510 de SEQ ID NO: 68, que corresponde a posiciones del aminoácido de aproximadamente 1070 a aproximadamente 2170 de SEQ ID NO: 69. La secuencia de nucleótidos para toda la región de ACP que contiene los 10 dominios en el presente documento se representa como SEQ ID NO: 78; mientras que la secuencia de aminoácidos para toda la región de ACP que contiene los seis dominios en el presente documento se representa como SEQ ID NO: 79. Se encontró que el intervalo de repetición para los 10 dominios de ACP dentro de SEQ ID NO: 78 se producía aproximadamente cada 327 nucleótidos (el número real de aminoácidos medido entre serinas de sitio activo adyacentes varía de 101 a 109 aminoácidos). Cada uno de los diez dominios de ACP contiene un motivo de unión a panteteína LGIDS* (SEQ ID NO: 47) en el que S* es la serina de sitio de unión a panteteína (S). La serina de sitio de unión a panteteína (S) está situada cerca del centro de cada secuencia de dominio de ACP. Las posiciones de los restos de serina de sitio activo (es decir, el sitio de unión a panteteína) para cada uno de los seis dominios de ACPD, con respecto a la secuencia de aminoácidos de SEQ ID NO: 69 son: ACP1 = S1135, ACP2 = S1244, ACP3 = S1353, ACP4 = S1462, ACP5 = S1571, ACP6 = S1680, ACP7 = S1789, ACP7 = S1789, ACP8 = S1898, ACP9 = S2007, y ACP10 = S2108.

El decimotercer dominio en Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de KR. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de KR de Pfa1 en el presente documento se representa como SEQ ID NO: 100, que corresponde a las posiciones 6808-8958 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene el dominio de KR de Pfa1 en el presente documento se representa como SEQ ID NO: 101, que corresponde a las posiciones 2270-2986 de SEQ ID NO: 69. Dentro del dominio de KR se encuentra una región núcleo (contenida dentro de la secuencia de nucleótidos de SEQ ID NO: 116, y la secuencia de aminoácidos de SEQ ID NO: 117) con homología para aldehído-deshidrogenasas de cadena corta (KR es un miembro de esta familia). Esta región núcleo se extiende desde aproximadamente la posición 5998 a aproximadamente 6900 de SEQ ID NO: 68, que corresponde a las posiciones 2000-2300 de los aminoácidos de SEQ ID NO: 69.

El decimocuarto dominio en Pfa1 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de DH. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH de Pfa1 en el presente documento se representa como SEQ ID NO: 118, que corresponde a las posiciones 7027-7065 de SEQ ID NO: 68. La secuencia de aminoácidos que contiene el dominio de DH de Pfa1 en el presente documento se representa como SEQ ID NO: 119, que corresponde a las posiciones 2343-2355 de SEQ ID NO: 69. El dominio de DH contiene un motivo de sitio activo conservado (véase, Donadio, S. y Katz., L., Gene 111 (1): 51-60 (1992)): LxxHxxxGxxxxP (SEQ ID NO: 50).

La Tabla 10 muestra los dominios y los sitios activos asociados con PFA2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Tabla 10: Análisis de Dominio de PFA2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Dominio	Posiciones de ADN	Posiciones de AA	Sitios	Posiciones de ADN	Posiciones de AA
KS	10-1320 de SEQ ID NO: 70	4-440 de SEQ ID NO: 71	DXAC*	571-573 de SEQ ID NO: 70	C191 de SEQ ID NO: 71
	(SEQ ID NO: 102)	(SEQ ID NO: 103)	(SEQ ID NO: 43)		

Dominio	Posiciones de ADN	Posiciones de AA	Sitios	Posiciones de ADN	Posiciones de AA
			Terminal - GFGG (SEQ ID NO: 44)	1312-1323 de SEQ ID NO: 70	423-426 de SEQ ID NO: 71
CLF	1378-2700 de SEQ ID NO: 70 (SEQ ID NO: 104)	460-900 de SEQ ID NO: 71 (SEQ ID NO: 105)			
AT	2848-4200 de SEQ ID NO: 70 (SEQ ID NO: 106)	950-1400 de SEQ ID NO: 71 (SEQ ID NO: 107)	GxS*xG (SEQ ID NO: 52)	3361-3363 de SEQ ID NO: 70	S1121 de SEQ ID NO: 71
ER	4498-5700 de SEQ ID NO: 70 (SEQ ID NO: 108)	1500-1900 de SEQ ID NO: 71 (SEQ ID NO: 109)			

El primer dominio en Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de KS. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de KS de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 102, que corresponde a las posiciones 10-1320 de SEQ ID NO: 70. La secuencia de aminoácidos que contiene el dominio de KS de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 103, que corresponde a las posiciones 4-440 de SEQ ID NO: 71. El dominio de KS contiene un motivo de sitio activo: DXAC* (SEQ ID NO: 43), con un *sitio de unión a acilo que corresponde a C191 de SEQ ID NO: 71. Además, un motivo característico está presente en el extremo del dominio de KS: GFGG (SEQ ID NO: 44), que corresponde a las posiciones 423-426 de SEQ ID NO: 71 y a las posiciones 1267-1278 de SEQ ID NO: 70.

El segundo dominio en Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de CLF. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de CLF de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 104, que corresponde a las posiciones 1378-2700 de SEQ ID NO: 70. La secuencia de aminoácidos que contiene el dominio de CLF de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 105, que corresponde a las posiciones 460-900 de SEQ ID NO: 71.

El tercer dominio en Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de AT. La secuencia de nucleótidos que contiene la secuencia que codifica the el dominio de AT de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 106, que corresponde a las posiciones 2848-4200 de SEQ ID NO: 70. La secuencia de aminoácidos que contiene el dominio de AT de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 107, que corresponde a las posiciones 950-1400 de SEQ ID NO: 71. El dominio de AT contiene un motivo de sitio activo de GxS*xG (SEQ ID NO: 50) que es característico de proteínas aciltransferasa (AT), con un resto de serina de sitio activo que corresponde a S1121 de SEQ ID NO: 71.

El cuarto dominio de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de ER. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de ER de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 108, que corresponde a las posiciones 4498-5700 de SEQ ID NO: 70. La secuencia de aminoácidos que contiene el dominio de ER de Pfa2 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 109, que corresponde a las posiciones 1500-1900 de SEQ ID NO: 71.

La Tabla 11 muestra los dominios y los sitios activos asociados con PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Tabla 11: Análisis de Dominio de PFA3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

Dominio	Posiciones de ADN	Posiciones de AA	Sitios	Posiciones de ADN	Posiciones de AA
DH1	1-1350 de SEQ ID NO: 72 (SEQ ID NO: 110)	1-450 de SEQ ID NO: 73 (SEQ ID NO: 111)	FxxH*F (SEQ ID NO: 53)	934-936 de SEQ ID NO: 72	H312 de SEQ ID NO: 73
DH2	1501-2700 de SEQ ID NO: 72 (SEQ ID NO: 112)	501-900 de SEQ ID NO: 73 (SEQ ID NO: 113)	FxxH*F (SEQ ID NO: 53)	2401-2403 de SEQ ID NO: 72	H801 de SEQ ID NO: 73
ER	2848-4212 de SEQ ID NO: 72 (SEQ ID NO: 114)	950-1404 de SEQ ID NO: 73 (SEQ ID NO: 115)			

El primer y segundo dominios de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* son dominios de DH, denominados DH1 y DH2 en el presente documento, respectivamente. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH1 de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 110, que corresponde a las posiciones 1-1350 de SEQ ID NO: 72. La secuencia de aminoácidos que contiene el dominio de DH1 de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 111, que corresponde a las posiciones 1-450 de SEQ ID NO: 73. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de DH2 de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 112, que corresponde a las posiciones 1501-2700 de SEQ ID NO: 72. La secuencia de aminoácidos que contiene el dominio de DH2 de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 113, que corresponde a las posiciones 501-900 de SEQ ID NO: 73. Los dominios de DH contienen un motivo de sitio activo: FxxH*F (SEQ ID NO: 53). La secuencia de nucleótidos que contiene el motivo de sitio activo en DH1 corresponde a las posiciones 934-936 de SEQ ID NO: 72, mientras que la secuencia de nucleótidos que contiene el motivo de sitio activo en DH2 corresponde a las posiciones 2401-2403 de SEQ ID NO: 72. El sitio activo H* en el motivo FxxH*F se basa en datos de Leesong *et al.*, Structure 4: 253-64 (1996) y Kimber *et al.* J Biol Chem. 279: 52593-602 (2004), con el sitio activo H* en DH1 que corresponde a H312 de SEQ ID NO: 73 y el sitio activo H* en DH2 que corresponde a H801 de SEQ ID NO: 73.

El tercer dominio de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* es un dominio de ER. La secuencia de nucleótidos que contiene la secuencia que codifica el dominio de ER de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 114, que corresponde a las posiciones 2848-4200 de SEQ ID NO: 72. La secuencia de aminoácidos que contiene el dominio de ER de Pfa3 de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* en el presente documento se representa como SEQ ID NO: 115, que corresponde a las posiciones 950-1400 de SEQ ID NO: 73.

EJEMPLO 7

La inactivación de los genes de PUFA sintasa nativa en la cepa 20888 de la ATCC de *Schizochytrium sp.*, para generar auxótrofos de PUFA, y la sustitución de los genes inactivados de este tipo con genes homólogos introducidos por vía exógena para restablecer la síntesis de PUFA se ha demostrado y descrito previamente. Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.217.856, incorporado por referencia en el presente documento en su totalidad. Los tres genes de PUFA sintasa de la cepa 20888 de la ATCC de *Schizochytrium sp.*, se han denominado previamente *orfA*, *orfB*, y *orfC*, que corresponden a la nomenclatura PFA1, PFA2, y PFA3 usada en el presente documento, respectivamente. *Del mismo modo.*

El gen *orfA* nativo en la cepa 20888 de la ATCC de *Schizochytrium sp.*, se sustituyó mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a ZeocinTM rodeado por secuencias de la región de flaqueo de *orfA*. Se generó una cepa mutante que carecía de un gen *orfA* funcional. La cepa mutante era auxotrófica y para su crecimiento requería suplemento de PUFA.

El PFA1 de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 1) se clonó en el vector de expresión pREZ37 para generar pREZ345. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flaqueo del locus del gen *orfA* nativo de la cepa 20888 de la ATCC de *Schizochytrium sp.* El *orfA* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, se transformó a través de electroporación con tratamiento enzimático previo (véase a continuación) con pREZ345 que contenía PFA1. Basándose en regiones homólogas que flanquean el marcador de resistencia a ZeocinTM en el mutante y que flanquea el gen PFA1 en

pREZ345, se produjo una recombinación de doble cruce de modo que *PFA1* se insertó en el locus de *orfA* nativo. La recombinación con *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 1) restableció la producción de PUFA en el *orfA* que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.* En resumen, las células se cultivaron en medio líquido M2B (véase el párrafo que sigue a continuación) a 30 °C con agitación a 200 rpm durante 3 días. Las células se cosecharon y los ácidos grasos se convirtieron en ésteres de metilo usando técnicas convencionales. Los perfiles de ácido graso se determinaron usando detección mediante cromatografía de gases con ionización a la llama (GC-FID) como ésteres de metilo de ácido graso (FAME). La cepa nativa 20888 de la ATCC de *Schizochytrium sp.*, que contiene un gen *orfA* funcional produjo DHA y DPA n-6 en una proporción de 2,3:1. La cepa recombinante que contenía *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 1) en lugar del gen *orfA* inactivado también produjo DHA y DPA n-6 en una proporción de 2,4:1. El contenido de EPA de la cepa recombinante era de un 2,7 % de ésteres de metilo de ácido graso (FAME), el contenido de DPA n-3 era de un 0,7 %, el contenido de DPA n-6 era de un 8,8 %, y el contenido de DHA era de un 21,2 %.

Medio de M2B - 10 g/l de glucosa, 0,8 g/l de $(\text{NH}_4)_2\text{SO}_4$, 5 g/l de Na_2SO_4 , 2 g/l de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,5 g/l de KH_2PO_4 , 0,5 g/l de KCl, 0,1 g/l de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, MES 0,1 M (pH 6,0) metales de PB26 al 0,1 %, y Vitaminas de PB26 al 0,1 % (v/v). Las vitaminas de PB26 consistían en 50 mg/ml de vitamina B12, 100 µg/ml de tiamina, y 100 µg/ml de Ca-pantotenato. Los metales de PB26 se ajustaron a pH 4,5 y consistían en 3 g/l de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 1 g/l de $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 800 mg/ml de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 20 mg/ml de $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 10 mg/ml de $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 600 mg/ml de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, y 800 mg/ml de $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$. Las soluciones de reserva de PB26 se esterilizaron mediante filtración por separado y se añadieron al caldo de cultivo después de tratamiento en autoclave. Cada uno de glucosa, KH_2PO_4 , y $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ se trataron en autoclave por separado a partir del resto de los ingredientes del caldo de cultivo antes de la mezcla para evitar la precipitación de sal y la caramelización del carbohidrato. Todos los ingredientes del medio se adquirieron en Sigma Chemical (St. Louis, MO).

Electroporación con tratamiento enzimático previo - Las células se cultivaron en 50 ml de medio M50-20 (véase la Publicación de Estados Unidos N.º 2008/0022422) en un agitador a 200 rpm durante 2 días a 30 °C. Las células se diluyeron a 1: 100 en medio M2B y se cultivaron durante una noche (16-24 h), intentando alcanzar el crecimiento de fase logarítmica medio (DO600 de 1,5-2,5). Las células se centrifugaron en un tubo cónico de 50 ml durante 5 min a aproximadamente 3000 x g. El sobrenadante se retiró y las células se volvieron a suspender en manitol 1 M, pH 5,5, en un volumen adecuado para alcanzar una concentración final de 2 unidades de DO₆₀₀. Se tomaron alícuotas de 5 ml en el matraz agitador de 25 ml y se modificó con CaCl_2 10 mM (solución de reserva 1,0 M, esterilizado mediante filtración) y 0,25 mg/ml de Protease XIV (10 mg/ml de solución de reserva, esterilizado mediante filtración; Sigma-Aldrich, St. Louis, MO). Los matraces se incubaron en un agitador a 30 °C y aproximadamente 100 rpm durante 4 h. Las células se controlaron bajo el microscopio para determinar el grado de formación de protoplastos, con las células individuales deseadas. Las células centrifugaron durante 5 min a aproximadamente 2500 x g en tubos de fondo redondo (es decir, tubos Falcon™ de 14 ml, BD Biosciences, San Jose, CA). El sobrenadante se retiró y las células se volvieron a suspender con cuidado con 5 ml de glicerol al 10 % enfriado en hielo. Las células se volverán a centrifugar durante 5 min a aproximadamente 2500 x g en tubos de fondo redondo. El sobrenadante se retiró y las células se volverán a suspender con cuidado con 500 µl de glicerol al 10 % enfriado en hielo, usando puntas de pipeta de calibre ancho. Se tomaron alícuotas de 90 µl de células en una electrocubeta enfriada previamente (cubeta Gene Pulser® - 0,1 cm de hueco, o 0,2 cm de hueco, Bio-Rad, Hercules, CA). De un µg a 5 µg de ADN (en un volumen inferior o igual a 10 µl de volumen) se añadieron a la cubeta, se mezcló con cuidado con una punta de pipeta, y se colocó en hielo durante 5 min. Las células se electroporaron a 200 ohmios (resistencia), 25 µF (capacitancia), y cualquiera de 250 V (para hueco de 0,1 cm) o 500 V (hueco de 0,2 cm). Se añadieron 0,5 ml de medio M50-20 inmediatamente a la cubeta. A continuación, las células se transfirieron a 4,5 ml de medio M50-20 en un matraz agitador de 25 ml y se incubaron durante 2-3 h a 30 °C y aproximadamente 100 rpm en un agitador. Las células se centrifugaron durante 5 min a aproximadamente 2500 x g en tubos de fondo redondo. El sobrenadante se retiró y el sedimento celular se volvió a suspender en 0,5 ml de medio M50-20. Las células se sembraron en un número apropiado (de 2 a 5) de placas M2B con la selección apropiada y se incubaron a 30 °C.

El *orfA* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, también se transforma con pREZ345 que contiene *PFA1*, de modo que *PFA1* se integra de manera aleatoria en el mutante y restablece la producción de PUFA.

EJEMPLO 8

El *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 68) se volvió a sintetizar (DNA2.0) y se optimizó con codón para su expresión en *Schizochytrium* (SEQ ID NO: 120) y se clonó en un vector de expresión para generar pLR95. La optimización con codón se produjo usando la tabla de uso de codón de *Schizochytrium* en la FIG. 22. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flanqueo del locus del gen *orfA* nativo de la cepa 20888 de la ATCC de *Schizochytrium sp.*

El *orfA* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, del Ejemplo 7 se transformó mediante electroporación con tratamiento enzimático previo (Véase el Ejemplo 7) con pLR95 que contiene *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 120).

Basándose en regiones homólogas que flanquean el marcador de resistencia a Zeocin™ en el mutante y que flanquean el gen *PFA1* en pLR95, se produjo recombinación de doble cruce de modo que el *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón se insertó en el locus de *orfA* nativo. La recombinación con el *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón (SEQ ID NO: 120) restableció la producción de PUFA en el *orfA* que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium* sp. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. La cepa 20888 nativa de la ATCC de *Schizochytrium* sp., que contiene un gen *orfA* funcional produjo DHA y EPA en una proporción de 25:1. La cepa recombinante que contiene *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón (SEQ ID NO: 120) en lugar del gen *orfA* inactivado produjo DHA y EPA en una proporción de 5,4:1, lo que demuestra adicionalmente que el perfil de PUFA de *Schizochytrium* se puede alterar con las moléculas de ácido nucleico que se describen en el presente documento. El contenido de EPA de la cepa recombinante era de un 4,4 % de FAME, el contenido de DPA n-3 era de un 2,3 %, el contenido de DPA n-6 era de un 4,9 %, y el contenido de DHA era de un 24,0 %.

El *orfA* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium* sp., también se transforma con pLR95 que contiene *PFA1*, de modo que *PFA1* se integra de manera aleatoria en el mutante y restablece la producción de PUFA.

EJEMPLO 9

El gene *orfB* nativo en la cepa 20888 de la ATCC de *Schizochytrium* sp., se reemplazó mediante recombinación homóloga después de transformación mediante electroporación con tratamiento enzimático previo (Véase el Ejemplo 7) con un vector que contenía el marcador de resistencia a Zeocin™ rodeado por secuencias de la región de flanco de *orfB*. Se generó una cepa mutante que carecía de un gen de *orfB* funcional. La cepa mutante era auxotrófica y requería suplemento de PUFA para su crecimiento.

El *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium* sp., (SEQ ID NO: 3) se clonó en el vector de expresión pDS04 para generar pREZ331. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flanco del locus del gen *orfB* nativo de la cepa 20888 de la ATCC de *Schizochytrium* sp.

El *orfB* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium* sp., se transformó con pREZ331 que contiene *PFA2*. Basándose en la integración aleatoria en el mutante, la producción de PUFA se restableció con *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium* sp., (SEQ ID NO: 3). Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. La cepa 20888 nativa de la ATCC de *Schizochytrium* sp., que contiene un gen *orfB* funcional produjo DHA y DPA n-6 en una proporción de 2,3: 1. La cepa recombinante que contiene *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium* sp., (SEQ ID NO: 3) como un reemplazo del gen *orfB* inactivado produjo DHA y DPA n-6 en una proporción de 3,5: 1. El contenido de EPA de la cepa recombinante era de un 0,8 % de FAME, el contenido de DPA n-3 era de un 0,1 %, el contenido de DPA n-6 era de un 7,1 %, y el contenido de DHA era de un 25,1 %.

El *orfB* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium* sp., también se transforma con pREZ331 que contiene *PFA2*, de modo que *PFA2* se inserta en el locus de *orfB* nativo y restablece la producción de PUFA.

EJEMPLO 10

El *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. (SEQ ID NO: 70) se volvió a sintetizar (DNA2.0) y se optimizó con codón para su expresión en *Schizochytrium* (SEQ ID NO: 121) y se clonó en un vector de expresión para generar pLR85. La optimización con codón se produjo usando la tabla de uso de codón de *Schizochytrium* en la FIG. 22. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flanco del locus del gen *orfB* nativo de la cepa 20888 de la ATCC de *Schizochytrium* sp.

El reemplazo de los genes *orf* también se estudió en una cepa hija de la cepa 20888 de la ATCC de *Schizochytrium* sp., con un aumento de la productividad de DHA. El gene *orfB* nativo en la cepa hija se reemplazó mediante recombinación homóloga después de transformación mediante electroporación con tratamiento enzimático previo (Véase el Ejemplo 7) con un vector que contenía el marcador de resistencia a Zeocin™ rodeado por secuencias de la región de flanco de *orfB*. Se generó una cepa mutante que carecía de un gen *orfB* funcional. La cepa mutante era auxotrófica y requería suplemento de PUFA para su crecimiento. La cepa mutante se transformó a través de electroporación con tratamiento enzimático previo (véase el Ejemplo 8) con pLR85 que contiene *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón (SEQ ID NO: 121). Basándose en regiones homólogas que flanquean el marcador de resistencia a Zeocin™ en el mutante y que flanquean el gen *PFA2* en pLR85, se produjo recombinación de doble cruce de modo que el *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón (SEQ ID NO: 121) se insertó en el locus de *orfB* nativo de la cepa mutante. La recombinación con el *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón (SEQ ID NO: 121) restableció la producción de PUFA en el *orfB* que carece de mutante de la cepa hija. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la

cepa recombinante era de un 1,0 % de FAME, el contenido de DPA n-3 era de un 0,3 %, el contenido de DPA n-6 era de un 7,0 %, y el contenido de DHA era de un 31,0 %.

En un experimento a realizar, el *orfB* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, del Ejemplo 9 se transforma mediante electroporación con tratamiento enzimático previo (véase el Ejemplo 8) con pLR85 que contiene *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 121). Basándose en regiones homólogas que flanquean el marcador de resistencia a Zeocin™ en el mutante que flanquea el gen *PFA2* en pLR85, se produce recombinación de doble cruce de modo que el *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 121) se inserta en el locus de *orfB* nativo. La recombinación con el *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 121) restablece la producción de PUFA en el *orfB* que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*

La cepa 20888 de la ATCC de *Schizochytrium sp.*, y *orfB* funcional que carece de mutantes de la cepa hija también se transforman con pLR85 que contiene *PFA2*, de modo que *PFA2* se integra de manera aleatoria en los mutantes y restablece la producción de PUFA en cada uno de los mutantes.

EJEMPLO 11

Un plásmido que contiene un casete de marcador de resistencia a paromomicina funcional en *Schizochytrium* se desarrolló para la cepa 20888 de la ATCC de *Schizochytrium sp.*, mediante el reemplazo de la región codificante del gen (*ble*) de resistencia a bleomicina/Zeoцин™ en pMON50000/pTUBZEO11-2 (documento de Patente de Estados Unidos N.º 7.001.772 B2) con el de la neomicina fosfotransferasa II (*npt*), originalmente a partir del transposón Tn5 bacteriano. En pMON50000, el gen de resistencia *ble* está dirigido por el promotor de α -tubulina de *Schizochytrium* iba seguido del terminador de la transcripción SV40. La región de *ble* en pMON50000 incluye un sitio de restricción *NcoI* en el codón de inicio de ATG y un sitio de restricción *PmlI* que sigue inmediatamente a la señal de parada de TGA. La PCR se usó para amplificar la región codificante de *npt* presente en pCaMVnpt (Shimizu *et al.*, *Plant J.* 26 (4): 375 (2001)) de modo que el producto incluía un sitio de restricción *BspHI* (subrayado a continuación, cebador CAX055) en el ATG de inicio (en negrita) y un sitio de restricción *PmlI* (subrayado a continuación, cebador CAX056) que sigue inmediatamente a la señal de parada (en negrita - complemento):

CAX055 (directo): GTCATGATTGAACAAGATGGATTGCAC (SEQ ID NO: 66)

CAX056 (inverso): CCACGTGTCAGAAGAAGCTCGTCAAGAA (SEQ ID NO: 67).

La PCR se realizó con el kit de polimerasa TaqMaster (5Prime), los productos se clonaron en pCR4-TOPO (Invitrogen), y los plásmidos resultantes se transformaron en TOP10 de *E. coli* (Invitrogen). El análisis de la secuencia de ADN usando cebadores de vector identificó múltiples planes que contienen la estructura de 805 pb deseada (es decir, las secuencias se emparejan con las del molde de la fuente más los sitios de restricción modificados por ingeniería). La región codificante *npt* modificada se aisló mediante digestión con enzimas de restricción *BspHI* más *PmlI*, y el fragmento de ADN purificado se ligó con un fragmento de vector de pMON50000 generado por digestión con enzimas *NcoI* más *PmlI*. Las enzimas de restricción *BspHI* y *NcoI* dejan extremos de solapamiento compatibles, y *PmlI* deja extremos romos. El plásmido resultante, pTS-NPT, contiene el gen *npt* de resistencia a neomicina/paromomicina en el contexto idéntico al del gen *ble* original en pMON50000.

Para evaluar la función del nuevo casete de resistencia a paromomicina en pTS-NPT se usó bombardeo de partículas de *Schizochytrium* (documento de Patente de Estados Unidos N.º 7.001.772 B2). La selección para resistencia a paromomicina (PAR) se realizó en placas de agar que contenían 50 μ g/ml de sulfato de paromomicina (Sigma). Los transformantes de *Schizochytrium* resistentes a paromomicina se encontraron a frecuencias similares a las de resistencia a Zeocin™ de pMON50000. El casete "terminador de promotor/*npt*/SV40 de α -tubulina" se puede liberar de pTS-NPT con diversas enzimas de restricción para su uso posterior en otros esfuerzos de desarrollo.

EJEMPLO 12

El gen *orfC* nativo en la cepa 20888 de la ATCC de *Schizochytrium sp.*, se reemplazó mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a paromomicina rodeado por secuencias de la región de flaqueo de *orfC*. Se generó una cepa mutante que carecía de un gen *orfC* funcional. La cepa mutante era auxotrófica y requería suplemento de PUFA para su crecimiento.

El *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5) se clonó en el vector de expresión pREZ22 para generar pREZ324. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flaqueo del locus del gen *orfC* nativo de la cepa 20888 de la ATCC de *Schizochytrium sp.*

El *orfC* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, se transformó con pREZ324 que contenía *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* Basándose en regiones homólogas que flanquean el marcador de resistencia a paromomicina en el mutante y que flanquea al gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* en pREZ324, se produjo recombinación de doble cruce de

modo que *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* se insertó en el locus de *orfC* nativo. La recombinación homóloga con *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5) restableció la producción de PUFA en el *orfC* que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.* Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. La cepa 20888 nativa de la ATCC de *Schizochytrium sp.*, que contenía un gen *orfC* funcional produjo DHA y DPA n-6 en una proporción de 2,3: 1. La cepa recombinante que contenía *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5) en lugar del gen *orfC* inactivado produjo DHA y DPA n-6 en una proporción de 14: 9, lo que demuestra adicionalmente que el perfil de PUFA de *Schizochytrium* se puede alterar con las moléculas de ácido nucleico que se describen en el presente documento. El contenido de EPA de la cepa recombinante era de un 1,2 % de FAME, el contenido de DPA n-3 era de un 0,2 %, el contenido de DPA n-6 era de un 2,9 %, y el contenido de DHA era de un 43,4 %.

El *orfC* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, también se transformó con pREZ324 que contenía *PFA3*, de modo que *PFA3* se integró de manera aleatoria en el mutante y restableció la producción de PUFA. El contenido de EPA de la cepa recombinante era de un 1,2 % de FAME, el contenido de DPA n-3 era de un 0,2 %, el contenido de DPA n-6 era de un 2,5 %, y el contenido de DHA era de un 39,1 %.

El gen *orfC* nativo en la cepa hija discutida en el Ejemplo 10 se reemplazó mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a paromomicina rodeado por secuencias de la región de flanco de *orfC*. Se generó una cepa mutante que carecía de un gen *orfC* funcional. La cepa mutante era auxotrófica y requería suplemento de PUFA para su crecimiento. El *orfC* funcional que carece de mutante se transformó con pREZ324. Se produjo recombinación de doble cruce de modo que *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* se insertó en el locus de *orfC* nativo de la cepa mutante. La recombinación homóloga con *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* (SEQ ID NO: 5) restableció la producción de PUFA en el *orfC* que carece de mutante de la cepa hija. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 1,2 % de FAME, el contenido de DPA n-3 era de un 0,3 %, el contenido de DPA n-6 era de un 2,8 %, y el contenido de DHA era de un 43,1 %.

El *orfC* funcional que carece de mutante de la cepa hija también se transforma con pREZ324 que contiene *PFA3*, de modo que *PFA3* se integra de manera aleatoria en el mutante y restablece la producción de PUFA.

EJEMPLO 13

El *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 72) se volvió a sintetizar (DNA2.0) y se optimizó con codón para su expresión en *Schizochytrium* (SEQ ID NO: 122) y se clonó en el vector de expresión pREZ22 para generar pREZ337. La optimización con codón se produjo usando la tabla de uso de codón de *Schizochytrium* en la FIG. 22. El vector de expresión contenía aproximadamente 2 kb de ADN de la región de flanco del locus del gen *orfC* nativo de la cepa 20888 de la ATCC de *Schizochytrium sp.*

El *orfC* funcional que carece de mutante de la cepa hija del Ejemplo 12 se transformó a través de electroporación con tratamiento enzimático previo (véase el Ejemplo 8) con pREZ337 que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122). Basándose en regiones homólogas que flanquean el marcador de resistencia a Zeocin™ en el mutante y que flanquean el gen *PFA3* en pREZ337, se produjo recombinación de doble cruce de modo que el *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122) se insertó en el locus de *orfC* nativo. La recombinación con *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122) restableció la producción de PUFA en el *orfC* que carece de mutante de la cepa hija. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 1,3 % de FAME, el contenido de DPA n-3 era de un 0,4 %, el contenido de DPA n-6 era de un 2,7 %, y el contenido de DHA era de un 50,2 %.

En un experimento a realizar, el *orfC* funcional que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*, del Ejemplo 12 se transforma mediante electroporación con tratamiento enzimático previo (véase el Ejemplo 8) con pREZ337 que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122). Basándose en regiones homólogas que flanquean el marcador de resistencia a Zeocin™ en el mutante y que flanquean el gen *PFA3* en pREZ337, se produce recombinación de doble cruce de modo que el *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122) se inserta en el locus de *orfC* nativo. La recombinación con *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón (SEQ ID NO: 122) restablece la producción de PUFA en el *orfC* que carece de mutante de la cepa 20888 de la ATCC de *Schizochytrium sp.*

La cepa 20888 de la ATCC de *Schizochytrium sp.*, y el *orfC* funcional que carece de mutantes de la cepa hija también se transforman con pREZ337 que contiene *PFA3*, de modo que *PFA3* se integra de manera aleatoria en los mutantes y restablece la producción de PUFA en cada uno de los mutantes.

EJEMPLO 14

Cualesquiera dos o los tres de genes *orfA*, *orfB*, y *orfC* en la cepa 20888 de la ATCC de *Schizochytrium sp.*, se reemplazan mediante recombinación homóloga después de transformación con vectores que contienen cualquiera del marcador de resistencia a Zeocin™ o paromomicina rodeado por secuencias de la región de flanco de *orf* apropiada. Se generan cepas mutantes que carecen de genes funcionales para cualesquiera dos o los tres de *orfA*, *orfB*, y *orfC*. Las cepas mutantes son auxotróficas y requieren suplemento de PUFA para su crecimiento.

Los mutantes de la cepa 20888 de la ATCC de *Schizochytrium sp.*, que carecen de genes funcionales de *orf* se transforman con uno o más vectores de expresión que contienen los genes *PFA* correspondientes (una o más de las SED ID NOs: 1, 3, 5, 120, 121, o 122). Basándose en regiones homólogas que flanquean los marcadores de resistencia a Zeocin™ o paromomicina en los mutantes y que flanquean los genes *PFA* en los respectivos vectores de expresión, se puede producir recombinación de doble cruce de modo que los genes *PFA* se insertan en los loci de *orf* nativos. La integración aleatoria de estos vectores de expresión también se puede producir con la selección de transformantes basándose únicamente en el restablecimiento de la producción de PUFA. La recombinación homóloga con genes *PFA* que establece la producción de PUFA en los mutantes, de modo que los perfiles de PUFA nativo se restablecen o se alteran basándose en la combinación de genes *PFA* insertados en los mutantes.

En un experimento realizado, la cepa 20888 de la ATCC de *Schizochytrium sp.*, del Ejemplo 12 que carece de un gen *orfC* funcional y que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, integrado de manera aleatoria (SEQ ID NO: 5) se usó para el reemplazo de los genes *orfA* y *orfB*. Los genes *orfA* y *orfB* nativos en la cepa se reemplazaron mediante recombinación homóloga después de transformación con un vector que contiene el marcador de resistencia a Zeocin™ rodeado por secuencias de las regiones de flanco de *orfA* y *orfB*. Se generó una cepa que carecía de *orfA*, *orfB*, y *orfC* funcional, y que contenía *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, integrado de manera aleatoria. La cepa se transformó con pREZ345 que contenía *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 1) y pREZ331 que contenía *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 3) de modo que se produjo la integración aleatoria de *PFA1* y *PFA2*. La cepa recombinante resultante carecía de *orfA*, *orfB*, y *orfC* funcional y contenía interacciones aleatorias de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 6,6 % de FAME, el contenido de DPA n-3 era de un 0,8 %, el contenido de DPA n-6 era de un 1,6 %, y el contenido de DHA era de un 20,9 %.

En otro experimento realizado, la cepa hija del Ejemplo 12 que carece de un gen *orfC* funcional y que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 5) insertado en el locus de *orfC* nativo se usó para el reemplazo de los genes *orfA* y *orfB*. Los genes *orfA* y *orfB* nativos en la cepa se reemplazaron mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a paromomicina rodeado por secuencias de las regiones de flanco de *orfA* y *orfB*. Se generó una cepa que carecía de *orfA*, *orfB*, y *orfC* funcional, y que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, insertado en el locus de *orfC* nativo. La cepa se transformó con pREZ345 que contenía *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 1) y pREZ331 que contenía *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 3). Se produjeron recombinaciones de doble cruce de modo que *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se insertó en el locus de *orfA* nativo y *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se insertó en el locus de *orfB* nativo. La cepa recombinante resultante carecía de *orfA*, *orfB*, y *orfC* funcional y contenía *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, insertados en los respectivos loci de *orfA*, *orfB*, y *orfC*. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 7,3 % de FAME, el contenido de DPA n-3 era de un 0,4 %, el contenido de DPA n-6 era de un 1,5 %, y el contenido de DHA era de un 23,9 %.

En otro experimento realizado, la cepa hija del Ejemplo 12 que carece de un gen *orfC* funcional y que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, integrado de manera aleatoria (SEQ ID NO: 5) se usó para el reemplazo de los genes *orfA* y *orfB*. Los genes *orfA* y *orfB* nativos en la cepa se reemplazaron mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a Zeocin™ rodeado por secuencias de las regiones de flanco de *orfA* y *orfB*. Se generó una cepa que carecía de *orfA*, *orfB*, y *orfC* funcional, y que contenía *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, integrado de manera aleatoria. La cepa se transformó con pREZ345 que contenía *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 1) y pREZ331 que contenía *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 3) de modo que se produjo la integración aleatoria de *PFA1* y *PFA2*. La cepa recombinante resultante carecía de *orfA*, *orfB*, y *orfC* funcional y contenía interacciones aleatorias de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.* Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 6,2 % de FAME, el contenido de DPA n-3 era de un 1,3 %, el contenido de DPA n-6 era de un 0,9 %, y el contenido de DHA era de un 16,6 %.

En otro experimento realizado, la cepa hija del Ejemplo 13 que carece de un gen *orfC* funcional y que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 122) insertado en el locus de *orfC* nativo se usó para el reemplazo de los genes *orfA* y *orfB*. Los genes *orfA* y *orfB* nativos en la cepa se reemplazaron mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a paromomicina rodeado por secuencias de las regiones de flanco de *orfA* y *orfB*. Se generó una cepa que carecía de *orfA*, *orfB*, y *orfC* funcional, y que contenía *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, insertado en el locus de *orfC* nativo. La cepa se transformó con pLR95 que contenía *PFA1* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 120) y pLR85 que contenía *PFA2* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 121). Se produjeron recombinaciones de doble cruce de modo que *Schizochytrium sp.*, ATCC PTA-10212 *PFA1* se insertó en el locus de *orfA* nativo y *PFA2* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, se insertó en el locus de *orfB* nativo. La cepa recombinante resultante carecía de *orfA*, *orfB*, y *orfC* funcional y contenía *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, insertados en los respectivos loci de *orfA*, *orfB*, y *orfC*. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 5,2 % de FAME, el contenido de DPA n-3 era de un 0,6 %, el contenido de DPA n-6 era de un 2,1 %, y el contenido de DHA era de un 47,1 %.

En otro experimento realizado, la cepa hija del Ejemplo 13 que carece de un gen *orfC* funcional y que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, Integrado de manera aleatoria (SEQ ID NO: 122) se usó para el reemplazo de los genes *orfA* y *orfB*. Los genes *orfA* y *orfB* nativos en la cepa se reemplazaron mediante recombinación homóloga después de transformación con un vector que contenía el marcador de resistencia a Zeocin™ rodeado por secuencias de las regiones de flanco de *orfA* y *orfB*. Se generó una cepa que carecía de *orfA*, *orfB*, y *orfC* funcional, y que contenía *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, integrado de manera aleatoria. La cepa se transformó con pLR95 que contenía *PFA1* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 120) y pLR85 que contenía *PFA2* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.*, optimizado por codón (SEQ ID NO: 121) de modo que se produjo la integración aleatoria de *PFA1* y *PFA2*. La cepa recombinante resultante carecía de *orfA*, *orfB*, y *orfC* funcional y contenía interacciones aleatorias de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-10212 de la ATCC de *Schizochytrium sp.* Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 7. El contenido de EPA de la cepa recombinante era de un 1,8 % de FAME, el contenido de DPA n-3 era de un 1,8 %, el contenido de DPA n-6 era de un 2,3 %, y el contenido de DHA era de un 34,1 %.

EJEMPLO 15

Los genes *orfA*, *orfB*, y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, se clonaron en una serie de vectores Duet (Novagen). Los vectores de expresión Duet son un conjunto de plásmidos compatibles en los que se clonan y se coexpresan múltiples genes diana a partir del promotor inducible de T7 en *E. coli*. El plásmido pREZ91 de Duet contenía que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.* en pETDuet-1; plásmido pREZ96 de Duet contenía la cepa 20888 de la ATCC de *Schizochytrium sp.*, *orfB* en pCDFDuet-1; y plásmido pREZ101 de Duet contenía la cepa 20888 de la ATCC de *Schizochytrium sp.*, *orfC* en pCOLADuet-1. Los plásmidos pREZ91, pREZ96, y pREZ101 de Duet, junto con el plásmido pJK737, que contenía el gen *HetI* auxiliar requerido (descrito en el documento de Patente de Estados Unidos N.º 7.217.856, incorporado por referencia en el presente documento en su totalidad), se transformaron en la cepa BLR(DE3) de *E. coli*, que contenía un gen de T7 ARN polimerasa inducible. Después del crecimiento celular y la adición de IPTG, de acuerdo con las instrucciones del fabricante (Novagen), se produjeron DHA y DPA n-6. En resumen, se añadió IPTG 1 mM a la inducción cuando las células alcanzaron una densidad óptica de aproximadamente 0,5 a 600 nm. Las células se cultivaron durante 12 horas a 30 °C en caldo de cultivo Luria y se cosecharon. Los ácidos grasos se convirtieron en ésteres de metilo usando técnicas convencionales. Los perfiles de ácido graso se determinaron usando cromatografía de gases con detección de ionización a la llama (GC-FID) como ésteres de metilo de ácido graso (FAME).

El gen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 1) se clonó en el vector de expresión pETDuet-1, generando pREZ346. Los plásmidos Duet pREZ346 (que contienen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ96 (que contiene *orfB*), y pREZ101 (contiene *orfC*) se transformaron en la cepa BLR(DE3) de *E. coli* junto con pJK737 (que contenía *HetI*). El gen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, se coexpresaba con los genes *orfB* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.* La expresión de *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en combinación con *orfB* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoyaban la producción de DHA en *E. coli* en condiciones de inducción. El contenido de DHA de la *E. coli* transformada era de un 2,8 % de FAME, el contenido de DPA n-6 era de un 1,1 %, el contenido de DPA n-3 era de un 0,6 %, y el contenido de EPA era de un 3,7 %.

EJEMPLO 16

El gen de *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 120) optimizado con codón se clonó en el vector de expresión pETDuet-1, generando pLR100. Los plásmidos Duet pLR100 (que contiene *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pREZ96 (que contiene *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ101 (que contiene *orfC* de la cepa 20888 de la

ATCC de *Schizochytrium sp.*) se transforman en la cepa BLR(DE3) de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. El gen *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* se coexpresa con los genes *orfB* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.* La expresión de *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en combinación con *orfB* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA y EPA en *E. coli* en condiciones de inducción.

EJEMPLO 17

El gen de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 5) se clonó en el vector de expresión pCOLADuet-1, generando pREZ326. Los plásmidos Duet pREZ326 (que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ96 (que contiene *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en combinación con *orfA* y *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,3 % de FAME.

EJEMPLO 18

El gen de *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 122) optimizado con codón se clonó en el vector de expresión pCOLADuet-1, generando pREZ348. Los plásmidos Duet pREZ348 (que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ96 (que contiene *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, en combinación con *orfA* y *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 2,9 % de FAME y el contenido de DPA n-6 era de un 0,4 %.

EJEMPLO 19

El gen de *PFA2* PTA-9695 de la ATCC de *Schizochytrium sp.*, (SEQ ID NO: 3) se clonó en el vector de expresión pCDFDuet-1, generando pREZ330. Los plásmidos Duet pREZ330 (que contienen *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ326 (que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), y pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 9. La expresión de *PFA2* y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en combinación con *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,8 % de FAME y el contenido de DPA n-6 era de un 0,2 %.

EJEMPLO 20

El gen de *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* (SEQ ID NO: 121) optimizado con codón se clonó en el vector de expresión pCDFDuet-1, generando pLR87. Los plásmidos Duet pLR87 (que contienen *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pREZ348 (que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), y pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA2* y *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón, en combinación con *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya niveles de DHA y niveles bajos de producción de EPA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 4,4 % de FAME, el contenido de DPA n-6 era de un 1,1 %, y el contenido de EPA era de un 0,1 %.

EJEMPLO 21

Los plásmidos Duet pREZ346 (que contienen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ330 (que contiene *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), y pREZ326 (que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,3 % de FAME y el contenido de EPA era de un 0,3 %.

EJEMPLO 22

Los plásmidos Duet pLR100 (que contienen *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pLR87 (que contiene *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), y pREZ348 (que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón) se transforman en la cepa BLR(DE3) de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA1*, *PFA2*, y *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*, optimizados con codón apoya la producción de DHA y EPA en *E. coli* en condiciones de inducción.

EJEMPLO 23

Los plásmidos Duet pREZ330 (que contienen *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ101 (que contiene *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en combinación con *orfA* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,6 % de FAME y el contenido de DPA n-6 era de un 0,3 %.

EJEMPLO 24

Los plásmidos Duet pLR87 (que contienen *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pREZ91 (que contiene *orfA* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ101 (que contiene *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón, en combinación con *orfA* y *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya niveles de DHA y niveles bajos de producción de EPA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 1,7 % de FAME, el contenido de DPA n-6 era de un 0,9 %, y el contenido de EPA era de un 0,1 %.

EJEMPLO 25

Los plásmidos Duet pREZ346 (que contienen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ330 (que contiene *PFA2* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), y pREZ101 (que contiene *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA1* y *PFA2*, en combinación con *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,3 % de FAME, el contenido de DPA n-6 era de un 0,1 %, y el contenido de EPA era de un 0,5 %.

EJEMPLO 26

Los plásmidos Duet pLR100 (que contienen *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), pLR87 (que contiene *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizado con codón), y pREZ101 (que contiene *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*) se transforman en la cepa BLR(DE3) de *E. coli* junto con pJK737 (que contiene *HetI*). Véase el Ejemplo 15. La expresión de *PFA1* y *PFA2* de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.* optimizados con codón, en combinación con *orfC* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA y EPA en *E. coli* en condiciones de inducción.

EJEMPLO 27

Los plásmidos Duet pREZ346 (que contienen *PFA1* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*), pREZ96 (que contiene *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*), y pREZ326 (que contiene *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*) se transformaron en BLR(DE3) de la cepa de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA1* y *PFA3* de la cepa PTA-9695 de la ATCC de *Schizochytrium sp.*, en combinación con *orfB* de la cepa 20888 de la ATCC de *Schizochytrium sp.*, apoya la producción de DHA en *E. coli* en condiciones de inducción. Las células se cultivaron y se analizaron para los FAME como se describe en el Ejemplo 15. El contenido de DHA de la *E. coli* transformada era de un 0,1 % de FAME y el contenido de EPA era de un 0,1 %.

EJEMPLO 28

Los plásmidos Duet pLR100 (que contienen *PFA1* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón), pREZ96 (que contiene *orfB* de la cepa 20888 de la ATCC de *Schizochytrium* sp.), y pREZ348 (que contiene *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. optimizado con codón) se transforman en la cepa BLR(DE3) de *E. coli* junto con pJK737 (que contenía *HetI*). Véase el Ejemplo 15. La expresión de *PFA1* y *PFA3* de la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp., optimizados con codón en combinación con *orfB* de la cepa 20888 de la ATCC de *Schizochytrium* sp., apoya la producción de DHA y EPA en *E. coli* en condiciones de inducción.

EJEMPLO 29

Las actividades de PUFA sintasa de Pfa1p, Pfa2p, y Pfa3p en la cepa PTA-9695 de la ATCC de *Schizochytrium* sp., y en la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp. presentan supresión genética de manera individual mediante procedimientos convencionales. Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 7.217.856, incorporado por referencia en el presente documento en su totalidad.

El marcador de resistencia a Zeocin™, higromicina, blastidina, u otro marcador de resistencia apropiado se inserta en un sitio de restricción del gen *PFA1* (SEQ ID NO: 1 o SEQ ID NO: 68) que está contenido en un plásmido. Después de la inserción del marcador de resistencia, el plásmido se introduce en la cepa PTA-9695 de la ATCC de *Schizochytrium* sp., o en la cepa PTA-10212 de la ATCC de *Thraustochytrium* sp., respectivamente, mediante bombardeo de partículas, electroporación, u otro método de transformación apropiado. La recombinación homóloga se produce, generando mutantes en los que el gen *PFA1* nativo se reemplaza o se interrumpe con el marcador de resistencia a Zeocin™, higromicina, blastidina, u otro marcador de resistencia apropiado. Los transformantes se seleccionan en placas que contienen Zeocin™, higromicina, blastidina, u otro agente de resistencia apropiado, suplementado con los PUFA. Las colonias se examinan adicionalmente para la capacidad para crecer en ausencia de suplemento de PUFA. El ADN genómico se aísla de las colonias que son resistentes al agente de selección e incapaces de crecer en ausencia de suplemento de PUFA. La PCR y el de análisis de Transferencia de Southern del ADN se realiza para demostrar que el gen *PFA1* ha experimentado delección o interrupción.

PFA2 presenta supresión genética con procedimientos similares. Se encuentra que los mutantes con supresión genética resultantes que requieren suplemento de PUFA carecen de *PFA2* de longitud completa.

PFA3 presenta supresión genética con procedimientos similares. Se encuentra que los mutantes con supresión genética resultantes que requieren suplemento de PUFA carecen de *PFA3* de longitud completa.

LISTADO DE SECUENCIAS

<110> APT, Kirk E.
 RICHTER, Leslie
 SIMPSON, David
 ZIRKLE, Ross

<120> Moléculas de ácido nucleico de ácido graso poliinsaturado sintasas, y polipéptidos, composiciones y métodos de preparación y usos de los mismos

<130> 2715.025PC02/JUK/SAS

<140> A asignar
 <141> 19-03-2009

<150> 61/296.460
 <151> 19-01-2010

<150> 61/161.742
 <151> 19-03-2009

<160> 127

<170> PatentIn versión 3.3

<210> 1
 <211> 7824
 <212> ADN
 <213> *Schizochytrium* sp.

ES 2 623 610 T3

<400> 1

```

atggatactc gcatcgcgat cgtgggggatg tcggcgatcc tgccgagcgg ggagAACgtg      60
cgcgagagct gggaggcgat ccgcatggg ctggattgcc tgagcgatct gccggcggac      120
cgcggtggacg tgacggccta ctacaacccg gagaagacga ccaaggacaa gatctactgc      180
aagcgcgggc gggtcatccc ggagtagcac ttcgacgcgc gtgagttcgg gctcaacatg      240
ttccagatgg aggactcggg cgccaaccag acgatctcgc tgctcaaggt gaaggaggcg      300
ctgacggacg ccaacatccc ggcgttctcg agcggtaaga agaacatcgg ctgcgtgctg      360
ggcatcggcg gcggccagaa ggcgagccac gagttctact cgcggtcaa ctacgtggtc      420
gtggacaagg tgctgcgcaa gatgggcctg ccggagggaag acgtggcggc ggcggtggac      480
aagtacaagg cgagtttccc cgagtggcgc ctcgactctt tccccgggtt cctgggcaac      540
gtcacggcgg ggcgctgctg caataccttc aacatggagg gcatgaactg cgtcgtggac      600
gcggcctgcg cgtcgtcgtc gatcgcggtc aaagtggcga tcgaggagct gctctacggc      660
gactgcgatg cgatgatcgc ggggtgccacc tgcacggaca actcgatcgg gatgtacatg      720
gccttctcca agacgcccggt gttttccacg gaccogagcg tcaaggcgta cgacgccgcc      780
accaaaggca tgctcatcgg cgagggctcg gcgatgctcg tgctgaagcg ctacgcggac      840
gccgtgcgcg acggcgacac cgtgcacgcc gtcatcaagg ggtgcgcgtc ctcgagcgac      900
ggcaaggcgg cgggcatcta cacgccgaca atctcggggc aggaggaggc cctgcgccgc      960
gcctacgccc gcgccaatgt cgaccgggcc actgtgacgc tgggtggaggg ccacggcacg     1020

```

ES 2 623 610 T3

ggtagcggcg tgggagacaa gatcgagctg acggcgctga gcaacctctt ctccaaggcg 1080
 ttttctgcca acggtggcg cgcgaggaa gcagagcagg tggcggtggg cagcatcaag 1140
 tcgcagatcg ggcacctcaa ggcggtggcc gggtggccg ggtggtcaa ggtggtgctg 1200
 gcgctcaagc acaagacgct gccgcagacg atcaacgtcg acaagccgc gtcgctggtg 1260
 gacgggaccc cgatccagca gtgcgcgctg tacgtcaaca cgatgaaccg cccctggttc 1320
 acgcccgtag ggggtgccgc cgcgcccggc gtgtcgctgt ttgggtttgg cgggtccaac 1380
 taccacgccc tgctggagga gtttgagccc gagcacgaga gcgctaccg gtacaacaac 1440
 ctgcgcgagg tggcgctgct gcacggggg gacgtcgca ccttggcggc gacggttcgc 1500
 gccaagctgg cgctggccac cgcgagcag gaagaggcg gtgtggtgaa gaacgggac 1560
 tacatcgct accaccggtt cctggacgag tgcaagttgc gcggcgctgt gccgcaggcg 1620
 cacgcgcggg tgggactgct cgtacgggac ctgagctcgc tcatcgccgt gctcgaggcc 1680
 gotgccgcca agctcgcggg cgaagagagc gcgacggagt ggacggtcag cgttgctacg 1740
 ggcgaggcg ccttcgcgt gcgcggtgtg gctacggagg ccaacgtggc ggcgctgttc 1800
 tcggggccagg gcgcgcagta cacgcacatg ttcagcgacg tggcgatgaa ctggcccccg 1860
 ttccgcgaga gcgtcgccgc catggaccgc gccagcgcg agcgcttcgg gcggcctgcc 1920
 aagcgctga gcagcgtgct gtaccgcgc aagcgtacg gcgacgaacc gcggcaggac 1980
 cacaaggaga tctcgaaaac gcgtactcg cagcccgcaa cgctcgctg ctcggtcggc 2040
 gcctttgaca tcttcaaagc ggcgggactg gcgcgagct ttgcggcggg ccaactcgctg 2100
 ggcgagtttg cggcgctcta cgcgcccggg tcgctcgatc gcgacgccgt cttcgacctg 2160
 gtctgcgcgc gcgccaaggc catgagcgac ttcacggccc aggcagcag cagcggtggc 2220
 gccatggcg cgtgattgg cgccaaggcg gaccagctct cgctgggtgg cgcgcccgac 2280
 gtgtggctcg ccaacagcaa ctgcacctcg cagaccgtga tcacgggaac cgccgaagca 2340
 gtggctgcgc cctctgacaa gttgcgctgc agcggaact tccgctcgt gcctctggcc 2400
 tgcgaggcg ccttccactc gccgcacatg cgcgcgcgcg agcagacgtt tgcgtcggcg 2460
 ctgcgcgagg cggccgtgtc ggcaccggcg gctgctcgt tctactctaa cgtgacgggg 2520
 ggcgcgcgg taacctcgcc cgcgacgtc aaaacgaacc tgggcaagca catgacgagc 2580
 cctgtgcagt tcgtgcagca ggtgcgagcc atgcacgcgg cgggcgcgcg tgtgtttgtg 2640
 gagtttgggc ccaagcaggt cctgtcgcg ctcgtcaagg agaccttg cgaggccggc 2700
 gacgtggtca cggtcgccgt caaccagac tgggccaagg acagcgacac gcagctgcgc 2760
 caggcggcgc tcacgttggc ggtcgccggc gtgcgctca aggactttga ccgctggcag 2820
 ctgcgggatg ccacgcgcct cgagcctgtc aagaagaaga agaccacgtt gcggctctcg 2880
 gcagccacct acgtctccgc caagacgtt cgccagcgcg aggcgtgct caacgacggc 2940

ES 2 623 610 T3

tacactgtca gtggtgccac ggcggtagtc aaggaagtgg acacggccaa cgaggagcgt 3000
 ctctgccgcc aagccagga tctccagcgc cagctcgcgg aggcctcgac ggcagcccag 3060
 gcggcgagct ccaaggtcgc ggagctcgag cgcacgatcc aggaacttga gcgcaaggtg 3120
 cagcagcagc agcaagagaa ggggtgagaac tcagacagca acgctgccgc cgaagtgtg 3180
 cggcgccaca aggagctgct ccagcgcgtg ctgcaggact gtgacgagca ggcagtgtcc 3240
 gtagccacgg tggttccgac acctacgtcc tccccgacgc ctacatctc acccgatatcc 3300
 ggcaacagca agagcactcg tggcagtgtc gatctgcaag cgctgctggc caaggcggag 3360
 actgtggtga tggctgtgct ggctgccaaag actggctacg aggcgcgacat ggttgaggcg 3420
 gacatggacc tggaggccga gctcggcatc gactcgatca agcgcgtgga gatcctttcc 3480
 gaggtgcagg gccagctggg cgtcgaggcc aaggacgtgg atgcgctgag ccgcacgcgc 3540
 acggtcggtg aggttgtgga cgcctgaag gcggagatcg tggctgcctc tgggtgtagt 3600
 gctcctgcgg ttccttcggc gcccgctgct tctgcagctc cgactccgc tgcctcgact 3660
 gcgccttctg ctgatctgca agcgtgtgtg tccaaggcgg agactgtggt gatggctgtg 3720
 ctggcgccca agactggcta cgaggccgac atggtcgagg cggacatgga cctggaggcc 3780
 gagctcgga tcgactcgat caagcgcgtg gagatcctc cggaggtgca gggccagctg 3840
 ggcgtcgagg ccaaggacgt ggatgcgctg agccgcacgc gcacggctcg tgaggttgtg 3900
 gatgccatga aggcggaaat cgtggctgcc tctgctggtg gtgctcctgc tctgctgtt 3960
 ccttcggcgc ccgtgtctt tgcagctccg actcccgtg cttcgactgc gccttctgct 4020
 gatctgcaag cgctgctgtc caaggcggag acggtggtga tggctgtgct ggcgcccaag 4080
 actggtacg aggcgcgacat ggtcgaggcg gacatggacc tggaggccga gctcggcatc 4140
 gactcgatca agcgcgtgga gatcctctcg gaggtgcagg gccagctggg cgtcgaggcc 4200
 aaggacgtgg atgcgctgag ccgcacgcgc acggtcgggtg aggttgtgga tgccatgaag 4260
 gcggaaatcg tggctgcctc tgggtgtagt gctcctgctc ctgcgggtcc ttggcgccc 4320
 gctgcttctg cagctccgac tcccgcggct gcgacagcgc cttctgctga tctgcaagcg 4380
 ctgctggcca aggcggagac tgtggtgatg gctgtgctgg cgccaagac tggctacgag 4440
 gccgacatgg tcgaggcgga catggacctg gaggcgcgac tcggcatcga ctogatcaag 4500
 cgcgtggaga tcctttccga ggtgcagggc cagctgggcg tcgaggccaa ggacgtagat 4560
 gcgtgagcc gcacgcgcac ggtcggtagg gttgtggatg ccatgaaggc ggagatcgtg 4620
 gctgcctctg ctggtagtgc tectgtcct gctgttcctt cggcgccgc tgcctctgca 4680
 gctccgactc ccgtgtctt gactgcgcct tctgtgctg tcgaagcgt gctgtccaag 4740
 gcggagactg tggtagtggt tgtgctggcg gccaaactg gctacgaggc cgacatggtc 4800
 gaggcggaca tggacctgga ggccgagctc ggcacgact cgatcaagcg cgtggagatc 4860
 ctctcggagg tgcaggccca gctgggcgtc gaggccaaag acgtggatgc gctgagccgc 4920

acgcgcacgg tcggtgaggt tgtggatgcc atgaaggcgg aaatcgtggc tgcctctggt 4980
 ggtagtgctc ctgctgctgc tgttccttcg gcgcccgtcg cttctgcagc tccgactcct 5040
 gcgactgcgc cttctgctga tctgcaagcg ctgctgtcca aggcggagac tgtggatgatg 5100
 gctgtgctgg cggccaagac tggctacgag gccgacatgg tcgagggcga catggacctg 5160
 gaggccgagc tcggcatcga ctcgatcaag cgcgtggaga tcctttccga ggtgcagggc 5220
 cagctgggcg tcgaggccaa ggacgtagat gcgctgagcc gcacgcgcac ggtcggtgaa 5280
 gtggtggacg ccatgaaggc ggagatcgtg gctgcctctg gtggtagtgc tcctgctgct 5340
 ccttcggcgc ccgcgttct tccaaacgtg tttgggtccg agtgcgagga cctgtctctg 5400
 acctttcccg tgataacgac cctgccgtct cctgcagagc ttgtgctggc cgagggcggc 5460
 gctcgcctcg tagtcgtggt ggatgatgga tctgcaactc cctcgtcgtc ggtgtcctcg 5520
 ctcgccgctc gtgcggtgct gctgcaggcg cagtcttctc ctgcctgctc gccgcgctcg 5580
 accacgcaca agttggtgac cgtacgagac cgctctgaag cggcgctaca ggcggcgctc 5640
 acgtccgtcg aggcgcagtt cggcaaggcg ggtggctttg tgttcagtt cggcgacgac 5700
 gacgtgcaag cgcagctcgg ctgggcgctg ctgcgggcca agcacctcaa aacttcgctg 5760
 tcagaacaga tcgagggcgg tcgcaccttt ttctggtccg tcgcggcgct cgacggccag 5820
 ctggggctct ccggcaagtc gacgaccgct accgttgatc tctcccgcg gcagcagggc 5880
 agcgtgttcg gcctgtgcaa gacactcgac ctggagtggc ccgctgtctt ctgcccgga 5940
 atcgacctgg ccgccgacct cgacgccgca caggcccgcg ggtgcctgct gggcgagctg 6000
 tcagaccccg acgtggcgtg gcgcgagctt ggttactccg cctcgggcca gcctgcacg 6060
 acaactacga agtcgctgac tacgggcaag ccgcaccagc cgatctctc gtccgacctc 6120
 tttctggtgt cgggcggcgc gcgcggcatc acccgctgt gcgtgcgga gctggcgag 6180
 cgcgtggcg gcggcaagta cgtgctcatc ggccgctcgg agctgccac gacggagcct 6240
 gcctgggcgg tcggcgtgga gtctggcaag ccgctggaga aggcgcgct ggcgttctg 6300
 aaggcggagt ttgcagcggg ccgcggggcc aagccgacgc cgatgctgca caagaagctc 6360
 gtgggcgcgg tggtcggagc gcgcgaggtg cgagcctcgc tcgccgagat cactgcacag 6420
 ggcgccacgg ctgtgtacga gtctgctgac gtgagctctg ccgccaaagg gcgtgagatg 6480
 gtagagcgcg tgcagcagca gggcgggcgg cgcgtgtcgg gcgtgttcca cgcgtcgggc 6540
 gtgctgcgcg acaagctcgt ggagaacaag tcgctggcgg acttcagcgc cgtgtacgac 6600
 accaagggtg gcggcctcat caacctgctg gcctgcgtgg acctggcgca gctgcgtcac 6660
 ctctgtctct tcagctcgtc cgcgggcttc caccgcaacg tcgggcagtc ggactacgca 6720
 atggccaacg aggcgtcaa caagctggcg gcgcacctgt cggcggtgca cccgcagctg 6780
 tcgcgcgct cgatctgctt cggaccgtgg gacggcgcca tggtagcccc cgcgtcaa 6840

ES 2 623 610 T3

gccaacttca	tccgcatggg	catccagatc	atcccgcgcc	aaggcggcgc	gcagaccgtc	6900
gccaacatgc	tcgtcagtag	ctcccccggt	cagctgctcg	tgggcaactg	gggcgtgcc	6960
cccgtcgtgc	cgagtgccac	cgagcacacc	gtgctgcaga	cgctccgcc	gagcgacaac	7020
cccttcctcg	actcgcaagt	gatccagggc	cgccgcgtgc	tgcccatgac	cctggccgtg	7080
ggctacatgg	cgcaccaggc	gcagagcatc	tacgcgggcc	accagctgtg	ggccgtcgag	7140
gacgcccagc	tcttcaaggg	catcgccatc	gacaatggcg	ccgacgtgcc	cgtgcgcgtg	7200
gagctgtcgc	gccgcaagga	ggagcaggag	gacgccggca	aggtcaaggt	caaggtgcag	7260
gtgctgtcga	aatcgcaagt	caacggcaag	tcggtgcccc	cgtacaaggc	gaccgtcgtg	7320
ctgtcccctg	cgccgcgcgc	cagcgatcatc	acgcgtgact	tcgacctcac	cccggaccgc	7380
gcctgcacgg	agcacgacct	ctacgacggc	aagacgctct	tccacggcaa	ggccttcacg	7440
ggcatcgagc	aggtgctctc	ggcgacgccc	aagcagctca	ccgccaagtg	ccgcaatttg	7500
ccctcacgc	ccgagcagcg	cgccagttc	gtcgttaacc	tcagccagca	ggaccggttc	7560
caggcggaca	ttgcgttcca	ggcgatgctc	gtctgggcgc	gcagctgtcg	ccaatcggcg	7620
gcctgcccc	acaactgcga	gcgcttcgac	ttttacaagc	cgatggcccc	gggcgcacc	7680
tactacacgt	cgggtcaagct	ggcctcgccc	tcacccttgg	tggactctgt	gtgcaagtgc	7740
accgtggcga	tgcacgatga	gcaaggtgag	gtgtactttt	ctgctcgtgc	cagcgtcgtc	7800
ctcaacaaga	ccctcacgta	ctaa				7824

<210> 2

<211> 2607

5 <212> PRT

<213> *Schizochytrium* sp.

<400> 2

ES 2 623 610 T3

Met Asp Thr Arg Ile Ala Ile Val Gly Met Ser Ala Ile Leu Pro Ser
1 5 10 15

Gly Glu Asn Val Arg Glu Ser Trp Glu Ala Ile Arg Asp Gly Leu Asp
20 25 30

Cys Leu Ser Asp Leu Pro Ala Asp Arg Val Asp Val Thr Ala Tyr Tyr
35 40 45

Asn Pro Glu Lys Thr Thr Lys Asp Lys Ile Tyr Cys Lys Arg Gly Gly
50 55 60

Phe Ile Pro Glu Tyr Asp Phe Asp Ala Arg Glu Phe Gly Leu Asn Met
65 70 75 80

Phe Gln Met Glu Asp Ser Asp Ala Asn Gln Thr Ile Ser Leu Leu Lys
85 90 95

Val	Lys	Glu	Ala	Leu	Thr	Asp	Ala	Asn	Ile	Pro	Ala	Phe	Ser	Ser	Gly	100	105	110
Lys	Lys	Asn	Ile	Gly	Cys	Val	Leu	Gly	Ile	Gly	Gly	Gly	Gln	Lys	Ala	115	120	125
Ser	His	Glu	Phe	Tyr	Ser	Arg	Leu	Asn	Tyr	Val	Val	Val	Asp	Lys	Val	130	135	140
Leu	Arg	Lys	Met	Gly	Leu	Pro	Glu	Glu	Asp	Val	Ala	Ala	Ala	Val	Asp	145	150	155
Lys	Tyr	Lys	Ala	Ser	Phe	Pro	Glu	Trp	Arg	Leu	Asp	Ser	Phe	Pro	Gly	165	170	175
Phe	Leu	Gly	Asn	Val	Thr	Ala	Gly	Arg	Cys	Cys	Asn	Thr	Phe	Asn	Met	180	185	190
Glu	Gly	Met	Asn	Cys	Val	Val	Asp	Ala	Ala	Cys	Ala	Ser	Ser	Leu	Ile	195	200	205
Ala	Val	Lys	Val	Ala	Ile	Glu	Glu	Leu	Leu	Tyr	Gly	Asp	Cys	Asp	Ala	210	215	220
Met	Ile	Ala	Gly	Ala	Thr	Cys	Thr	Asp	Asn	Ser	Ile	Gly	Met	Tyr	Met	225	230	235
Ala	Phe	Ser	Lys	Thr	Pro	Val	Phe	Ser	Thr	Asp	Pro	Ser	Val	Lys	Ala	245	250	255
Tyr	Asp	Ala	Ala	Thr	Lys	Gly	Met	Leu	Ile	Gly	Glu	Gly	Ser	Ala	Met	260	265	270
Leu	Val	Leu	Lys	Arg	Tyr	Ala	Asp	Ala	Val	Arg	Asp	Gly	Asp	Thr	Val	275	280	285
His	Ala	Val	Ile	Lys	Gly	Cys	Ala	Ser	Ser	Ser	Asp	Gly	Lys	Ala	Ala	290	295	300
Gly	Ile	Tyr	Thr	Pro	Thr	Ile	Ser	Gly	Gln	Glu	Glu	Ala	Leu	Arg	Arg	305	310	315
Ala	Tyr	Ala	Arg	Ala	Asn	Val	Asp	Pro	Ala	Thr	Val	Thr	Leu	Val	Glu	325	330	335
Gly	His	Gly	Thr	Gly	Thr	Pro	Val	Gly	Asp	Lys	Ile	Glu	Leu	Thr	Ala	340	345	350

Leu Ser Asn Leu Phe Ser Lys Ala Phe Ser Ala Asn Gly Gly Gly Ala
 355 360 365
 Glu Glu Ala Glu Gln Val Ala Val Gly Ser Ile Lys Ser Gln Ile Gly
 370 375 380
 His Leu Lys Ala Val Ala Gly Leu Ala Gly Leu Val Lys Val Val Leu
 385 390 395 400
 Ala Leu Lys His Lys Thr Leu Pro Gln Thr Ile Asn Val Asp Lys Pro
 405 410 415
 Pro Ser Leu Val Asp Gly Thr Pro Ile Gln Gln Ser Pro Leu Tyr Val
 420 425 430
 Asn Thr Met Asn Arg Pro Trp Phe Thr Pro Val Gly Val Pro Arg Arg
 435 440 445
 Ala Gly Val Ser Ser Phe Gly Phe Gly Gly Ala Asn Tyr His Ala Val
 450 455 460
 Leu Glu Glu Phe Glu Pro Glu His Glu Ser Ala Tyr Arg Tyr Asn Asn
 465 470 475 480
 Leu Pro Gln Val Ala Leu Leu His Ala Gly Asp Val Ala Thr Leu Ala
 485 490 495
 Ala Thr Val Arg Ala Lys Leu Ala Leu Ala Thr Ala Glu Gln Glu Glu
 500 505 510
 Ala Arg Val Val Lys Asn Ala Asp Tyr Ile Ala Tyr His Arg Phe Leu
 515 520 525
 Asp Glu Cys Lys Leu Arg Gly Ala Val Pro Gln Ala His Ala Arg Val
 530 535 540
 Gly Leu Leu Val Arg Asp Leu Ser Ser Leu Ile Ala Val Leu Glu Ala
 545 550 555 560
 Ala Ala Ala Lys Leu Ala Gly Glu Glu Ser Ala Thr Glu Trp Thr Val
 565 570 575
 Ser Val Ala Thr Gly Glu Ala Ala Phe Arg Val Arg Gly Val Ala Thr
 580 585 590
 Glu Ala Asn Val Ala Ala Leu Phe Ser Gly Gln Gly Ala Gln Tyr Thr
 595 600 605

ES 2 623 610 T3

His Met Phe Ser Asp Val Ala Met Asn Trp Pro Pro Phe Arg Glu Ser
 610 615 620
 Val Ala Ala Met Asp Arg Ala Gln Arg Glu Arg Phe Gly Arg Pro Ala
 625 630 635 640
 Lys Arg Val Ser Ser Val Leu Tyr Pro Arg Lys Pro Tyr Gly Asp Glu
 645 650 655
 Pro Arg Gln Asp His Lys Glu Ile Ser Gln Thr Arg Tyr Ser Gln Pro
 660 665 670
 Ala Thr Leu Ala Cys Ser Val Gly Ala Phe Asp Ile Phe Lys Ala Ala
 675 680 685
 Gly Leu Ala Pro Ser Phe Ala Ala Gly His Ser Leu Gly Glu Phe Ala
 690 695 700
 Ala Leu Tyr Ala Ala Gly Ser Leu Asp Arg Asp Ala Val Phe Asp Leu
 705 710 715 720
 Val Cys Ala Arg Ala Lys Ala Met Ser Asp Phe Thr Ala Gln Ala Ser
 725 730 735
 Ser Ser Gly Gly Ala Met Ala Ala Val Ile Gly Ala Lys Ala Asp Gln
 740 745 750
 Leu Ser Leu Gly Gly Ala Pro Asp Val Trp Leu Ala Asn Ser Asn Ser
 755 760 765
 Pro Ser Gln Thr Val Ile Thr Gly Thr Ala Glu Ala Val Ala Ala Ala
 770 775 780
 Ser Asp Lys Leu Arg Cys Ser Gly Asn Phe Arg Val Val Pro Leu Ala
 785 790 795 800
 Cys Glu Ala Ala Phe His Ser Pro His Met Arg Gly Ala Glu Gln Thr
 805 810 815
 Phe Ala Ser Ala Leu Ala Gln Ala Pro Val Ser Ala Pro Ala Ala Ala
 820 825 830
 Arg Phe Tyr Ser Asn Val Thr Gly Gly Ala Ala Val Thr Ser Pro Ala
 835 840 845
 Asp Val Lys Thr Asn Leu Gly Lys His Met Thr Ser Pro Val Gln Phe
 850 855 860
 Val Gln Gln Val Arg Ala Met His Ala Ala Gly Ala Arg Val Phe Val

865		870		875		880
Glu Phe Gly Pro Lys Gln Val Leu Ser Arg Leu Val Lys Glu Thr Leu						
		885		890		895
Gly Glu Ala Gly Asp Val Val Thr Val Ala Val Asn Pro Asp Ser Ala						
		900		905		910
Lys Asp Ser Asp Thr Gln Leu Arg Gln Ala Ala Leu Thr Leu Ala Val						
		915		920		925
Ala Gly Val Pro Leu Lys Asp Phe Asp Arg Trp Gln Leu Pro Asp Ala						
		930		935		940
Thr Arg Leu Glu Pro Val Lys Lys Lys Lys Thr Thr Leu Arg Leu Ser						
		945		950		955
				955		960
Ala Ala Thr Tyr Val Ser Ala Lys Thr Leu Arg Gln Arg Glu Ala Val						
		965		970		975
Leu Asn Asp Gly Tyr Thr Val Ser Gly Ala Thr Ala Val Val Lys Glu						
		980		985		990
Val Asp Thr Ala Asn Glu Glu Arg Leu Val Arg Gln Ala Gln Asp Leu						
		995		1000		1005
Gln Arg Gln Leu Ala Glu Ala Ser Thr Ala Ala Gln Ala Ala Gln						
		1010		1015		1020
Ser Lys Val Ala Glu Leu Glu Arg Thr Ile Gln Asp Leu Glu Arg						
		1025		1030		1035
Lys Val Gln Gln Gln Gln Gln Glu Lys Gly Glu Asn Ser Asp Ser						
		1040		1045		1050
Asn Ala Ala Ala Glu Val Leu Arg Arg His Lys Glu Leu Leu Gln						
		1055		1060		1065
Arg Met Leu Gln Asp Cys Asp Glu Gln Ala Val Pro Val Ala Thr						
		1070		1075		1080
Val Val Pro Thr Pro Thr Ser Ser Pro Thr Pro Thr Ser Ser Pro						
		1085		1090		1095
Val Ser Gly Asn Ser Lys Ser Thr Arg Gly Ser Ala Asp Leu Gln						
		1100		1105		1110
Ala Leu Leu Ala Lys Ala Glu Thr Val Val Met Ala Val Leu Ala						
		1115		1120		1125

Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala Asp Met Asp 1130 1135 1140
Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile 1145 1150 1155
Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp Val 1160 1165 1170
Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala 1175 1180 1185
Met Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser Ala Pro Ala 1190 1195 1200
Val Pro Ser Ala Pro Ala Ala Ser Ala Ala Pro Thr Pro Ala Ala 1205 1210 1215
Ser Thr Ala Pro Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala 1220 1225 1230
Glu Thr Val Val Met Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu 1235 1240 1245
Ala Asp Met Val Glu Ala Asp Met Asp Leu Glu Ala Glu Leu Gly 1250 1255 1260
Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln Gly 1265 1270 1275
Gln Leu Gly Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr 1280 1285 1290
Arg Thr Val Gly Glu Val Val Asp Ala Met Lys Ala Glu Ile Val 1295 1300 1305
Ala Ala Ser Ala Gly Ser Ala Pro Ala Pro Ala Val Pro Ser Ala 1310 1315 1320
Pro Ala Ala Ser Ala Ala Pro Thr Pro Ala Ala Ser Thr Ala Pro 1325 1330 1335
Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val 1340 1345 1350
Met Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val 1355 1360 1365

Glu 1370	Ala	Asp	Met	Asp	Leu	Glu 1375	Ala	Glu	Leu	Gly	Ile 1380	Asp	Ser	Ile
Lys 1385	Arg	Val	Glu	Ile	Leu	Ser 1390	Glu	Val	Gln	Gly	Gln 1395	Leu	Gly	Val
Glu 1400	Ala	Lys	Asp	Val	Asp	Ala 1405	Leu	Ser	Arg	Thr	Arg 1410	Thr	Val	Gly
Glu 1415	Val	Val	Asp	Ala	Met	Lys 1420	Ala	Glu	Ile	Val	Ala 1425	Ala	Ser	Gly
Gly 1430	Ser	Ala	Pro	Ala	Pro	Ala 1435	Val	Pro	Ser	Ala	Pro 1440	Ala	Ala	Ser
Ala 1445	Ala	Pro	Thr	Pro	Ala	Ala 1450	Ala	Thr	Ala	Pro	Ser 1455	Ala	Asp	Leu
Gln 1460	Ala	Leu	Leu	Ala	Lys	Ala 1465	Glu	Thr	Val	Val	Met 1470	Ala	Val	Leu
Ala 1475	Ala	Lys	Thr	Gly	Tyr	Glu 1480	Ala	Asp	Met	Val	Glu 1485	Ala	Asp	Met
Asp 1490	Leu	Glu	Ala	Glu	Leu	Gly 1495	Ile	Asp	Ser	Ile	Lys 1500	Arg	Val	Glu
Ile 1505	Leu	Ser	Glu	Val	Gln	Gly 1510	Gln	Leu	Gly	Val	Glu 1515	Ala	Lys	Asp
Val 1520	Asp	Ala	Leu	Ser	Arg	Thr 1525	Arg	Thr	Val	Gly	Glu 1530	Val	Val	Asp
Ala 1535	Met	Lys	Ala	Glu	Ile	Val 1540	Ala	Ala	Ser	Ala	Gly 1545	Ser	Ala	Pro
Ala 1550	Pro	Ala	Val	Pro	Ser	Ala 1555	Pro	Ala	Ala	Ser	Ala 1560	Ala	Pro	Thr
Pro 1565	Ala	Ala	Ser	Thr	Ala	Pro 1570	Ser	Ala	Asp	Leu	Gln 1575	Ala	Leu	Leu
Ser 1580	Lys	Ala	Glu	Thr	Val	Val 1585	Met	Ala	Val	Leu	Ala 1590	Ala	Lys	Thr
Gly 1595	Tyr	Glu	Ala	Asp	Met	Val 1600	Glu	Ala	Asp	Met	Asp 1605	Leu	Glu	Ala

Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu
1610						1615					1620			
Val	Gln	Gly	Gln	Leu	Gly	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu
1625						1630					1635			
Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Val	Asp	Ala	Met	Lys	Ala
1640						1645					1650			
Glu	Ile	Val	Ala	Ala	Ser	Gly	Gly	Ser	Ala	Pro	Ala	Ala	Ala	Val
1655						1660					1665			
Pro	Ser	Ala	Pro	Ala	Ala	Ser	Ala	Ala	Pro	Thr	Pro	Ala	Thr	Ala
1670						1675					1680			
Pro	Ser	Ala	Asp	Leu	Gln	Ala	Leu	Leu	Ser	Lys	Ala	Glu	Thr	Val
1685						1690					1695			
Val	Met	Ala	Val	Leu	Ala	Ala	Lys	Thr	Gly	Tyr	Glu	Ala	Asp	Met
1700						1705					1710			
Val	Glu	Ala	Asp	Met	Asp	Leu	Glu	Ala	Glu	Leu	Gly	Ile	Asp	Ser
1715						1720					1725			
Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Gly	Gln	Leu	Gly
1730						1735					1740			
Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val
1745						1750					1755			
Gly	Glu	Val	Val	Asp	Ala	Met	Lys	Ala	Glu	Ile	Val	Ala	Ala	Ser
1760						1765					1770			
Gly	Gly	Ser	Ala	Pro	Ala	Ala	Pro	Ser	Ala	Pro	Ala	Leu	Leu	Pro
1775						1780					1785			
Thr	Leu	Phe	Gly	Ser	Glu	Cys	Glu	Asp	Leu	Ser	Leu	Thr	Phe	Pro
1790						1795					1800			
Val	Ile	Thr	Thr	Leu	Pro	Leu	Pro	Ala	Glu	Leu	Val	Leu	Ala	Glu
1805						1810					1815			
Gly	Gly	Ala	Arg	Pro	Val	Val	Val	Val	Asp	Asp	Gly	Ser	Ala	Leu
1820						1825					1830			
Thr	Ser	Ser	Leu	Val	Ser	Ser	Leu	Gly	Asp	Arg	Ala	Val	Leu	Leu
1835						1840					1845			
Gln	Val	Gln	Ser	Ser	Ser	Ala	Cys	Ser	Pro	Arg	Ser	Thr	Thr	His

ES 2 623 610 T3

1850	1855	1860
Lys Leu Val Thr Val Ala Asp 1865	Arg Ser Glu Ala Ala 1870	Leu Gln Ala 1875
Ala Leu Thr Ser Val Glu Ala 1880	Gln Phe Gly Lys Val 1885	Gly Gly Phe 1890
Val Phe Gln Phe Gly Asp 1895	Asp Val Gln Ala Gln 1900	Leu Gly Trp 1905
Ala Leu Leu Ala Ala Lys 1910	His Leu Lys Thr Ser 1915	Leu Ser Glu Gln 1920
Ile Glu Gly Gly Arg Thr 1925	Phe Phe Val Ala Val 1930	Arg Leu Asp 1935
Gly Gln Leu Gly Leu Ser 1940	Gly Lys Ser Thr Thr 1945	Ala Thr Val Asp 1950
Leu Ser Arg Ala Gln Gln 1955	Gly Ser Val Phe Gly 1960	Leu Cys Lys Thr 1965
Leu Asp Leu Glu Trp Pro 1970	Ala Val Phe Cys Arg 1975	Gly Ile Asp Leu 1980
Ala Ala Asp Leu Asp Ala 1985	Ala Gln Ala Ala Arg 1990	Cys Leu Leu Gly 1995
Glu Leu Ser Asp Pro Asp 2000	Val Ala Val Arg Glu 2005	Ser Gly Tyr Ser 2010
Ala Ser Gly Gln Arg Cys 2015	Thr Thr Thr Thr Lys 2020	Ser Leu Thr Thr 2025
Gly Lys Pro His Gln Pro 2030	Ile Ser Ser Ser Asp 2035	Leu Phe Leu Val 2040
Ser Gly Gly Ala Arg Gly 2045	Ile Thr Pro Leu Cys 2050	Val Arg Glu Leu 2055
Ala Gln Arg Val Gly Gly 2060	Gly Thr Tyr Val Leu 2065	Ile Gly Arg Ser 2070
Glu Leu Pro Thr Thr Glu 2075	Pro Ala Trp Ala Val 2080	Gly Val Glu Ser 2085
Gly Lys Pro Leu Glu Lys 2090	Ala Ala Leu Ala Phe 2095	Leu Lys Ala Glu 2100

ES 2 623 610 T3

Phe	Ala	Ala	Gly	Arg	Gly	Ala	Lys	Pro	Thr	Pro	Met	Leu	His	Lys
2105						2110					2115			
Lys	Leu	Val	Gly	Ala	Val	Val	Gly	Ala	Arg	Glu	Val	Arg	Ala	Ser
2120						2125					2130			
Leu	Ala	Glu	Ile	Thr	Ala	Gln	Gly	Ala	Thr	Ala	Val	Tyr	Glu	Ser
2135						2140					2145			
Cys	Asp	Val	Ser	Ser	Ala	Ala	Lys	Val	Arg	Glu	Met	Val	Glu	Arg
2150						2155					2160			
Val	Gln	Gln	Gln	Gly	Gly	Arg	Arg	Val	Ser	Gly	Val	Phe	His	Ala
2165						2170					2175			
Ser	Gly	Val	Leu	Arg	Asp	Lys	Leu	Val	Glu	Asn	Lys	Ser	Leu	Ala
2180						2185					2190			
Asp	Phe	Ser	Ala	Val	Tyr	Asp	Thr	Lys	Val	Gly	Gly	Leu	Ile	Asn
2195						2200					2205			
Leu	Leu	Ala	Cys	Val	Asp	Leu	Ala	Gln	Leu	Arg	His	Leu	Val	Leu
2210						2215					2220			
Phe	Ser	Ser	Leu	Ala	Gly	Phe	His	Gly	Asn	Val	Gly	Gln	Ser	Asp
2225						2230					2235			
Tyr	Ala	Met	Ala	Asn	Glu	Ala	Leu	Asn	Lys	Leu	Ala	Ala	His	Leu
2240						2245					2250			
Ser	Ala	Val	His	Pro	Gln	Leu	Cys	Ala	Arg	Ser	Ile	Cys	Phe	Gly
2255						2260					2265			
Pro	Trp	Asp	Gly	Gly	Met	Val	Thr	Pro	Ala	Leu	Lys	Ala	Asn	Phe
2270						2275					2280			
Ile	Arg	Met	Gly	Ile	Gln	Ile	Ile	Pro	Arg	Gln	Gly	Gly	Ala	Gln
2285						2290					2295			
Thr	Val	Ala	Asn	Met	Leu	Val	Ser	Ser	Ser	Pro	Gly	Gln	Leu	Leu
2300						2305					2310			
Val	Gly	Asn	Trp	Gly	Val	Pro	Pro	Val	Val	Pro	Ser	Ala	Thr	Glu
2315						2320					2325			
His	Thr	Val	Leu	Gln	Thr	Leu	Arg	Gln	Ser	Asp	Asn	Pro	Phe	Leu
2330						2335					2340			

Asp	Ser	His	Val	Ile	Gln	Gly	Arg	Arg	Val	Leu	Pro	Met	Thr	Leu
2345						2350					2355			
Ala	Val	Gly	Tyr	Met	Ala	His	Gln	Ala	Gln	Ser	Ile	Tyr	Ala	Gly
2360						2365					2370			
His	Gln	Leu	Trp	Ala	Val	Glu	Asp	Ala	Gln	Leu	Phe	Lys	Gly	Ile
2375						2380					2385			
Ala	Ile	Asp	Asn	Gly	Ala	Asp	Val	Pro	Val	Arg	Val	Glu	Leu	Ser
2390						2395					2400			
Arg	Arg	Lys	Glu	Glu	Gln	Glu	Asp	Ala	Gly	Lys	Val	Lys	Val	Lys
2405						2410					2415			
Val	Gln	Val	Leu	Leu	Lys	Ser	Gln	Val	Asn	Gly	Lys	Ser	Val	Pro
2420						2425					2430			
Ala	Tyr	Lys	Ala	Thr	Val	Val	Leu	Ser	Pro	Ala	Pro	Arg	Pro	Ser
2435						2440					2445			
Val	Ile	Thr	Arg	Asp	Phe	Asp	Leu	Thr	Pro	Asp	Pro	Ala	Cys	Thr
2450						2455					2460			
Glu	His	Asp	Leu	Tyr	Asp	Gly	Lys	Thr	Leu	Phe	His	Gly	Lys	Ala
2465						2470					2475			
Phe	Gln	Gly	Ile	Glu	Gln	Val	Leu	Ser	Ala	Thr	Pro	Lys	Gln	Leu
2480						2485					2490			
Thr	Ala	Lys	Cys	Arg	Asn	Leu	Pro	Leu	Thr	Pro	Glu	Gln	Arg	Gly
2495						2500					2505			
Gln	Phe	Val	Val	Asn	Leu	Ser	Gln	Gln	Asp	Pro	Phe	Gln	Ala	Asp
2510						2515					2520			
Ile	Ala	Phe	Gln	Ala	Met	Leu	Val	Trp	Ala	Arg	Met	Leu	Arg	Gln
2525						2530					2535			
Ser	Ala	Ala	Leu	Pro	Asn	Asn	Cys	Glu	Arg	Phe	Asp	Phe	Tyr	Lys
2540						2545					2550			
Pro	Met	Ala	Pro	Gly	Ala	Thr	Tyr	Tyr	Thr	Ser	Val	Lys	Leu	Ala
2555						2560					2565			
Ser	Ala	Ser	Pro	Leu	Val	Asp	Ser	Val	Cys	Lys	Cys	Thr	Val	Ala
2570						2575					2580			

ES 2 623 610 T3

Met His Asp Glu Gln Gly Glu Val Tyr Phe Ser Ala Arg Ala Ser
 2585 2590 2595

Val Val Leu Asn Lys Thr Leu Thr Tyr
 2600 2605

<210> 3

<211> 5778

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 3

atgccgtgcg	ataacattgc	ggtcgtgggc	atggcggtgc	agtatgccgg	atgcaagaac	60
caggacgagt	tctgggatac	gctgatgcgt	aaggagatca	actcgagccc	gatctcggcg	120
gagcgcctcg	gtacgcgcta	cgcgcacctc	cacttcacac	cgcagcgcag	caagtacgcc	180
gacaccttct	gcaacgatcg	ctacggctgc	gtcgatgcca	gcgtcgacaa	cgagcacgac	240
ctcctcgccg	acctggcccc	gcgcgccttg	ctcgacgcgc	gaattaacct	cgacgacgcc	300
agcaccaccg	ccaacctacg	cgacttcggc	atcgtgagcg	gctgcctgtc	gttccccatg	360
gacaatctgc	agggcgagct	gctcaatctg	taccaagtgc	atgtggagaa	ccgcgtgggc	420
gcccagcgct	tccgcgactc	gcgcccctgg	tcggagcgcc	cgcgcgctgt	ctcgcccagag	480
gccagcgacc	cgcgcgtgta	ctccgacccg	gcgtccttcg	tggccaacca	gctcggcctg	540
gggcccgtgc	gctacagcct	cgatgcagcc	tgcgcgtcgg	cgctgtactg	cctcaagctg	600
gcgtccgacc	acttgctctc	gcgcagcgcg	gacgtgatgc	tgtgcggcgc	cacatgcttt	660
ccggacccgt	tcttcattct	ctcggggttc	tccaccttc	aggcgatgcc	gctgggcgga	720
ccggacgata	accactgtc	cgtgccgctg	cggcagggca	gccagggcct	gacgcccgga	780
gagggcgggc	ccatcatggt	gctgaagcgc	ctcgaggacg	ccgtgcgcga	cggcgaccgc	840
atctacggca	ccttgctcgg	cacgagtctg	agcaacgcgc	ggtgcggcct	gcgctgagc	900
ccgcacctgc	cgagcgagaa	gtcgtgcatg	gaggacctgt	acacgagcgt	cggcacgcac	960
ccaagcgagg	tgcagtacgt	ggagtgcac	gccacgggca	ctccgcaggg	cgacgtcgtg	1020
gaggtagagg	cgtgcgcga	ctgctttcga	ggtaacacgg	accaccgcgc	gcgcatgggc	1080
tccaccaagg	gcaactttgg	ccacactctc	gtggcgggcg	ggttcgcagg	catggccaag	1140
gtgctgctgt	cgatgcagca	cggcacgac	ccgccacgc	ccggtgtcga	ccgtccaac	1200
tgcacgcacc	cgtcgtcgt	ggacgaggcc	atcccttggc	cgtactcgtc	ggcgcaggcg	1260
cgggcaggca	aaccaggcga	tgagctcaag	tgcgcctcgc	tctccgcctt	tggctttggt	1320
ggaaccaacg	cgcactgtgt	cttcgctgag	caccgccaaa	ttgctgctac	tgcgacagcc	1380
tcgccggtgc	ttcccagggt	gactcctgga	ccgattgcca	tcatcgggat	ggacgcgacg	1440
tttgggtacc	tcaagggcct	ggacgcgttt	gagcaggcca	tctacaaggg	cacggacggc	1500

gccagcgacc tgcgagcaa gcgctggcgg ttcttggggc cgcacacgga cttcttgacc 1560
gccatggggc tcgacgccgt gccgcgcggg tgctacgtgc gcgacgtgga cgtggactac 1620
aagcggctgc ggtcgccgat gatccctgag gacgtcctgc gcccgcaaca gctgctggcg 1680
gtggctacga tggaccgcgc gctgcaggac gctggaatgg cgcaggagg caaggtggcg 1740
gtgctggtgg ggctcgccac ggacaccgag ctgtaccggc accgcgcgcg cgtgacactc 1800
aaggagcggc tcgaccggc cgcgttctcg cccgagcagg tgcaggagat gatggactac 1860
atcaacgact gcggcacctc gacgtcgtac acgtcgtaca tcggcaacct cgtggccacg 1920
cgcgtgtcct cgcagtgggg ctttacgggc ccgtccttca ccgtcaccga aggcgcaaac 1980
tcggtctacc gctgcctcga gctgggcaag ttcttgcctc acacgcacca ggtggacgcc 2040
gtcgtggtgg ccggcgctga cctctgtgcc accgcgcaga acctttacct caaggcgcgc 2100
cgctccgcca tcagccgaca ggaccaccct cgcgccaaact ttgaggccag cgcgcaggg 2160
tactttgccg gcgagggcag cggcgccctg gtctcaagc gccaggccga cgttggtca 2220
gacgacaagg totacgccag tgcgcggggc ctcacgtgcg ccgcgcagcc cgtgaagcc 2280
gtgtcgccgc tactactcca agtccacaac gacgacaacg agaagagggt ggtggagatg 2340
gtggagctcg ccgccgactc gggtcgccat gcgccgcaact tggccaactc gccgctgagc 2400
gccgagtcgc agctggagca agtgtccaag ttgtctgcgc accaggtgcc gggtcggtg 2460
gccatcggca gcgtgcgcgc caacgtggga gacgtcgggt acgcctcggg cgcgcgagc 2520
ctcatcaaga cggcgctgtg cctccacaac cgtacctcc cggccaacct gcagtgggag 2580
cggccggtgg cgcgggtctc cgaggcgctg tttacttgcc cgcgctcgcg tgccctggctg 2640
aagaacccgg gcgagtcgcg actggcggtc gtgcgcagtg cctccgagag cgggtcctgc 2700
tttgcgctgc tcctcacaga cgagtaagcc actcatgaga gcagcaaccg cctctcgtg 2760
gatgacgccg cccccaagct catcgcgatc cgtggcgaca ccgttgacga tcatatggcc 2820
aaggtcaacg ccgagctggc gctcctccga gcgcacgccg aaaccgggtc tgctactgac 2880
gacgaccag ctgctgctgt cgttttcaact gctcatcgct tgcgcttttt gcggtcgtg 2940
ggggagacgg tggttagtca cggtgccacg gcgacctgt gtttgccct gctgacaacg 3000
ccggagaagc tggagaagga gttggagctg gcagccaagg gtgtaccgcg aagcgccaag 3060
gccggggcga actggatgtc gccatcgggc agcgcccttg cgcgcacacc tgtgaccagc 3120
gaccgcgtcg cgttcatgta cggcgagggc cgcagccct actacggcgt cgggctcgac 3180
ctgcaccgcc tgtggccggc tttgcaagag cgcatacaacg acaagaccgc ggcgctgtgg 3240
gagaacggcg actcgtggct catgcgcgc gcggtggatg ccgactcgca gcgcgcgtg 3300
cagacggcct ttgacgcgga ccagatcgag atgttccgca cgggcacctt cgtgtccatc 3360
tgctcaccg actacgcgcg cgacgtgctc ggggtgcagc ccaaggcgtg cttcggcctc 3420
agcctcggcg agatctccat gctctttgcg ctgtcgcgac gcaactgcg cctgtcggac 3480

cagctcacgc agcgccctacg cacctcgccg gtgtgggtcga cacagctggc ggtggagttc 3540
 caggccttgc gcaagctatg gaacgtgccg gcggacgccc ccgtggagtc cttctggcag 3600
 ggctacttgg ttgcgcgcag ccgcgccgaa atcgagaagg cgatcgggcc cgacaaccgc 3660
 ttctgtgccc tgctgatcgt caacgactcg agcagcgccg tgatcgccgg caaacctgcc 3720
 gagtgtctgc gcgtgctgga gcgcctgggc gggcggttgc cggcgatgcc cgtcaagcaa 3780
 ggcatgattg ggcactgccc cgaagtggcg ccctacacgc cgggcatcgc gcacatccac 3840
 gagatttttg agattccgga cagccccgtc aagatgtaca cctcggtcac caacgccgag 3900
 ctgcgcgggg gcagcaacag cagcatcacc gaggctcgtc agaagttgta cagcgccatc 3960
 gccgactttc cgggcatcgt cgacaaggtc agccgtgacg gccacgatgt ctctgtcgag 4020
 gtggggccga acaacatgcg ctccgcccgcg gtcagtgaaca ttcttggaac ggetgccacc 4080
 ccgcatgtct ccgtggcgct ggaccgcccc agtgagtcgg cgtggacgca gacctcaag 4140
 tcgctggcgc tgctgaccgc ccaccgcgtg cccctgcaca acccgactct gtttgccgac 4200
 ctgtaccacc ccacgttccg gacggctatc gactctgcga tgcaggagcc cccgcccaag 4260
 cccaaccgct tccttcgcag cgtagaggtc aacgggtact tttgccccga cggcatcagc 4320
 aagcagggtt ctgctgcaag tgccaaaccc tcgacgcatt gcatggttcg tttgcaccca 4380
 gccaaaggcag ttgtggttgc tgctgctggt gctgtggttg ctgattcgac gcccggtggtc 4440
 aaggccaagc agacgtcgtc gtgcttgttg gttggggatg acgcctttct gcgctgctac 4500
 gacgtggact ggccgctcta catgggcgcc atggcggaag gcatctcgtc ggtagacctg 4560
 gtggtcgtcg ccgccgaggc ccgcatgctg gcatcattcg gaggggcccg cttgcctatg 4620
 gaccaggtgg aactccagat ccgtgagatc cagcaacgca cctccaacgc ctttgcctgc 4680
 aacctgatgc cgggtcctga cgaggccgcg acgggtggacg cgctgctgcg caccggcgctc 4740
 tcaatcgctg aggcacgagg ctacaccggc gcgctctctg cagacctggt gcgctaccgt 4800
 gtcacgggtc tgcgacgaac tagttgcggg gcttctgtgt cggcgactca ccgtgtggtc 4860
 gccaaagggt cgcgccaccga ggtggccgag cactttctgc gcccggcgcc ggccgcgcta 4920
 ctagaggctt tggtcgccgc caaacagatt acgcccagac aggcgcgcgt ggccagccgc 4980
 gtcgccatgg ccgacgacgt cgcggtggag gccgactcgg gcgggcacac cgacaaccga 5040
 ccgatccacg tgctgctgcc gctcgtggtg gcgcagcgca accgctggcg ccacctggtg 5100
 gacacgccag tgcgcgtcgg cgccggcggc gggatcgccg gtccgcgcgc cgcgctgctc 5160
 gccttttccc tgggcgcgcg ctttgtggtc accgggtccg tcaaccaact ggcccgcgag 5220
 gctggcacca gcgacgggt ccgaactact ctggcgacgg ccacctactc ggacgtggcc 5280
 atggcgccgg gcggcgctca ggtgctcaag aagcagacca tgctcgccgc gcggggccacg 5340
 atgctcggcc agctgcaggc caagttcggc tcctttgacg ccgtgccgga gccgcagctg 5400

cgcaagctcg agcgctcgt gttcaagcag tccgtggcgg acgtgtgggc tgctgcacgc 5460
gaaaagtttg gtgtcgacgc taccgctgca agtccgcagg agaggatggc gctctgtgtg 5520
cgctggtaca tgctgcagtc gtcgcgatgg gctaccgagg cgacgtccgc gcgcaaggcg 5580
gactaccaga tctgggtggg ccccgccatc ggcagcttca acgacttcgt tcgcggcacc 5640
aagctggacg cgaccgctgg caccggcgag ttccgcgcg tcgtggacat caaccagcac 5700
atcctcctcg gagcctcgca ctaccgcgc gtgcagcaac aacaacagga cgacgacgta 5760
gaatacatca tcgtataa 5778

<210> 4

<211> 1925

<212> PRT

<213> *Schizochytrium* sp.

<400> 4

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

ES 2 623 610 T3

Gln Leu Gly Leu Gly Pro Val Arg Tyr Ser Leu Asp Ala Ala Cys Ala
 180 185 190

Ser Ala Leu Tyr Cys Leu Lys Leu Ala Ser Asp His Leu Leu Ser Arg
 195 200 205

Ser Ala Asp Val Met Leu Cys Gly Ala Thr Cys Phe Pro Asp Pro Phe
 210 215 220

Phe Ile Leu Ser Gly Phe Ser Thr Phe Gln Ala Met Pro Leu Gly Gly
 225 230 235 240

Pro Asp Asp Asn Pro Leu Ser Val Pro Leu Arg Gln Gly Ser Gln Gly
 245 250 255

Leu Thr Pro Gly Glu Gly Gly Ala Ile Met Val Leu Lys Arg Leu Glu
 260 265 270

Asp Ala Val Arg Asp Gly Asp Arg Ile Tyr Gly Thr Leu Leu Gly Thr
 275 280 285

Ser Leu Ser Asn Ala Gly Cys Gly Leu Pro Leu Ser Pro His Leu Pro
 290 295 300

Ser Glu Lys Ser Cys Met Glu Asp Leu Tyr Thr Ser Val Gly Ile Asp
 305 310 315 320

Pro Ser Glu Val Gln Tyr Val Glu Cys His Ala Thr Gly Thr Pro Gln
 325 330 335

Gly Asp Val Val Glu Val Glu Ala Leu Arg His Cys Phe Arg Gly Asn
 340 345 350

Thr Asp His Pro Pro Arg Met Gly Ser Thr Lys Gly Asn Phe Gly His
 355 360 365

Thr Leu Val Ala Ala Gly Phe Ala Gly Met Ala Lys Val Leu Leu Ser
 370 375 380

Met Gln His Gly Thr Ile Pro Pro Thr Pro Gly Val Asp Arg Ser Asn
 385 390 395 400

Cys Ile Asp Pro Leu Val Val Asp Glu Ala Ile Pro Trp Pro Tyr Ser
 405 410 415

Ser Ala Gln Ala Arg Ala Gly Lys Pro Gly Asp Glu Leu Lys Cys Ala
 420 425 430

Ser Leu Ser Ala Phe Gly Phe Gly Gly Thr Asn Ala His Cys Val Phe
435 440 445

Arg Glu His Arg Gln Ile Ala Ala Thr Ala Thr Ala Ser Pro Val Leu
450 455 460

Pro Glu Val Thr Pro Gly Pro Ile Ala Ile Ile Gly Met Asp Ala Thr
465 470 475 480

Phe Gly Thr Leu Lys Gly Leu Asp Ala Phe Glu Gln Ala Ile Tyr Lys
485 490 495

Gly Thr Asp Gly Ala Ser Asp Leu Pro Ser Lys Arg Trp Arg Phe Leu
500 505 510

Gly Ala Asp Thr Asp Phe Leu Thr Ala Met Gly Leu Asp Ala Val Pro
515 520 525

Arg Gly Cys Tyr Val Arg Asp Val Asp Val Asp Tyr Lys Arg Leu Arg
530 535 540

Ser Pro Met Ile Pro Glu Asp Val Leu Arg Pro Gln Gln Leu Leu Ala
545 550 555 560

Val Ala Thr Met Asp Arg Ala Leu Gln Asp Ala Gly Met Ala Thr Gly
565 570 575

Gly Lys Val Ala Val Leu Val Gly Leu Gly Thr Asp Thr Glu Leu Tyr
580 585 590

Arg His Arg Ala Arg Val Thr Leu Lys Glu Arg Leu Asp Pro Ala Ala
595 600 605

Phe Ser Pro Glu Gln Val Gln Glu Met Met Asp Tyr Ile Asn Asp Cys
610 615 620

Gly Thr Ser Thr Ser Tyr Thr Ser Tyr Ile Gly Asn Leu Val Ala Thr
625 630 635 640

Arg Val Ser Ser Gln Trp Gly Phe Thr Gly Pro Ser Phe Thr Val Thr
645 650 655

Glu Gly Ala Asn Ser Val Tyr Arg Cys Leu Glu Leu Gly Lys Phe Leu
660 665 670

Leu Asp Thr His Gln Val Asp Ala Val Val Val Ala Gly Val Asp Leu
675 680 685

ES 2 623 610 T3

Cys Ala Thr Ala Glu Asn Leu Tyr Leu Lys Ala Arg Arg Ser Ala Ile
 690 695 700
 Ser Arg Gln Asp His Pro Arg Ala Asn Phe Glu Ala Ser Ala Asp Gly
 705 710 715 720
 Tyr Phe Ala Gly Glu Gly Ser Gly Ala Leu Val Leu Lys Arg Gln Ala
 725 730 735
 Asp Val Gly Ser Asp Asp Lys Val Tyr Ala Ser Val Ala Gly Leu Thr
 740 745 750
 Cys Ala Ala Gln Pro Ala Glu Ala Val Ser Pro Leu Leu Leu Gln Val
 755 760 765
 His Asn Asp Asp Asn Glu Lys Arg Val Val Glu Met Val Glu Leu Ala
 770 775 780
 Ala Asp Ser Gly Arg His Ala Pro His Leu Ala Asn Ser Pro Leu Ser
 785 790 795 800
 Ala Glu Ser Gln Leu Glu Gln Val Ser Lys Leu Leu Ala His Gln Val
 805 810 815
 Pro Gly Ser Val Ala Ile Gly Ser Val Arg Ala Asn Val Gly Asp Val
 820 825 830
 Gly Tyr Ala Ser Gly Ala Ala Ser Leu Ile Lys Thr Ala Leu Cys Leu
 835 840 845
 His Asn Arg Tyr Leu Pro Ala Asn Pro Gln Trp Glu Arg Pro Val Ala
 850 855 860
 Pro Val Ser Glu Ala Leu Phe Thr Cys Pro Arg Ser Arg Ala Trp Leu
 865 870 875 880
 Lys Asn Pro Gly Glu Ser Arg Leu Ala Ala Val Ala Ser Ala Ser Glu
 885 890 895
 Ser Gly Ser Cys Phe Gly Val Leu Leu Thr Asp Glu Tyr Ala Thr His
 900 905 910
 Glu Ser Ser Asn Arg Leu Ser Leu Asp Asp Ala Ala Pro Lys Leu Ile
 915 920 925
 Ala Ile Arg Gly Asp Thr Val Asp Asp Ile Met Ala Lys Val Asn Ala
 930 935 940
 Glu Leu Ala Leu Leu Arg Ala His Ala Glu Thr Gly Ser Ala Thr Asp

ES 2 623 610 T3

945		950		955		960
Asp Asp Pro Ala Ala Ala Val Ala Phe Thr Ala His Arg Leu Arg Phe						
		965		970		975
Leu Arg Leu Val Gly Glu Thr Val Ala Ser His Gly Ala Thr Ala Thr						
		980		985		990
Leu Cys Leu Ala Leu Leu Thr Thr Pro Glu Lys Leu Glu Lys Glu Leu						
		995		1000		1005
Glu Leu Ala Ala Lys Gly Val Pro Arg Ser Ala Lys Ala Gly Arg						
		1010		1015		1020
Asn Trp Met Ser Pro Ser Gly Ser Ala Phe Ala Pro Thr Pro Val						
		1025		1030		1035
Thr Ser Asp Arg Val Ala Phe Met Tyr Gly Glu Gly Arg Ser Pro						
		1040		1045		1050
Tyr Tyr Gly Val Gly Leu Asp Leu His Arg Leu Trp Pro Ala Leu						
		1055		1060		1065
His Glu Arg Ile Asn Asp Lys Thr Ala Ala Leu Trp Glu Asn Gly						
		1070		1075		1080
Asp Ser Trp Leu Met Pro Arg Ala Val Asp Ala Asp Ser Gln Arg						
		1085		1090		1095
Ala Val Gln Thr Ala Phe Asp Ala Asp Gln Ile Glu Met Phe Arg						
		1100		1105		1110
Thr Gly Ile Phe Val Ser Ile Cys Leu Thr Asp Tyr Ala Arg Asp						
		1115		1120		1125
Val Leu Gly Val Gln Pro Lys Ala Cys Phe Gly Leu Ser Leu Gly						
		1130		1135		1140
Glu Ile Ser Met Leu Phe Ala Leu Ser Arg Arg Asn Cys Gly Leu						
		1145		1150		1155
Ser Asp Gln Leu Thr Gln Arg Leu Arg Thr Ser Pro Val Trp Ser						
		1160		1165		1170
Thr Gln Leu Ala Val Glu Phe Gln Ala Leu Arg Lys Leu Trp Asn						
		1175		1180		1185
Val Pro Ala Asp Ala Pro Val Glu Ser Phe Trp Gln Gly Tyr Leu						
		1190		1195		1200

Val	Arg	Ala	Ser	Arg	Ala	Glu	Ile	Glu	Lys	Ala	Ile	Gly	Pro	Asp
1205						1210					1215			
Asn	Arg	Phe	Val	Arg	Leu	Leu	Ile	Val	Asn	Asp	Ser	Ser	Ser	Ala
1220						1225					1230			
Leu	Ile	Ala	Gly	Lys	Pro	Ala	Glu	Cys	Leu	Arg	Val	Leu	Glu	Arg
1235						1240					1245			
Leu	Gly	Gly	Arg	Leu	Pro	Pro	Met	Pro	Val	Lys	Gln	Gly	Met	Ile
1250						1255					1260			
Gly	His	Cys	Pro	Glu	Val	Ala	Pro	Tyr	Thr	Pro	Gly	Ile	Ala	His
1265						1270					1275			
Ile	His	Glu	Ile	Leu	Glu	Ile	Pro	Asp	Ser	Pro	Val	Lys	Met	Tyr
1280						1285					1290			
Thr	Ser	Val	Thr	Asn	Ala	Glu	Leu	Arg	Gly	Gly	Ser	Asn	Ser	Ser
1295						1300					1305			
Ile	Thr	Glu	Phe	Val	Gln	Lys	Leu	Tyr	Thr	Arg	Ile	Ala	Asp	Phe
1310						1315					1320			
Pro	Gly	Ile	Val	Asp	Lys	Val	Ser	Arg	Asp	Gly	His	Asp	Val	Phe
1325						1330					1335			
Val	Glu	Val	Gly	Pro	Asn	Asn	Met	Arg	Ser	Ala	Ala	Val	Ser	Asp
1340						1345					1350			
Ile	Leu	Gly	Lys	Ala	Ala	Thr	Pro	His	Val	Ser	Val	Ala	Leu	Asp
1355						1360					1365			
Arg	Pro	Ser	Glu	Ser	Ala	Trp	Thr	Gln	Thr	Leu	Lys	Ser	Leu	Ala
1370						1375					1380			
Leu	Leu	Thr	Ala	His	Arg	Val	Pro	Leu	His	Asn	Pro	Thr	Leu	Phe
1385						1390					1395			
Ala	Asp	Leu	Tyr	His	Pro	Thr	Phe	Leu	Thr	Ala	Ile	Asp	Ser	Ala
1400						1405					1410			
Met	Gln	Glu	Pro	Pro	Pro	Lys	Pro	Asn	Arg	Phe	Leu	Arg	Ser	Val
1415						1420					1425			
Glu	Val	Asn	Gly	Tyr	Phe	Cys	Pro	Asp	Gly	Ile	Ser	Lys	Gln	Val
1430						1435					1440			

ES 2 623 610 T3

Ala	Ala	Ala	Ser	Ala	Lys	Pro	Ser	Thr	His	Cys	Met	Val	Arg	Leu
1445						1450					1455			
His	Pro	Ala	Lys	Ala	Val	Val	Val	Ala	Ala	Ala	Gly	Ala	Val	Val
1460						1465					1470			
Ala	Asp	Ser	Thr	Pro	Val	Val	Lys	Ala	Lys	Gln	Thr	Ser	Ser	Ser
1475						1480					1485			
Leu	Leu	Val	Gly	Asp	Asp	Ala	Phe	Leu	Arg	Cys	Tyr	Asp	Val	Asp
1490						1495					1500			
Trp	Pro	Leu	Tyr	Met	Gly	Ala	Met	Ala	Glu	Gly	Ile	Ser	Ser	Val
1505						1510					1515			
Asp	Leu	Val	Val	Ala	Ala	Ala	Glu	Ala	Arg	Met	Leu	Ala	Ser	Phe
1520						1525					1530			
Gly	Ala	Ala	Arg	Leu	Pro	Met	Asp	Gln	Val	Glu	Leu	Gln	Ile	Arg
1535						1540					1545			
Glu	Ile	Gln	Gln	Arg	Thr	Ser	Asn	Ala	Phe	Ala	Val	Asn	Leu	Met
1550						1555					1560			
Pro	Gly	Pro	Asp	Glu	Ala	Ala	Thr	Val	Asp	Ala	Leu	Leu	Arg	Thr
1565						1570					1575			
Gly	Val	Ser	Ile	Val	Glu	Ala	Ser	Gly	Tyr	Thr	Gly	Ala	Leu	Ser
1580						1585					1590			
Ala	Asp	Leu	Val	Arg	Tyr	Arg	Val	Thr	Gly	Leu	Arg	Arg	Thr	Ser
1595						1600					1605			
Cys	Gly	Ala	Ser	Val	Ser	Ala	Thr	His	Arg	Val	Val	Ala	Lys	Val
1610						1615					1620			
Ser	Arg	Thr	Glu	Val	Ala	Glu	His	Phe	Leu	Arg	Pro	Ala	Pro	Ala
1625						1630					1635			
Ala	Val	Leu	Glu	Ala	Leu	Val	Ala	Ala	Lys	Gln	Ile	Thr	Pro	Glu
1640						1645					1650			
Gln	Ala	Ala	Leu	Ala	Ser	Arg	Val	Ala	Met	Ala	Asp	Asp	Val	Ala
1655						1660					1665			
Val	Glu	Ala	Asp	Ser	Gly	Gly	His	Thr	Asp	Asn	Arg	Pro	Ile	His
1670						1675					1680			

Val	Leu	Leu	Pro	Leu	Val	Val	Ala	Gln	Arg	Asn	Arg	Trp	Arg	His
1685						1690					1695			
Leu	Val	Asp	Thr	Pro	Val	Arg	Val	Gly	Ala	Gly	Gly	Gly	Ile	Ala
1700						1705					1710			
Cys	Pro	Arg	Ala	Ala	Leu	Leu	Ala	Phe	Ser	Leu	Gly	Ala	Ala	Phe
1715						1720					1725			
Val	Val	Thr	Gly	Ser	Val	Asn	Gln	Leu	Ala	Arg	Glu	Ala	Gly	Thr
1730						1735					1740			
Ser	Asp	Ala	Val	Arg	Leu	Leu	Leu	Ala	Thr	Ala	Thr	Tyr	Ser	Asp
1745						1750					1755			
Val	Ala	Met	Ala	Pro	Gly	Gly	Val	Gln	Val	Leu	Lys	Lys	Gln	Thr
1760						1765					1770			
Met	Phe	Ala	Ala	Arg	Ala	Thr	Met	Leu	Ala	Gln	Leu	Gln	Ala	Lys
1775						1780					1785			
Phe	Gly	Ser	Phe	Asp	Ala	Val	Pro	Glu	Pro	Gln	Leu	Arg	Lys	Leu
1790						1795					1800			
Glu	Arg	Ser	Val	Phe	Lys	Gln	Ser	Val	Ala	Asp	Val	Trp	Ala	Ala
1805						1810					1815			
Ala	Arg	Glu	Lys	Phe	Gly	Val	Asp	Ala	Thr	Ala	Ala	Ser	Pro	Gln
1820						1825					1830			
Glu	Arg	Met	Ala	Leu	Cys	Val	Arg	Trp	Tyr	Met	Ser	Gln	Ser	Ser
1835						1840					1845			
Arg	Trp	Ala	Thr	Glu	Ala	Thr	Ser	Ala	Arg	Lys	Ala	Asp	Tyr	Gln
1850						1855					1860			
Ile	Trp	Cys	Gly	Pro	Ala	Ile	Gly	Ser	Phe	Asn	Asp	Phe	Val	Arg
1865						1870					1875			
Gly	Thr	Lys	Leu	Asp	Ala	Thr	Ala	Gly	Thr	Gly	Glu	Phe	Pro	Arg
1880						1885					1890			
Val	Val	Asp	Ile	Asn	Gln	His	Ile	Leu	Leu	Gly	Ala	Ser	His	Tyr
1895						1900					1905			
Arg	Arg	Val	Gln	Gln	Gln	Gln	Gln	Asp	Asp	Asp	Val	Glu	Tyr	Ile
1910						1915					1920			
Ile	Val													

1925

5 <210> 5
 <211> 4308
 <212> ADN
 <213> *Schizochytrium* sp.

 <400> 5

atgacatcat	cgaagaagac	tcccggtgtgg	gagatgagca	aggaggagct	gctggacggc	60
aagacggtgg	tcttcgacta	caacgagctg	ctcgaattcg	ccgagggcga	cgtgggcca	120
gtgttcggac	ccgagttcga	catcatcgac	aagtaccggc	gtcgcgtgcg	gctgccggcg	180
cgcgagtacc	tgctcgtgtc	gcgcgtgacg	ctgatggacg	ccgaggtgaa	caacttccgc	240
gtcgggtcgc	gcatggtgac	cgagtacgac	gtgcccgta	acggggagct	gtcggagggc	300
ggggacgtgc	cgtgggcggg	gctggtggag	tggggcagt	gcgacctgat	gctcatctcg	360
tacatgggca	tgcacttcca	gtgcaagggc	gaccgcgtgt	accgcctgct	caacacatcg	420
ctcaccttct	tgggggtggc	gcacgagggc	gagacgctgg	tgtacgacat	ccgcgtcacg	480
gggttcgcca	agggcgcggg	cggggagatc	tcgatgttct	tcttcgagta	cgactgcttc	540
gtggacggcc	gcctgctgat	cgagatgcgc	gacgggtgcg	ccgggttctt	cacggacgcc	600
gagctggccg	ccggcaaggg	cgtgcttaag	accaagggcg	agctggcggc	gcgcgcgacg	660
atccagaagc	aggacatcgc	gccctttgcg	ccggcgccgt	gtcgcacaa	gacctcgctg	720
gacgcgcgcg	agatgcggct	gctcgtggac	cgccagtggg	cgcgcgtctt	cggcagcggc	780
atggcgggca	tgcactacaa	gttgtgcgct	cgcaagatgc	tcatgatcga	ccgcgtcacg	840
cacctcgacc	cgcgcggcgg	cgcgcacggc	ctcgggctgc	tgatcgggga	gaaggtgctg	900
gagcgcgacc	actggtactt	cccctgccac	tttgtgcgcg	acgaggtgat	ggccgggtcg	960
ctggtcagcg	acggctgctc	gcagctcctc	aaggtgtaca	tgctgtggct	cggcctgcac	1020
acgaccgtgg	gcgcgttcga	ctttcgtccc	gtgagcgggc	acgccaaaca	ggtgcggtgc	1080
cgcgggcaga	tctcaccgca	caagggcaag	ctcgtgtacg	tgatggagat	caaggaaatg	1140
ggctttgacg	cgaagacggg	cgatccgttt	gcgatcgcgg	acgtggacat	catcgacgtc	1200
aacttcgagg	agggacaggc	gtttgcggga	gtggaagacc	tgcacagcta	cggccagggc	1260
gacctcgcga	agaagatcgt	cgtcgacttc	aagggcacgc	cgtctccct	gcagaagcgg	1320
aaggagcagc	agaaggaaag	catgaccgtg	actacgacga	cgacgacgac	gagccgggtg	1380
attgcgccgc	ccagcgggtg	cctcaagggc	gacccgacgg	cgcgcgacgag	cgtgacgtgg	1440
caccgatgg	cggagggcaa	cggcggggcc	ggaccgacgc	cgtcgttctc	gcogtccgcg	1500
taccgcgccg	gggcgggtgtg	cttctcgccg	ttccccaaca	acccgcttga	caacgaccac	1560
acgccggggc	agatgccgtt	gacctggttc	aacatgtccg	aattcatgtg	cggcaaagtg	1620
tccaactgcc	tgggccccga	gtttgcgcgc	ttcgacgcga	gcaagacgag	ccgcagcccc	1680

gcctttgacc tggcgctcgt gacgcgggtg acgagcgtgg cggacatgga gcacggggccg 1740
 ttctacaacg tggacgtcaa cccggggccag ggcacgatgg tgggcgagtt cgactgtccc 1800
 gcggacgcgt ggtttcttcgg cgcttcgagc cgcgacgacc acatgccgta ctcgatcctg 1860
 atggagatcg cgctgcagac gtcggggcgtc ctcacctcgg tgctcaaggc gccgctgacg 1920
 atggacaagg acgacatcct cttccgcaac ctgcacgcag acgccgagct cgtgggcgac 1980
 gccatgccgg acgtgcgcgg caagacgatc cgcaacttca ccaagtgcac aggctacagc 2040
 atgctcggca agatgggcat ccaccgcttc acctttgagc tcagcgtcga cggcgccgtc 2100
 ttctacaagg gcagcacctc gtttggtcgg ttctgtcccc aggtcttcga gtcgcagacc 2160
 ggtctcgaca acggcaagcc gcgcctgcct tggtaaccgg agaacaacgt cgccgtcgac 2220
 acgctctccg cgcgcgcctc cgcttctctc gcgcaaggtc agctgcagct gcagcgacgc 2280
 gggctcgacg cgcagttcct ggacacaatc cacctggcgg gcagcggcgc cggcgtgcac 2340
 gggcagggct acgcgcacgg ggagaaggcc gtgaacaagc aagattggtt cttctcgtgc 2400
 cacttctggt tcgaccccggt gatgcccggt tccctgggca tcgagtcgat gttccagctc 2460
 gtcgaggcgt ggtgcgtgaa gcagggaactc gcggcgccgc acggcatcgc tcaccagtg 2520
 ttgcgcgacg cgcgcggggc cagcagctgg aagtaccggg ggcagctaac cccaagaac 2580
 gaccgcatgg acagcgaggt gcacatcaag tcggtggcgg ccttctctc ctgggtcgac 2640
 gtcgtcgcgg acgggttctt cttcgtcgac ggcctccggg tctactcggc agacaacctc 2700
 cggtccgca tccagaccgg cgcggggccac gttgaagagc aagaggttgc tgccaaggcc 2760
 acaaccaaga acagcagtat tgctgatgtg gacgtggcgg acctgcaagc gctcaagcag 2820
 gcgttgctga cgctggagcg accgctgcag ctggacgcgg ggagcgaggt gcccgcctgc 2880
 gcggtgagcg acctgggcga taggggcttc atggagacgt acggggtggt ggcgcgcgtg 2940
 tacagcgggg cgatggccaa gggcatcgcg tcggcggacc tggtgatcgc gatgggcccag 3000
 cgcaagatgc tggggtcgtt tggcgcgggc gggctcccga tgcacgtcgt gcgcgcgggg 3060
 attgagaaga tccaggcagc gctgccagcg gggccatacg cggccaacct gattcactcg 3120
 ccttttgacg ccaacctgga gaagggaac gtggacctct tcctggagaa gggcgtgcgc 3180
 gtcgtggagg cgtcggcctt catggagctc acgccccagg tggtgcgcta ccgcgcgacg 3240
 ggcctctctc gcgacgcgcg cggcggctcc gtgcgcacgg cccacaagat catcggaag 3300
 gtcagccgca ccgagctggc cgagatgttt atccggcccc cgcgcgaagc cattctcgac 3360
 aagcttggtg cgtccggcga gatcaccccc gagcaggcgg cgctggcgct cgaggtgccc 3420
 atggcggacg acatcgccgt cgaggccgat tcgggcgggc acaccgacaa ccgccccatc 3480
 cacgtcatcc tgccctcat cctcagcctg cgcaaccgcc tccagcgcga gctcaagtac 3540
 cctgcgcgac accgcgtgcg cgtcggcgcc gggggcggca tcgggtgccc gcaagcggct 3600

```

ctggggcgctt tccacatggg cgcgcggttt gtggtgacgg gcacgggtcaa ccagctgagc 3660
cggcaggccg ggacatgcga caatgtgcgg cggcagctgt cgcgcgcgac gtactcggac 3720
atcacgatgg cgcgggcggc ggacatgttc gagcagggcg tcgagctgca ggtgctcaag 3780
aagggcacga tgtttccctc gcgcgcgaag aagctgttcg agctgtttca caagtacgac 3840
tcgttcgagg cgatgccggc ggacgagctg ggcgcgctcg agaagcgcat cttcagcaag 3900
tcaactcgcc aggtgtgggc cgagaccaag gactttctaca tcacgcggct caacaaccgc 3960
gagaagatcc gcaaggcgga gaacgaggac cccaagctca agatgtcact ctgcttccgc 4020
tggtacctcg ggctcagctc gttctggggc aacaacggca tcggggaccg cactgatggc 4080
taccagatct ggtgcggccc tgccatcggc gccttcaacg acttcategc cgactcgtac 4140
ctcgacgtgg ccgtctcgga cgagttcccc gacgtcgtgc agatcaacct gcagatcctg 4200
tcggggcgag cctacctcca gcgcctctc tcgtcaagc tcgcaccgcg gatcgacgtc 4260
gacaccgagg acgacctctt cactaccgc cccgaccacg cactctaa 4308

```

<210> 6

<211> 1435

<212> PRT

<213> *Schizochytrium* sp.

<400> 6

5

ES 2 623 610 T3

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

Lys Gly Asp Arg Val Tyr Arg Leu Leu Asn Thr Ser Leu Thr Phe Phe
 130 135 140
 Gly Val Ala His Glu Gly Glu Thr Leu Val Tyr Asp Ile Arg Val Thr
 145 150 155 160
 Gly Phe Ala Lys Gly Ala Gly Gly Glu Ile Ser Met Phe Phe Phe Glu
 165 170 175
 Tyr Asp Cys Phe Val Asp Gly Arg Leu Leu Ile Glu Met Arg Asp Gly
 180 185 190
 Cys Ala Gly Phe Phe Thr Asp Ala Glu Leu Ala Ala Gly Lys Gly Val
 195 200 205
 Leu Lys Thr Lys Ala Glu Leu Ala Ala Arg Ala Gln Ile Gln Lys Gln
 210 215 220
 Asp Ile Ala Pro Phe Ala Pro Ala Pro Cys Ser His Lys Thr Ser Leu
 225 230 235 240
 Asp Ala Arg Glu Met Arg Leu Leu Val Asp Arg Gln Trp Ala Arg Val
 245 250 255
 Phe Gly Ser Gly Met Ala Gly Ile Asp Tyr Lys Leu Cys Ala Arg Lys
 260 265 270
 Met Leu Met Ile Asp Arg Val Thr His Leu Asp Pro Arg Gly Gly Ala
 275 280 285
 His Gly Leu Gly Leu Leu Ile Gly Glu Lys Val Leu Glu Arg Asp His
 290 295 300
 Trp Tyr Phe Pro Cys His Phe Val Arg Asp Glu Val Met Ala Gly Ser
 305 310 315 320
 Leu Val Ser Asp Gly Cys Ser Gln Leu Leu Lys Val Tyr Met Leu Trp
 325 330 335
 Leu Gly Leu His Thr Thr Val Gly Ala Phe Asp Phe Arg Pro Val Ser
 340 345 350
 Gly His Ala Asn Lys Val Arg Cys Arg Gly Gln Ile Ser Pro His Lys
 355 360 365
 Gly Lys Leu Val Tyr Val Met Glu Ile Lys Glu Met Gly Phe Asp Ala
 370 375 380
 Lys Thr Gly Asp Pro Phe Ala Ile Ala Asp Val Asp Ile Ile Asp Val

ES 2 623 610 T3

385		390		395		400
Asn Phe Glu Glu Gly Gln Ala Phe Ala Gly Val Glu Asp Leu His Ser						
		405		410		415
Tyr Gly Gln Gly Asp Leu Arg Lys Lys Ile Val Val Asp Phe Lys Gly						
		420		425		430
Ile Ala Leu Ser Leu Gln Lys Arg Lys Glu Gln Gln Lys Glu Ser Met						
		435		440		445
Thr Val Thr Thr Thr Thr Thr Thr Thr Ser Arg Val Ile Ala Pro Pro						
		450		455		460
Ser Gly Cys Leu Lys Gly Asp Pro Thr Ala Pro Thr Ser Val Thr Trp						
465		470		475		480
His Pro Met Ala Glu Gly Asn Gly Gly Pro Gly Pro Thr Pro Ser Phe						
		485		490		495
Ser Pro Ser Ala Tyr Pro Pro Arg Ala Val Cys Phe Ser Pro Phe Pro						
		500		505		510
Asn Asn Pro Leu Asp Asn Asp His Thr Pro Gly Gln Met Pro Leu Thr						
		515		520		525
Trp Phe Asn Met Ser Glu Phe Met Cys Gly Lys Val Ser Asn Cys Leu						
		530		535		540
Gly Pro Glu Phe Ala Arg Phe Asp Ala Ser Lys Thr Ser Arg Ser Pro						
545		550		555		560
Ala Phe Asp Leu Ala Leu Val Thr Arg Val Thr Ser Val Ala Asp Met						
		565		570		575
Glu His Gly Pro Phe Tyr Asn Val Asp Val Asn Pro Gly Gln Gly Thr						
		580		585		590
Met Val Gly Glu Phe Asp Cys Pro Ala Asp Ala Trp Phe Phe Gly Ala						
		595		600		605
Ser Ser Arg Asp Asp His Met Pro Tyr Ser Ile Leu Met Glu Ile Ala						
		610		615		620
Leu Gln Thr Ser Gly Val Leu Thr Ser Val Leu Lys Ala Pro Leu Thr						
625		630		635		640
Met Asp Lys Asp Asp Ile Leu Phe Arg Asn Leu Asp Ala Asp Ala Glu						
		645		650		655

ES 2 623 610 T3

Leu Val Gly Asp Ala Met Pro Asp Val Arg Gly Lys Thr Ile Arg Asn
 660 665 670
 Phe Thr Lys Cys Thr Gly Tyr Ser Met Leu Gly Lys Met Gly Ile His
 675 680 685
 Arg Phe Thr Phe Glu Leu Ser Val Asp Gly Ala Val Phe Tyr Lys Gly
 690 695 700
 Ser Thr Ser Phe Gly Trp Phe Val Pro Glu Val Phe Glu Ser Gln Thr
 705 710 715 720
 Gly Leu Asp Asn Gly Lys Pro Arg Leu Pro Trp Tyr Arg Glu Asn Asn
 725 730 735
 Val Ala Val Asp Thr Leu Ser Ala Pro Ala Ser Ala Ser Ser Ala Gln
 740 745 750
 Gly Gln Leu Gln Leu Gln Arg Arg Gly Ser Gln Ala Gln Phe Leu Asp
 755 760 765
 Thr Ile His Leu Ala Gly Ser Gly Ala Gly Val His Gly Gln Gly Tyr
 770 775 780
 Ala His Gly Glu Lys Ala Val Asn Lys Gln Asp Trp Phe Phe Ser Cys
 785 790 795 800
 His Phe Trp Phe Asp Pro Val Met Pro Gly Ser Leu Gly Ile Glu Ser
 805 810 815
 Met Phe Gln Leu Val Glu Ala Trp Cys Val Lys Gln Gly Leu Ala Ala
 820 825 830
 Arg His Gly Ile Ala His Pro Val Phe Ala His Ala Pro Gly Ala Thr
 835 840 845
 Ser Trp Lys Tyr Arg Gly Gln Leu Thr Pro Lys Asn Asp Arg Met Asp
 850 855 860
 Ser Glu Val His Ile Lys Ser Val Ala Ala Phe Ser Ser Trp Val Asp
 865 870 875 880
 Val Val Ala Asp Gly Phe Leu Phe Val Asp Gly Leu Arg Val Tyr Ser
 885 890 895
 Ala Asp Asn Leu Arg Val Arg Ile Gln Thr Gly Ala Gly His Val Glu
 900 905 910

Glu Gln Glu Val Ala Ala Lys Ala Thr Thr Lys Asn Ser Ser Ile Ala
 915 920 925
 Asp Val Asp Val Ala Asp Leu Gln Ala Leu Lys Gln Ala Leu Leu Thr
 930 935 940
 Leu Glu Arg Pro Leu Gln Leu Asp Ala Gly Ser Glu Val Pro Ala Cys
 945 950 955 960
 Ala Val Ser Asp Leu Gly Asp Arg Gly Phe Met Glu Thr Tyr Gly Val
 965 970 975
 Val Ala Pro Leu Tyr Ser Gly Ala Met Ala Lys Gly Ile Ala Ser Ala
 980 985 990
 Asp Leu Val Ile Ala Met Gly Gln Arg Lys Met Leu Gly Ser Phe Gl
 995 1000 1005
 Ala Gly Gly Leu Pro Met His Val Val Arg Ala Gly Ile Glu Lys
 1010 1015 1020
 Ile Gln Ala Ala Leu Pro Ala Gly Pro Tyr Ala Val Asn Leu Ile
 1025 1030 1035
 His Ser Pro Phe Asp Ala Asn Leu Glu Lys Gly Asn Val Asp Leu
 1040 1045 1050
 Phe Leu Glu Lys Gly Val Arg Val Val Glu Ala Ser Ala Phe Met
 1055 1060 1065
 Glu Leu Thr Pro Gln Val Val Arg Tyr Arg Ala Thr Gly Leu Ser
 1070 1075 1080
 Arg Asp Ala Arg Gly Gly Ser Val Arg Thr Ala His Lys Ile Ile
 1085 1090 1095
 Gly Lys Val Ser Arg Thr Glu Leu Ala Glu Met Phe Ile Arg Pro
 1100 1105 1110
 Ala Pro Gln Ala Ile Leu Asp Lys Leu Val Ala Ser Gly Glu Ile
 1115 1120 1125
 Thr Pro Glu Gln Ala Ala Leu Ala Leu Glu Val Pro Met Ala Asp
 1130 1135 1140
 Asp Ile Ala Val Glu Ala Asp Ser Gly Gly His Thr Asp Asn Arg
 1145 1150 1155

Pro Ile	His Val	Ile Leu	Pro	Leu Ile	Leu Ser	Leu	Arg Asn	Arg
1160			1165			1170		
Leu Gln	Arg Glu	Leu Lys	Tyr	Pro Ala	Arg His	Arg	Val Arg	Val
1175			1180			1185		
Gly Ala	Gly Gly	Gly Ile	Gly	Cys Pro	Gln Ala	Ala	Leu Gly	Ala
1190			1195			1200		
Phe His	Met Gly	Ala Ala	Phe	Val Val	Thr Gly	Thr	Val Asn	Gln
1205			1210			1215		
Leu Ser	Arg Gln	Ala Gly	Thr	Cys Asp	Asn Val	Arg	Arg Gln	Leu
1220			1225			1230		
Ser Arg	Ala Thr	Tyr Ser	Asp	Ile Thr	Met Ala	Pro	Ala Ala	Asp
1235			1240			1245		
Met Phe	Glu Gln	Gly Val	Glu	Leu Gln	Val Leu	Lys	Lys Gly	Thr
1250			1255			1260		
Met Phe	Pro Ser	Arg Ala	Lys	Lys Leu	Phe Glu	Leu	Phe His	Lys
1265			1270			1275		
Tyr Asp	Ser Phe	Glu Ala	Met	Pro Ala	Asp Glu	Leu	Ala Arg	Val
1280			1285			1290		
Glu Lys	Arg Ile	Phe Ser	Lys	Ser Leu	Ala Glu	Val	Trp Ala	Glu
1295			1300			1305		
Thr Lys	Asp Phe	Tyr Ile	Thr	Arg Leu	Asn Asn	Pro	Glu Lys	Ile
1310			1315			1320		
Arg Lys	Ala Glu	Asn Glu	Asp	Pro Lys	Leu Lys	Met	Ser Leu	Cys
1325			1330			1335		
Phe Arg	Trp Tyr	Leu Gly	Leu	Ser Ser	Phe Trp	Ala	Asn Asn	Gly
1340			1345			1350		
Ile Ala	Asp Arg	Thr Met	Asp	Tyr Gln	Ile Trp	Cys	Gly Pro	Ala
1355			1360			1365		
Ile Gly	Ala Phe	Asn Asp	Phe	Ile Ala	Asp Ser	Tyr	Leu Asp	Val
1370			1375			1380		
Ala Val	Ser Gly	Glu Phe	Pro	Asp Val	Val Gln	Ile	Asn Leu	Gln
1385			1390			1395		
Ile Leu	Ser Gly	Ala Ala	Tyr	Leu Gln	Arg Leu	Leu	Ser Val	Lys

1400	1405	1410
Leu Ala Pro Arg Ile Asp Val Asp Thr Glu Asp Asp Leu Phe Thr		
1415	1420	1425
Tyr Arg Pro Asp His Ala Leu		
1430	1435	

<210> 7
 <211> 1395
 5 <212> ADN
 <213> *Schizochytrium* sp.
 <400>7

```

actcgcatcg cgatcgtggg gatgtcggcg atcctgccga gcggggagaa cgtgcgcgag      60
agctgggagg cgatccgcga tgggctggat tgcctgagcg atctgccggc ggaccgcgtg      120
gacgtgacgg cctactacaa cccggagaag acgaccaagg acaagatcta ctgcaagcgc      180
ggcgggttca tccccgagta cgacttcgac gcgcgtagt tggggtcaa catgttccag      240
atggaggact cggacgcaa ccagacgata tcgctgctca aggtgaagga ggcgctgacg      300
gacgccaaca tccccgcgtt ctcgagcggg aagaagaaca tcggctgcgt gctgggcatac      360
ggcggcgggc agaaggcgag ccacgagttc tactcgcggc tcaactacgt ggtcgtggac      420
aagggtgctgc gcaagatggg cctgccggag gaagacgtgg cggcggcggg ggacaagtac      480
aaggcgagtt tccccgagtg gcgcctcgac tctttccccg ggttcctggg caacgtcacg      540
gcggggcgct gctgcaatac cttcaacatg gagggcatga actgcgtcgt ggacgcggcc      600
tgcgcgctcg cgctgatcgc ggtcaaagtg gcgatcgagg agctgctcta cggcgactgc      660
gatgcgatga tcgcgggtgc cacctgcacg gacaactcga tcgggatgta catggccttc      720
tccaagacgc ccgtgttttc cacggacccg agcgtcaagg cgtacgacgc cgccaccaa      780
ggcatgctca tcggcgaggg ctcgcgatg ctcgctgctga agcgctacgc ggacgccgtg      840
cgcgacggcg acaccgtgca cgccgtcatc aagggtgctg cgtcctcgag cgacggcaag      900
gcggcgggca tctacacgcc gacaatctcg ggccaggagg aggccctgcg ccgcgcctac      960
gcccgcgcca atgtcgaccc ggccactgtg acgctggtgg agggccacgg cacgggtacg     1020
ccggtgggcg acaagatcga gctgacggcg ctgagcaacc tcttctccaa ggcgttttct     1080
gccaacgggtg gcggcgcgga ggaagcagag caggtggcgg tgggcagcat caagtgcag      1140
atcgggcacc tcaaggcggg ggccgggctg gccgggctgg tcaaggtggt gctggcgctc     1200
aagcacaaga cgctgccgca gacgatcaac gtcgacaagc cgccgtcgct ggtggacggg     1260
accccgatcc agcagtcgcc gctgtacgtc aacacgatga accgcccctg gttcacgccc     1320
gtaggggtgc cgcgccgcgc cggcgtgtcg tcgtttgggt ttggcgggtg caactaccac     1380
gccgtgctgg aggag                                         1395
    
```

<210> 8
<211> 465
<212> PRT
<213> *Schizochytrium* sp.

5

<400> 8

ES 2 623 610 T3

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

ES 2 623 610 T3

225		230		235		240									
Ser	Lys	Thr	Pro	Val	Phe	Ser	Thr	Asp	Pro	Ser	Val	Lys	Ala	Tyr	Asp
				245					250					255	
Ala	Ala	Thr	Lys	Gly	Met	Leu	Ile	Gly	Glu	Gly	Ser	Ala	Met	Leu	Val
			260					265					270		
Leu	Lys	Arg	Tyr	Ala	Asp	Ala	Val	Arg	Asp	Gly	Asp	Thr	Val	His	Ala
		275					280					285			
Val	Ile	Lys	Gly	Cys	Ala	Ser	Ser	Ser	Asp	Gly	Lys	Ala	Ala	Gly	Ile
	290					295					300				
Tyr	Thr	Pro	Thr	Ile	Ser	Gly	Gln	Glu	Glu	Ala	Leu	Arg	Arg	Ala	Tyr
305					310					315					320
Ala	Arg	Ala	Asn	Val	Asp	Pro	Ala	Thr	Val	Thr	Leu	Val	Glu	Gly	His
			325						330					335	
Gly	Thr	Gly	Thr	Pro	Val	Gly	Asp	Lys	Ile	Glu	Leu	Thr	Ala	Leu	Ser
			340					345					350		
Asn	Leu	Phe	Ser	Lys	Ala	Phe	Ser	Ala	Asn	Gly	Gly	Gly	Ala	Glu	Glu
		355					360					365			
Ala	Glu	Gln	Val	Ala	Val	Gly	Ser	Ile	Lys	Ser	Gln	Ile	Gly	His	Leu
	370					375					380				
Lys	Ala	Val	Ala	Gly	Leu	Ala	Gly	Leu	Val	Lys	Val	Val	Leu	Ala	Leu
385					390					395					400
Lys	His	Lys	Thr	Leu	Pro	Gln	Thr	Ile	Asn	Val	Asp	Lys	Pro	Pro	Ser
				405					410					415	
Leu	Val	Asp	Gly	Thr	Pro	Ile	Gln	Gln	Ser	Pro	Leu	Tyr	Val	Asn	Thr
			420					425					430		
Met	Asn	Arg	Pro	Trp	Phe	Thr	Pro	Val	Gly	Val	Pro	Arg	Arg	Ala	Gly
		435					440					445			
Val	Ser	Ser	Phe	Gly	Phe	Gly	Gly	Ala	Asn	Tyr	His	Ala	Val	Leu	Glu
	450					455					460				
Glu															
465															

ES 2 623 610 T3

<210> 9
<211> 903
<212> ADN
<213> *Schizochytrium* sp.

5

<400> 9

ttctcgggce agggcgcgca gtacacgcac atgttcagcg acgtggcgat gaactggccc	60
ccgttcgcgc agagcgtcgc cgccatggac cgcgcccagc gcgagcgctt cgggcggcct	120
gccaaagcgc tgagcagcgt gctgtaccgc cgcaagccgt acggcgacga accgcggcag	180
gaccacaagg agatctcgca aacgcgctac tcgcagcccg caacgcctgc gtgctcggtc	240
ggcgcccttg acatcttcaa agcggcgga ctggcgccga gctttgcggc gggccaactcg	300
ctgggcgagt ttgcggcgct ctacgcggcc gggtcgctcg atcgcgacgc cgtcttcgac	360
ctgggtctgc cgcgcgcca ggccatgagc gacttcacgg ccagggccag cagcagcggt	420
ggcgccatgg cggcggtgat tggcgccaag gcggaccagc tctcgctggg tggcgcgccc	480
gacgtgtggc tcgccaacag caactcgccc tcgcagaccg tgatcacggg aaccgccgaa	540
gcagtggctg cggcctctga caagttgcgc tgcagcggca acttcgcgt cgtgcctctg	600
gcctgcgagg cggccttcca ctgcgcgcac atgcgcggcg cggagcagac gtttgcgtcg	660
gcgctcgcgc aggcgcgcgt gtgggcaccg gcggctgctc ggttctactc taacgtgacg	720
gggggcgcgc cggtaacctc gccgcggac gtcaaaacga acctgggcaa gcacatgacg	780
agccctgtgc agttcgtgca gcaggtgcga gccatgcacg cggcgggcgc gcgtgtgttt	840
gtggagtttg ggcccaagca ggtcctgtcg cgctcgtca aggagaccct tggcgaggcc	900
ggc	903

10 <210> 10
<211> 301
<212> PRT
<213> *Schizochytrium* sp.

15 <400> 10

ES 2 623 610 T3

Phe Ser Gly Gln Gly Ala Gln Tyr Thr His Met Phe Ser Asp Val Ala
1 5 10 15

Met Asn Trp Pro Pro Phe Arg Glu Ser Val Ala Ala Met Asp Arg Ala
20 25 30

Gln Arg Glu Arg Phe Gly Arg Pro Ala Lys Arg Val Ser Ser Val Leu
35 40 45

Tyr Pro Arg Lys Pro Tyr Gly Asp Glu Pro Arg Gln Asp His Lys Glu
50 55 60

Ile Ser Gln Thr Arg Tyr Ser Gln Pro Ala Thr Leu Ala Cys Ser Val
65 70 75 80

Gly Ala Phe Asp Ile Phe Lys Ala Ala Gly Leu Ala Pro Ser Phe Ala

ES 2 623 610 T3

				85						90					95
Ala	Gly	His	Ser	Leu	Gly	Glu	Phe	Ala	Ala	Leu	Tyr	Ala	Ala	Gly	Ser
			100					105					110		
Leu	Asp	Arg	Asp	Ala	Val	Phe	Asp	Leu	Val	Cys	Ala	Arg	Ala	Lys	Ala
		115					120					125			
Met	Ser	Asp	Phe	Thr	Ala	Gln	Ala	Ser	Ser	Ser	Gly	Gly	Ala	Met	Ala
	130					135					140				
Ala	Val	Ile	Gly	Ala	Lys	Ala	Asp	Gln	Leu	Ser	Leu	Gly	Gly	Ala	Pro
145					150					155					160
Asp	Val	Trp	Leu	Ala	Asn	Ser	Asn	Ser	Pro	Ser	Gln	Thr	Val	Ile	Thr
				165					170					175	
Gly	Thr	Ala	Glu	Ala	Val	Ala	Ala	Ala	Ser	Asp	Lys	Leu	Arg	Cys	Ser
			180					185					190		
Gly	Asn	Phe	Arg	Val	Val	Pro	Leu	Ala	Cys	Glu	Ala	Ala	Phe	His	Ser
		195					200					205			
Pro	His	Met	Arg	Gly	Ala	Glu	Gln	Thr	Phe	Ala	Ser	Ala	Leu	Ala	Gln
	210					215					220				
Ala	Pro	Val	Ser	Ala	Pro	Ala	Ala	Ala	Arg	Phe	Tyr	Ser	Asn	Val	Thr
225					230					235					240
Gly	Gly	Ala	Ala	Val	Thr	Ser	Pro	Ala	Asp	Val	Lys	Thr	Asn	Leu	Gly
				245					250					255	
Lys	His	Met	Thr	Ser	Pro	Val	Gln	Phe	Val	Gln	Gln	Val	Arg	Ala	Met
			260					265					270		
His	Ala	Ala	Gly	Ala	Arg	Val	Phe	Val	Glu	Phe	Gly	Pro	Lys	Gln	Val
		275					280					285			
Leu	Ser	Arg	Leu	Val	Lys	Glu	Thr	Leu	Gly	Glu	Ala	Gly			
	290					295					300				

<210> 11
 <211> 2103
 <212> ADN
 <213> *Schizochytrium* sp.

<400> 11

ES 2 623 610 T3

tccggcaaca gcaagagcac tcgtggcagt gctgatctgc aagcgtgct ggccaaggcg 60
gagactgtgg tgatggetgt gctggetgcc aagactggct acgaggccga catggttgag 120

ES 2 623 610 T3

gcggacatgg acctggaggg cgagctcggc atcgactcga tcaagcgcggt ggagatcctt 180
 tccgaggtgc agggccagct gggcgctcag gccaaaggacg tggatgcgct gagccgcacg 240
 cgcacggctc gtgaggttgt ggacgccatg aaggcggaga tcgtggctgc ctctggtggt 300
 agtgcctctg cggttccttc ggcgcccgct gcttctgcag ctccgactcc cgtgcttcg 360
 actgcgcctt ctgctgatct gcaagcgctg ctgtccaagg cggagactgt ggtgatggct 420
 gtgctggcgg ccaagactgg ctacgaggcc gacatggctg aggcggacat ggacctggag 480
 gccgagctcg gcacgcactc gatcaagcgc gtggagatcc tctcggaggt gcagggccag 540
 ctgggcgctc aggccaaagga cgtggatgcg ctgagccgca cgcgcacgggt cggtagggtt 600
 gtggatgcca tgaaggcgga aatcgtggct gcctctgctg gtagtgctcc tgcctctgct 660
 gttccttcgg cgcccgctgc ttctgcagct ccgactcccg ctgcttcgac tgcgccttct 720
 gctgatctgc aagcgctgct gtccaaggcg gagacgggtg tgatggctgt gctggcgcc 780
 aagactggct acgaggccga catggctcag gcggacatgg acctggaggg cgagctcggc 840
 atcgactcga tcaagcgcggt ggagatcctc tcggaggtgc agggccagct gggcgctcag 900
 gccaaaggacg tggatgcgct gagccgcacg cgcacggctc gtgaggttgt ggatgccatg 960
 aaggcggaaa tcgtggctgc ctctggtggt agtgcctctg ctctgcgggt tccttcggcg 1020
 cccgctgctt ctgcagctcc gactcccgcg gctgcgacag cgccttctgc tgatctgcaa 1080
 gcgctgctgg ccaaggcgga gactgtggtg atggctgtgc tggcgggcaa gactggctac 1140
 gagggcgaca tggctcagggc ggacatggac ctggaggccg agctcgccat cgactcgatc 1200
 aagcgcggtg agatcctttc cgaggtgcag ggccagctgg gcgtcgaggc caaggacgta 1260
 gatgcgctga gccgcacgcg cacggctcgg gaggttgttg atgccatgaa ggccgagatc 1320
 gtggctgcct ctgctggtag tgcctctgct cctgctgttc ctccggcgcc cgtgcttct 1380
 gcagctccga ctcccgctgc ttcgactgcg ccttctgctg atctgcaagc gctgctgtcc 1440
 aaggcggaga ctgtggtgat ggctgtgctg gcggccaaga ctggctacga ggccgacatg 1500
 gtcgagggcg acatggacct ggaggccgag ctccggcatc actcgatcaa gcgcgtggag 1560
 atcctctcgg aggtgcaggg ccagctgggc gtcgaggcca aggacgtgga tgcgctgagc 1620
 cgcacgcgca cggctcggta ggttgtggat gccatgaagg cggaaatcgt ggctgcctct 1680
 ggtggttagt ctctgctgc tgotgttctc tcggcgcccg ctgcttctgc agctccgact 1740
 cctgcgactg cgccttctgc tgatctgcaa gcgctgctgt ccaaggcgga gactgtggtg 1800
 atggctgtgc tggcgggcaa gactggctac gagggcgaca tggctcagggc ggacatggac 1860
 ctggaggccg agctcgccat cgactcgatc aagcgcggtg agatcctttc cgaggtgcag 1920
 ggccagctgg gcgtcgaggc caaggacgta gatgcgctga gccgcacgcg cacggctcgg 1980
 gaagtgggtg accccatgaa ggccgagatc gtggctgcct ctggtggtag tgcctctgct 2040
 gctccttcgg cgcccgctc tcttccaacg ctgtttggtt ccgagtgcga ggacctgtct 2100
 ctg 2103

<210> 12
 <211> 701
 <212> PRT
 <213> *Schizochytrium* sp.

5

<400> 12

```

Ser Gly Asn Ser Lys Ser Thr Arg Gly Ser Ala Asp Leu Gln Ala Leu
1      5      10      15

Leu Ala Lys Ala Glu Thr Val Val Met Ala Val Leu Ala Ala Lys Thr
20      25      30

Gly Tyr Glu Ala Asp Met Val Glu Ala Asp Met Asp Leu Glu Ala Glu
35      40      45

Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln
50      55      60

Gly Gln Leu Gly Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr
65      70      75      80

Arg Thr Val Gly Glu Val Val Asp Ala Met Lys Ala Glu Ile Val Ala
85      90      95

Ala Ser Gly Gly Ser Ala Pro Ala Val Pro Ser Ala Pro Ala Ala Ser
100     105     110

Ala Ala Pro Thr Pro Ala Ala Ser Thr Ala Pro Ser Ala Asp Leu Gln
115     120     125

Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met Ala Val Leu Ala Ala
130     135     140

Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala Asp Met Asp Leu Glu
145     150     155     160

Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu
165     170     175

Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp Val Asp Ala Leu Ser
180     185     190

Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala Met Lys Ala Glu Ile
195     200     205

Val Ala Ala Ser Ala Gly Ser Ala Pro Ala Pro Ala Val Pro Ser Ala
210     215     220

```

Pro Ala Ala Ser Ala Ala Pro Thr Pro Ala Ala Ser Thr Ala Pro Ser
 225 230 235 240
 Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met Ala
 245 250 255
 Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala Asp
 260 265 270
 Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu
 275 280 285
 Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp Val
 290 295 300
 Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala Met
 305 310 315 320
 Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser Ala Pro Ala Pro Ala
 325 330 335
 Val Pro Ser Ala Pro Ala Ala Ser Ala Ala Pro Thr Pro Ala Ala Ala
 340 345 350
 Thr Ala Pro Ser Ala Asp Leu Gln Ala Leu Leu Ala Lys Ala Glu Thr
 355 360 365
 Val Val Met Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met
 370 375 380
 Val Glu Ala Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile
 385 390 395 400
 Lys Arg Val Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu
 405 410 415
 Ala Lys Asp Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val
 420 425 430
 Val Asp Ala Met Lys Ala Glu Ile Val Ala Ala Ser Ala Gly Ser Ala
 435 440 445
 Pro Ala Pro Ala Val Pro Ser Ala Pro Ala Ala Ser Ala Ala Pro Thr
 450 455 460
 Pro Ala Ala Ser Thr Ala Pro Ser Ala Asp Leu Gln Ala Leu Leu Ser
 465 470 475 480

Lys Ala Glu Thr Val Val Met Ala Val Leu Ala Ala Lys Thr Gly Tyr
485 490 495

Glu Ala Asp Met Val Glu Ala Asp Met Asp Leu Glu Ala Glu Leu Gly
500 505 510

Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln Gly Gln
515 520 525

Leu Gly Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr Arg Thr
530 535 540

Val Gly Glu Val Val Asp Ala Met Lys Ala Glu Ile Val Ala Ala Ser
545 550 555 560

Gly Gly Ser Ala Pro Ala Ala Ala Val Pro Ser Ala Pro Ala Ala Ser
565 570 575

Ala Ala Pro Thr Pro Ala Thr Ala Pro Ser Ala Asp Leu Gln Ala Leu
580 585 590

Leu Ser Lys Ala Glu Thr Val Val Met Ala Val Leu Ala Ala Lys Thr
595 600 605

Gly Tyr Glu Ala Asp Met Val Glu Ala Asp Met Asp Leu Glu Ala Glu
610 615 620

Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln
625 630 635 640

Gly Gln Leu Gly Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr
645 650 655

Arg Thr Val Gly Glu Val Val Asp Ala Met Lys Ala Glu Ile Val Ala
660 665 670

Ala Ser Gly Gly Ser Ala Pro Ala Ala Pro Ser Ala Pro Ala Leu Leu
675 680 685

Pro Thr Leu Phe Gly Ser Glu Cys Glu Asp Leu Ser Leu
690 695 700

<210> 13
<211> 276
5 <212> ADN
<213> *Schizochytrium* sp.

<400> 13

```

agtgtgtgatc tgcaagcgct gctggccaag gcgagactg tggatgatggc tgtgtgtggct      60
gccaagactg gctacgaggg cgacatgggt gagggcgaca tggacctgga ggccgagctc      120
ggcatcgact cgatcaagcg cgtggagatc ctttcggagg tgcagggcca gctggggcgtc      180
gaggccaagg acgtggatgc gctgagccgc acgcgcacgg tcggtgaggt tgtggacgcc      240
atgaaggcgg agatcgtggc tgcctctggt ggtagt                                  276

```

<210> 14
 <211> 92
 <212> PRT
 <213> *Schizochytrium* sp.

<400> 14

```

Ser Ala Asp Leu Gln Ala Leu Leu Ala Lys Ala Glu Thr Val Val Met
1          5          10          15

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20          25          30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35          40          45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50          55          60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65          70          75          80

Met Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser
85          90

```

<210> 15
 <211> 276
 <212> ADN
 <213> *Schizochytrium* sp.

<400> 15

```

tctgtgtgatc tgcaagcgct gctgtccaag gcgagactg tggatgatggc tgtgtgtggcg      60
gccaagactg gctacgaggg cgacatgggt gagggcgaca tggacctgga ggccgagctc      120
ggcatcgact cgatcaagcg cgtggagatc ctctcggagg tgcagggcca gctggggcgtc      180
gaggccaagg acgtggatgc gctgagccgc acgcgcacgg tcggtgaggt tgtggatgcc      240
atgaaggcgg aaatcgtggc tgcctctgct ggtagt                                  276

```

<210> 16
 <211> 92
 <212> PRT

ES 2 623 610 T3

<213> *Schizochytrium* sp.

<400> 16

Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met

1 5 10 15

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20 25 30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35 40 45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50 55 60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65 70 75 80

Met Lys Ala Glu Ile Val Ala Ala Ser Ala Gly Ser
85 90

5

<210> 17

<211> 276

<212> ADN

10 <213> *Schizochytrium* sp.

<400> 17

tctgctgata tgcaagcgct gctgtccaag gcggagacgg tggatgatggc tgtgctggcg 60
gccaagactg gctacgaggg cgacatgggc gaggcggaca tggacactgga ggccgagctc 120
ggcatcgact cgatcaagcg cgtggagatc ctctcggagg tgcagggcca gctgggctgc 180
gaggccaagg acgtggatgc gctgagccgc acgcgcacgg tcggtgaggt tgtggatgcc 240
atgaaggcgg aaatcgtggc tgcctctggt ggtagt 276

15

<210> 18

<211> 92

<212> PRT

20 <213> *Schizochytrium* sp.

<400> 18

ES 2 623 610 T3

Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met
1 5 10 15

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20 25 30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35 40 45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50 55 60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65 70 75 80

Met Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser
85 90

<210> 19

<211> 276

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 19

tctgctgata tgcaagcgct gctggccaag gcggagactg tggatgatggc tgtgctggcg 60

gccaagactg gctacgaggc cgacatgggc gaggcggaca tggacctgga ggccgagctc 120

ggcatcgact cgatcaagcg cgtggagatc ctttccgagg tgcagggcca gctgggcgctc 180

gaggccaagg acgtagatgc gctgagccgc acgcgcacgg tcggtgaggt tgtggatgcc 240

atgaaggcgg agatcgtggc tgcctctgct ggtagt 276

10

<210> 20

<211> 92

<212> PRT

15 <213> *Schizochytrium* sp.

<400> 20

ES 2 623 610 T3

Ser Ala Asp Leu Gln Ala Leu Leu Ala Lys Ala Glu Thr Val Val Met
1 5 10 15

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20 25 30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35 40 45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50 55 60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65 70 75 80

Met Lys Ala Glu Ile Val Ala Ala Ser Ala Gly Ser
85 90

<210> 21

<211> 276

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 21

tctgctgac tgcaagcgct gctgtccaag gcggagactg tggatgatggc tgtgctggcg 60

gccaagactg gctacgaggc cgacatgggc gaggcggaca tggacctgga ggccgagctc 120

ggcatcgact cgatcaagcg cgtggagatc ctctcggagg tgcagggcca gctgggctc 180

gaggccaagg acgtggatgc gctgagccgc acgcgcacgg tcggtgaggt tgtggatgcc 240

atgaaggcgg aaatcgtggc tgccctctggt ggtagt 276

10

<210> 22

<211> 92

<212> PRT

15 <213> *Schizochytrium* sp.

<400> 22

Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met
1 5 10 15

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20 25 30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35 40 45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50 55 60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65 70 75 80

Met Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser
85 90

<210> 23

<211> 276

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 23

tctgctgacg tgcaagcgct gctgtccaag gcgagactg tggatggc tgtgctggcg 60

gccaagactg gctacgagcg cgacatgggc gagcgagaca tggacctgga ggccgagctc 120

ggcatcgact cgatcaagcg cgtggagatc ctttcgagg tgcagggcca gctgggcgctc 180

gaggccaagg acgtagatgc gctgagccgc acgcgcacgg tcgggtgaagt ggtggacgcc 240

atgaaggcgg agatcgtggc tgcctctggt ggtagt 276

10

<210> 24

<211> 92

<212> PRT

15 <213> *Schizochytrium* sp.

<400> 24

Ser Ala Asp Leu Gln Ala Leu Leu Ser Lys Ala Glu Thr Val Val Met
1 5 10 15

ES 2 623 610 T3

Ala Val Leu Ala Ala Lys Thr Gly Tyr Glu Ala Asp Met Val Glu Ala
20 25 30

Asp Met Asp Leu Glu Ala Glu Leu Gly Ile Asp Ser Ile Lys Arg Val
35 40 45

Glu Ile Leu Ser Glu Val Gln Gly Gln Leu Gly Val Glu Ala Lys Asp
50 55 60

Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Val Asp Ala
65 70 75 80

Met Lys Ala Glu Ile Val Ala Ala Ser Gly Gly Ser
85 90

<210> 25

<211> 2178

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 25

ES 2 623 610 T3

gcgctacagg cggcgctcac gtccgtcgag gcgcagttcg gcaagggtggg tggctttgtg	60
ttccagttcg gcgacgacga cgtgcaagcg cagctcggct gggcgctgct cgcggccaag	120
cacctcaaaa cttcgctgtc agaacagatc gagggcggtc gcaccttttt cgtggccgtc	180
gcgcggctcg acggccagct ggggtctctc ggcaagtoga cgaccgctac cgttgatctc	240
tcccgcgcgc agcagggcag cgtgttcggc ctgtgcaaga cactcgacct ggagtggccc	300
gctgtcttct gccgcggaat cgacctggcc gccgacctcg acgccgcaca ggccgcgcgg	360
tgcctgctgg gcgagctgtc agaccccgac gtggccgtgc gcgagtctgg ttactccgcc	420
tcgggccagc gctgcacgac aactacgaag tcgctgacta cgggcaagcc gcaccagccg	480
atctcctcgt cggacctctt tctggtgtcg ggcggcgcgc gcggcatcac cccgctgtgc	540
gtgcgcgagc tggcgcgagc cgtgggcggc ggcacgtacg tgctcatcgg ccgctcggag	600
ctgcccacga cggagcctgc ctgggcggtc ggcgtggagt ctggcaagcc gctggagaag	660
gccgcgctgg cgttcctgaa ggcggaagtt gcagcgggce gcggggccaa gccgacgccg	720
atgctgcaca agaagctcgt gggcgccgtg gtcggagcgc gcgaggtgcg agcctcgctc	780
gccgagatca ctgcacagg cgccacggct gtgtacgagt cgtgcgacgt gagctctgcc	840
gccaaagggtc gtgagatggt agagcgcgtg cagcagcagg gcgggcggcg cgtgtcgggc	900
gtgttccacg cgtcggggct gctgcgcgac aagctcgtgg agaacaagtc gctggcggac	960
ttcagcgccg tgtacgacac caagggtggc ggcctcatca acctgctggc ctgcgtggac	1020
ctggcgcagc tgcgtcacct cgtgctcttc agctcgcctc cgggcttcca cggcaacgtc	1080
gggcagtcgg actacgcaat ggccaacgag gcgctcaaca agctggcggc gcacctgtcg	1140
gcggtgcacc cgcagctgtg cgcgcgctcg atctgcttcg gaccgtggga cggcggcatg	1200

ES 2 623 610 T3

```

gtgacccccg cgctcaaggc caacttcata cgcattgggca tccagatcat cccgcgccaa 1260
ggcgggcgcgc agaccgtcgc caacatgctc gtcagtagct cccccggtca gctgctcgtg 1320
ggcaactggg gcgtgccacc cgtcgtgccg agtgccaccg agcacaccgt gctgcagacg 1380
ctccgccaga gcgacaaccc ctccctcgac tcgcacgtga tccagggccg ccgcgtgctg 1440
cccatgaccc tggccgtggg ctacatggcg caccaggcgc agagcatcta cgcgggccac 1500
cagctgtggg ccgtcgagga cggccagctc ttcaagggca tcgccatcga caatggcgcc 1560
gacgtgcccg tgcgcgtgga gctgtcgcgc cgcaaggagg agcaggagga cgccggcaag 1620
gtcaaggcca aggtgcaggt gctgctcaaa tcgcaggcca acggcaagtc ggtgcccgcg 1680
tacaaggcga ccgtcgtgct gtccctcgcg ccgcgcccca gcgtcatcac gcgtgacttc 1740
gacctcacc cggacccggc ctgcacggag cagcactct acgacggcaa gacgtcttc 1800
cacggcaagg cctccaggg catcgagcag gtgctctcgg cgacgccaa gcagctcacc 1860
gccaagtgcc gcaatttgc cctcacgcc gagcagcgcg gccagtctgt cgttaacctc 1920
agccagcagg acccgttcca ggccgacatt gcgttccagg cgatgctcgt ctgggcgcgc 1980
atgtgcgcc aatcggcggc cctgcccaac aactgcgagc gcttcgactt ttacaagccg 2040
atggccccgg gcgccaccta ctacacgtcg gtcaagctgg cctcggcctc acccttggtg 2100
gactctgtgt gcaagtgcac cgtggcgatg cagcatgagc aaggtgaggt gtacttttct 2160
gctcgtgcca gcgtcgtc 2178

```

<210> 26
 <211> 726
 5 <212> PRT
 <213> *Schizochytrium* sp.
 <400> 26

ES 2 623 610 T3

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

Leu Glu Trp Pro Ala Val Phe Cys Arg Gly Ile Asp Leu Ala Ala Asp
 100 105 110

Leu Asp Ala Ala Gln Ala Ala Arg Cys Leu Leu Gly Glu Leu Ser Asp
 115 120 125

Pro Asp Val Ala Val Arg Glu Ser Gly Tyr Ser Ala Ser Gly Gln Arg
 130 135 140

Cys Thr Thr Thr Thr Lys Ser Leu Thr Thr Gly Lys Pro His Gln Pro
 145 150 155 160

Ile Ser Ser Ser Asp Leu Phe Leu Val Ser Gly Gly Ala Arg Gly Ile
 165 170 175

Thr Pro Leu Cys Val Arg Glu Leu Ala Gln Arg Val Gly Gly Gly Thr
 180 185 190

Tyr Val Leu Ile Gly Arg Ser Glu Leu Pro Thr Thr Glu Pro Ala Trp
 195 200 205

Ala Val Gly Val Glu Ser Gly Lys Pro Leu Glu Lys Ala Ala Leu Ala
 210 215 220

Phe Leu Lys Ala Glu Phe Ala Ala Gly Arg Gly Ala Lys Pro Thr Pro
 225 230 235 240

Met Leu His Lys Lys Leu Val Gly Ala Val Val Gly Ala Arg Glu Val
 245 250 255

Arg Ala Ser Leu Ala Glu Ile Thr Ala Gln Gly Ala Thr Ala Val Tyr
 260 265 270

Glu Ser Cys Asp Val Ser Ser Ala Ala Lys Val Arg Glu Met Val Glu
 275 280 285

Arg Val Gln Gln Gln Gly Gly Arg Arg Val Ser Gly Val Phe His Ala
 290 295 300

Ser Gly Val Leu Arg Asp Lys Leu Val Glu Asn Lys Ser Leu Ala Asp
 305 310 315 320

Phe Ser Ala Val Tyr Asp Thr Lys Val Gly Gly Leu Ile Asn Leu Leu
 325 330 335

Ala Cys Val Asp Leu Ala Gln Leu Arg His Leu Val Leu Phe Ser Ser
 340 345 350

Leu Ala Gly Phe His Gly Asn Val Gly Gln Ser Asp Tyr Ala Met Ala
 355 360 365
 Asn Glu Ala Leu Asn Lys Leu Ala Ala His Leu Ser Ala Val His Pro
 370 375 380
 Gln Leu Cys Ala Arg Ser Ile Cys Phe Gly Pro Trp Asp Gly Gly Met
 385 390 395 400
 Val Thr Pro Ala Leu Lys Ala Asn Phe Ile Arg Met Gly Ile Gln Ile
 405 410 415
 Ile Pro Arg Gln Gly Gly Ala Gln Thr Val Ala Asn Met Leu Val Ser
 420 425 430
 Ser Ser Pro Gly Gln Leu Leu Val Gly Asn Trp Gly Val Pro Pro Val
 435 440 445
 Val Pro Ser Ala Thr Glu His Thr Val Leu Gln Thr Leu Arg Gln Ser
 450 455 460
 Asp Asn Pro Phe Leu Asp Ser His Val Ile Gln Gly Arg Arg Val Leu
 465 470 475 480
 Pro Met Thr Leu Ala Val Gly Tyr Met Ala His Gln Ala Gln Ser Ile
 485 490 495
 Tyr Ala Gly His Gln Leu Trp Ala Val Glu Asp Ala Gln Leu Phe Lys
 500 505 510
 Gly Ile Ala Ile Asp Asn Gly Ala Asp Val Pro Val Arg Val Glu Leu
 515 520 525
 Ser Arg Arg Lys Glu Glu Gln Glu Asp Ala Gly Lys Val Lys Val Lys
 530 535 540
 Val Gln Val Leu Leu Lys Ser Gln Val Asn Gly Lys Ser Val Pro Ala
 545 550 555 560
 Tyr Lys Ala Thr Val Val Leu Ser Pro Ala Pro Arg Pro Ser Val Ile
 565 570 575
 Thr Arg Asp Phe Asp Leu Thr Pro Asp Pro Ala Cys Thr Glu His Asp
 580 585 590
 Leu Tyr Asp Gly Lys Thr Leu Phe His Gly Lys Ala Phe Gln Gly Ile
 595 600 605

ES 2 623 610 T3

Glu Gln Val Leu Ser Ala Thr Pro Lys Gln Leu Thr Ala Lys Cys Arg
610 615 620

Asn Leu Pro Leu Thr Pro Glu Gln Arg Gly Gln Phe Val Val Asn Leu
625 630 635 640

Ser Gln Gln Asp Pro Phe Gln Ala Asp Ile Ala Phe Gln Ala Met Leu
645 650 655

Val Trp Ala Arg Met Leu Arg Gln Ser Ala Ala Leu Pro Asn Asn Cys
660 665 670

Glu Arg Phe Asp Phe Tyr Lys Pro Met Ala Pro Gly Ala Thr Tyr Tyr
675 680 685

Thr Ser Val Lys Leu Ala Ser Ala Ser Pro Leu Val Asp Ser Val Cys
690 695 700

Lys Cys Thr Val Ala Met His Asp Glu Gln Gly Glu Val Tyr Phe Ser
705 710 715 720

Ala Arg Ala Ser Val Val
725

<210> 27
<211> 39
<212> ADN
<213> *Schizochytrium* sp.

<400> 27
ctcgactcgc acgtgatcca gggccgccgc gtgctgccc 39

<210> 28
<211> 13
<212> PRT
<213> *Schizochytrium* sp.

<400> 28

Leu Asp Ser His Val Ile Gln Gly Arg Arg Val Leu Pro
1 5 10

<210> 29
<211> 1341
<212> ADN
<213> *Schizochytrium* sp.

<400> 29

gataacattg cggtcgtggg catggcgggtg cagtatgccg gatgcaagaa ccaggacgag 60
 ttctgggata cgctgatgcg taaggagatc aactcgagcc cgatctcggc ggagcgccctc 120
 ggtacgcgct accgcgacct ccacttccac ccgcagcgca gcaagtacgc cgacaccttc 180
 tgcaacgata gctacggctg cgtcgatgcc agcgtcgaca acgagcacga cctcctcgcc 240
 gacctggccc ggcgcgccct gctcgacgcc ggaattaacc tcgacgacgc cagcaccacc 300
 gccaacctac gcgacttcgg catcgtgagc ggctgcctgt cgttccccat ggacaatctg 360
 cagggcgagc tgctcaatct gtaccaagtg catgtggaga accgcgtggg cggccagcgcc 420
 ttccgcgact cgcgccccctg gtcggagcgc ccgcgcgctg tctcgccga ggccagcgac 480
 ccgcgcgtgt actccgaccc ggcgtccttc gtggccaacc agctcggcct ggggcccgtg 540
 cgctacagcc tcgatgcagc ctgcgcgtcg gcgctgtact gcctcaagct ggcgctccgac 600
 cacttgctct cgcgcagcgc ggacgtgatg ctgtggggcg ccacatgctt tccggacccg 660
 ttcttcattc tctcgggggt ctccaccttc caggcgatgc cgctgggcgg accggacgat 720
 aaccactgt ccgtgccgct gcggcagggc agccagggcc tgacgcccgg agagggcggc 780
 gccatcatgy tgctgaagcg cctcgaggac gccgtgcgcg acggcgaccg catctacggc 840
 accttgctcg gcacgagtct gagcaacgcc ggggtgcggcc tgccgctgag ccgcgacctg 900
 ccgagcgaga agtcgtgcat ggaggacctg tacacgagcg tcggcatcga cccaagcgag 960
 gtgcagtacg tggagtgcc cgcacggggc actccgcagg gcgacgtcgt ggaggtagag 1020
 gcgctgcgcc actgctttcg aggtaacacg gaccacccgc cgcgcattgg ctccaccaag 1080
 ggcaactttg gccacactct cgtggcggcc gggttcgcag gcatggccaa ggtgctgctg 1140
 tcgatgcagc acggcacgat cccgcccacg cccggtgtcg accgtccaa ctgcatcgac 1200
 ccgctcgtcg tggacgaggc catcccttgg ccgtactcgt cggcgaggc gcgggcaggc 1260
 aaaccaggcg atgagctcaa gtgcgcctcg ctctccgcct ttggctttgg tggaaaccaac 1320
 gcgcactgtg tcttcggtga g 1341

<210> 30
 <211> 447
 <212> PRT
 <213> *Schizochytrium* sp.
 <400> 30

ES 2 623 610 T3

Asp Asn Ile Ala Val Val Gly Met Ala Val Gln Tyr Ala Gly Cys Lys
1 5 10 15

Asn Gln Asp Glu Phe Trp Asp Thr Leu Met Arg Lys Glu Ile Asn Ser
20 25 30

Ser Pro Ile Ser Ala Glu Arg Leu Gly Thr Arg Tyr Arg Asp Leu His
35 40 45

Phe His Pro Gln Arg Ser Lys Tyr Ala Asp Thr Phe Cys Asn Asp Arg
50 55 60

Tyr Gly Cys Val Asp Ala Ser Val Asp Asn Glu His Asp Leu Leu Ala
65 70 75 80

ES 2 623 610 T3

Asp Leu Ala Arg Arg Ala Leu Leu Asp Ala Gly Ile Asn Leu Asp Asp
 85 90 95

Ala Ser Thr Thr Ala Asn Leu Arg Asp Phe Gly Ile Val Ser Gly Cys
 100 105 110

Leu Ser Phe Pro Met Asp Asn Leu Gln Gly Glu Leu Leu Asn Leu Tyr
 115 120 125

Gln Val His Val Glu Asn Arg Val Gly Ala Gln Arg Phe Arg Asp Ser
 130 135 140

Arg Pro Trp Ser Glu Arg Pro Arg Ala Val Ser Pro Glu Ala Ser Asp
 145 150 155 160

Pro Arg Val Tyr Ser Asp Pro Ala Ser Phe Val Ala Asn Gln Leu Gly
 165 170 175

Leu Gly Pro Val Arg Tyr Ser Leu Asp Ala Ala Cys Ala Ser Ala Leu
 180 185 190

Tyr Cys Leu Lys Leu Ala Ser Asp His Leu Leu Ser Arg Ser Ala Asp
 195 200 205

Val Met Leu Cys Gly Ala Thr Cys Phe Pro Asp Pro Phe Phe Ile Leu
 210 215 220

Ser Gly Phe Ser Thr Phe Gln Ala Met Pro Leu Gly Gly Pro Asp Asp
 225 230 235 240

Asn Pro Leu Ser Val Pro Leu Arg Gln Gly Ser Gln Gly Leu Thr Pro
 245 250 255

Gly Glu Gly Gly Ala Ile Met Val Leu Lys Arg Leu Glu Asp Ala Val
 260 265 270

Arg Asp Gly Asp Arg Ile Tyr Gly Thr Leu Leu Gly Thr Ser Leu Ser
 275 280 285

Asn Ala Gly Cys Gly Leu Pro Leu Ser Pro His Leu Pro Ser Glu Lys
 290 295 300

Ser Cys Met Glu Asp Leu Tyr Thr Ser Val Gly Ile Asp Pro Ser Glu
 305 310 315 320

Val Gln Tyr Val Glu Cys His Ala Thr Gly Thr Pro Gln Gly Asp Val
 325 330 335

ES 2 623 610 T3

Val Glu Val Glu Ala Leu Arg His Cys Phe Arg Gly Asn Thr Asp His
340 345 350

Pro Pro Arg Met Gly Ser Thr Lys Gly Asn Phe Gly His Thr Leu Val
355 360 365

Ala Ala Gly Phe Ala Gly Met Ala Lys Val Leu Leu Ser Met Gln His
370 375 380

Gly Thr Ile Pro Pro Thr Pro Gly Val Asp Arg Ser Asn Cys Ile Asp
385 390 395 400

Pro Leu Val Val Asp Glu Ala Ile Pro Trp Pro Tyr Ser Ser Ala Gln
405 410 415

Ala Arg Ala Gly Lys Pro Gly Asp Glu Leu Lys Cys Ala Ser Leu Ser
420 425 430

Ala Phe Gly Phe Gly Gly Thr Asn Ala His Cys Val Phe Arg Glu
435 440 445

<210> 31

<211> 1293

5 <212> ADN

<213> *Schizochytrium* sp.

<400> 31

```

ggaccgattg ccatcatcgg gatggacgcg acgttttgga coetcaaggg cctggacgcg      60
tttgagcagg ccatctacaa gggcacggac ggcgccagcg aactgccgag caagcgctgg      120
cggttcctgg gcgccgacac ggacttcttg accgccatgg gcctcgacgc cgtgccgcgc      180
gggtgctacg tgcgcgacgt ggacgtggac tacaagcggc tgcggtcgcc gatgatecct      240
gaggacgtcc tgcgcccga acagctgctg gcggtggcta cgatggaccg cgcgctgcag      300
gacgtggaa tggcgacggg aggcaaggtg gcggtgctgg tggggctcgg cacggacacc      360
gagctgtacc ggcaccgcgc gcgcgtgaca ctcaaggagc ggctcgaccg ggccgcgttc      420
tcgcccgagc aggtgcagga gatgatggac tacatcaacg actgcggcac ctcgacgtcg      480
tacacgtcgt acatcggcaa cctcgtggcc acgcgcgtgt cctcgcagtg gggttttacg      540
ggcccgteet tcaccgtcac cgaaggcgca aactcgggtet accgctgeet cgagctgggc      600
aagttcctgc tcgacacgca ccaggtggac gccgtcgtgg tggccggcgt cgacctctgt      660
gccaccgcgc agaaccctta cctcaaggcg cgccgctccg ccatcagccg acaggaccac      720
cctcgcgcca actttgaggg cagcgccgac gggtactttg ccggcgaggg cagcggcgcc      780
ctggtcctca agcgccaggc cgacgttggc tcagacgaca aggtctacgc cagtgtcgcg      840
ggcctcacgt gcgccgcgca gcccgctgaa gccgtgtcgc cgctactact ccaagtccac      900

aacgacgaca acgagaagag ggtggtggag atggtggagc tcgccgccga ctcggttcgc      960
catgcgccgc acttggccaa ctgcgccgtg agcgccgagt cgcagctgga gcaagtgtcc     1020
aagttgctcg cgcaccaggt gccgggctcg gtggccatcg gcagcgtgcg cgccaacgtg     1080
ggagacgtcg ggtacgcctc gggcgccgcg agcctcatca agacggcgct gtgcctccac     1140
aaccgctacc tcccggccaa cccgcagtgg gagcggccgg tggcgccggt ctccgaggcg     1200
ctgtttactt gcccgcgctc gcgtgcctgg ctgaagaacc cgggcgagtc gcgactggcg     1260
gctgtcgcca gtgcctccga gagcgggtcc tgc                                     1293

```

<210> 32

<211> 431

<212> PRT

<213> *Schizochytrium* sp.

<400> 32

5

ES 2 623 610 T3

Gly Pro Ile Ala Ile Ile Gly Met Asp Ala Thr Phe Gly Thr Leu Lys
1 5 10 15

Gly Leu Asp Ala Phe Glu Gln Ala Ile Tyr Lys Gly Thr Asp Gly Ala
20 25 30

Ser Asp Leu Pro Ser Lys Arg Trp Arg Phe Leu Gly Ala Asp Thr Asp
35 40 45

Phe Leu Thr Ala Met Gly Leu Asp Ala Val Pro Arg Gly Cys Tyr Val
50 55 60

Arg Asp Val Asp Val Asp Tyr Lys Arg Leu Arg Ser Pro Met Ile Pro
65 70 75 80

Glu Asp Val Leu Arg Pro Gln Gln Leu Leu Ala Val Ala Thr Met Asp
85 90 95

Arg Ala Leu Gln Asp Ala Gly Met Ala Thr Gly Gly Lys Val Ala Val
100 105 110

Leu Val Gly Leu Gly Thr Asp Thr Glu Leu Tyr Arg His Arg Ala Arg
115 120 125

Val Thr Leu Lys Glu Arg Leu Asp Pro Ala Ala Phe Ser Pro Glu Gln
130 135 140

Val Gln Glu Met Met Asp Tyr Ile Asn Asp Cys Gly Thr Ser Thr Ser
145 150 155 160

Tyr Thr Ser Tyr Ile Gly Asn Leu Val Ala Thr Arg Val Ser Ser Gln
165 170 175

Trp Gly Phe Thr Gly Pro Ser Phe Thr Val Thr Glu Gly Ala Asn Ser
 180 185 190

Val Tyr Arg Cys Leu Glu Leu Gly Lys Phe Leu Leu Asp Thr His Gln
 195 200 205

Val Asp Ala Val Val Val Ala Gly Val Asp Leu Cys Ala Thr Ala Glu
 210 215 220

Asn Leu Tyr Leu Lys Ala Arg Arg Ser Ala Ile Ser Arg Gln Asp His
 225 230 235 240

Pro Arg Ala Asn Phe Glu Ala Ser Ala Asp Gly Tyr Phe Ala Gly Glu
 245 250 255

Gly Ser Gly Ala Leu Val Leu Lys Arg Gln Ala Asp Val Gly Ser Asp
 260 265 270

Asp Lys Val Tyr Ala Ser Val Ala Gly Leu Thr Cys Ala Ala Gln Pro
 275 280 285

Ala Glu Ala Val Ser Pro Leu Leu Leu Gln Val His Asn Asp Asp Asn
 290 295 300

Glu Lys Arg Val Val Glu Met Val Glu Leu Ala Ala Asp Ser Gly Arg
 305 310 315 320

His Ala Pro His Leu Ala Asn Ser Pro Leu Ser Ala Glu Ser Gln Leu
 325 330 335

Glu Gln Val Ser Lys Leu Leu Ala His Gln Val Pro Gly Ser Val Ala
 340 345 350

Ile Gly Ser Val Arg Ala Asn Val Gly Asp Val Gly Tyr Ala Ser Gly
 355 360 365

Ala Ala Ser Leu Ile Lys Thr Ala Leu Cys Leu His Asn Arg Tyr Leu
 370 375 380

Pro Ala Asn Pro Gln Trp Glu Arg Pro Val Ala Pro Val Ser Glu Ala
 385 390 395 400

Leu Phe Thr Cys Pro Arg Ser Arg Ala Trp Leu Lys Asn Pro Gly Glu
 405 410 415

Ser Arg Leu Ala Ala Val Ala Ser Ala Ser Glu Ser Gly Ser Cys
 420 425 430

ES 2 623 610 T3

<210> 33
 <211> 1203
 <212> ADN
 <213> *Schizochytrium* sp.

5

<400> 33

```

acgccggaga agctggagaa ggagttggag ctggcagcca aggggtgtacc gcgaagcgcc      60
aaggccgggc gcaactggat gtcgccatcg ggcagcgccct ttgcgccgac acctgtgacc      120
agcgaccgcg tcgcgttcat gtacggcgag ggccgcagcc cctactacgg cgtcgggctc      180
gacctgcacc gcctgtggcc ggctttgcac gagcgcatca acgacaagac cgcggcgctg      240
tgggagaacg gcgactcgtg gtcctatgccg cgcgcggtgg atgccgactc gcagcgcgcc      300
gtgcagacgg cctttgacgc ggaccagatc gagatgttcc gcacgggcat cttegtgtcc      360
atctgectca ccgactacgc gcgcgacgtg ctcggggtgc agcccaagge gtgcttcggc      420
ctcagcctcg gcgagatctc catgctcttt gcgctgtcgc gacgcaactg cggcctgtcg      480
gaccagctca cgcagcgccct acgcacctcg ccggtgtggt cgacacagct ggcggtggag      540
ttccaggcct tgcgcaagct atggaacgtg ccggcgggacg ccccggtgga gtccttctgg      600
cagggctact tggttcgcg cagccgcgcc gaaatcgaga aggcgatcgg gcccgacaac      660
cgcttcgtgc gcctgctgat cgtcaacgac tcgagcagcg cgctgatcgc cggcaaacct      720
gccgagtgtc tgcgcgtgct ggagcgccctg ggccggcggt tgcgcgccat gcccgtaag      780
caaggcatga ttgggcactg ccccgaaagtg gcgcacctaca cgcggggcat cgcgcacatc      840
cacgagattt tggagattcc ggacagcccc gtcaagatgt acacctcggg caccaacgcc      900
gagctgcgcg ggggcagcaa cagcagcatc accgagttcg tgcagaagtt gtacacgcgc      960
atcgccgact ttccgggcat cgtcgacaag gtcagccgtg acggccacga tgtcttcgtc     1020
gaggtggggc cgaacaacat gcgctccgcc gcggtcagtg acattcttgg caaggctgcc     1080
accccgcatg tctccgtggc gctggaccgc cccagtgagt cggcgtggac gcagaccctc     1140
aagtcgctgg cgctgctgac cgcccaccgc gtgcccctgc acaaccgac tctgtttgcg     1200
gac                                                                                   1203
  
```

10

<210> 34
 <211> 401
 <212> PRT
 <213> *Schizochytrium* sp.

15

<400> 34

ES 2 623 610 T3

Thr Pro Glu Lys Leu Glu Lys Glu Leu Glu Leu Ala Ala Lys Gly Val
 1 5 10 15

Pro Arg Ser Ala Lys Ala Gly Arg Asn Trp Met Ser Pro Ser Gly Ser
 20 25 30

Ala Phe Ala Pro Thr Pro Val Thr Ser Asp Arg Val Ala Phe Met Tyr

ES 2 623 610 T3

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

ES 2 623 610 T3

Gly Ser Asn Ser Ser Ile Thr Glu Phe Val Gln Lys Leu Tyr Thr Arg
305 310 315 320

Ile Ala Asp Phe Pro Gly Ile Val Asp Lys Val Ser Arg Asp Gly His
325 330 335

Asp Val Phe Val Glu Val Gly Pro Asn Asn Met Arg Ser Ala Ala Val
340 345 350

Ser Asp Ile Leu Gly Lys Ala Ala Thr Pro His Val Ser Val Ala Leu
355 360 365

Asp Arg Pro Ser Glu Ser Ala Trp Thr Gln Thr Leu Lys Ser Leu Ala
370 375 380

Leu Leu Thr Ala His Arg Val Pro Leu His Asn Pro Thr Leu Phe Ala
385 390 395 400

Asp

<210> 35

<211> 1273

<212> ADN

<213> *Schizochytrium* sp.

<400> 35

```

tacgacgtgg actggcgct ctacatgggc gccatggcgg aaggcatctc gtcggtagac      60
ctggtggtcg ctgcgcgcga ggcgcgcgtg ctggcatcat tcggagcggc ccgcttgctt      120
atggaccagg tggaactcca gatccgtgag atccagcaac gcacctcaa cgcctttgct      180
gtcaacctga tgccgggtcc tgacgaggcc gcgacgggtg acgcgctgct gcgcacgggc      240
gtctcaatcg tcgaggcatc gggctacacc ggcgcgctct ctgcagacct ggtgcgctac      300
cgtgtcacgg gtctgcgacg aactagttgc ggtgcttctg tgtcggcgac tcacctgtg      360
gtcgccaagg tgtcgcgcac cgaggtggcc gagcactttc tgcgcccggc gccggccgcc      420
gtactagagg ctttggtcgc cgccaaacag attacgcccg agcaggccgc gctggccagc      480
cgcgtcgcca tggccgacga cgtcgcggtg gaggccgact cgggcgggca caccgacaac      540
cgaccgatcc acgtgctgct gccgctcgtg gtggcgcgag gcaaccgctg gcgccacctg      600
gtggacacgc cagtgcgcgt cggcgccggc ggcgggatcg cctgtccgcg cgcgcgctg      660
ctcgcctttt ccctgggcgc cgcctttgtg gtcaccgggt ccgtcaacca actggcccgc      720
gaggctggca ccagcgacgc ggtccgacta ctgctggcga cggccacctc ctcgacgtg      780
gccatggcgc cgggcggcgt ccaggtgctc aagaagcaga ccatgttcgc cgcgcgggcc      840
acgatgctcg ccagctgca ggccaagttc ggctcctttg acgccgtgcc ggagccgcag      900

ctgcgcaagc tcgagcgctc cgtgttcaag cagtccgtgg cggacgtgtg ggctgctgca      960
cgcgaaaagt ttggtgtcga cgctaccgct gcaagtccgc aggagaggat ggcgctctgt      1020
gtgcgctggt acatgtcgca gtcgtcgcga tgggctaccg aggcgacgtc cgcgcgcaag      1080
gcggactacc agatctggtg cggccccgcc atcggcagct tcaacgactt cgttcgcggc      1140
accaagctgg acgcgaccgc tggcaccggc gagtttccgc gcgtcgtgga catcaaccag      1200
cacatcctcc tcggagcctc gcactaccgc cgcgtgcagc aacaacaaca ggacgacgac      1260
gtagaataca tca                                                    1273

```

<210> 36

<211> 401

5 <212> PRT

<213> *Schizochytrium* sp.

<400> 36

ES 2 623 610 T3

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

ES 2 623 610 T3

His Thr Asp Asn Arg Pro Ile His Val Leu Leu Pro Leu Val Val Ala
 180 185 190
 Gln Arg Asn Arg Trp Arg His Leu Val Asp Thr Pro Val Arg Val Gly
 195 200 205
 Ala Gly Gly Gly Ile Ala Cys Pro Arg Ala Ala Leu Leu Ala Phe Ser
 210 215 220
 Leu Gly Ala Ala Phe Val Val Thr Gly Ser Val Asn Gln Leu Ala Arg
 225 230 235 240
 Glu Ala Gly Thr Ser Asp Ala Val Arg Leu Leu Leu Ala Thr Ala Thr
 245 250 255
 Tyr Ser Asp Val Ala Met Ala Pro Gly Gly Val Gln Val Leu Lys Lys
 260 265 270
 Gln Thr Met Phe Ala Ala Arg Ala Thr Met Leu Ala Gln Leu Gln Ala
 275 280 285
 Lys Phe Gly Ser Phe Asp Ala Val Pro Glu Pro Gln Leu Arg Lys Leu
 290 295 300
 Glu Arg Ser Val Phe Lys Gln Ser Val Ala Asp Val Trp Ala Ala Ala
 305 310 315 320
 Arg Glu Lys Phe Gly Val Asp Ala Thr Ala Ala Ser Pro Gln Glu Arg
 325 330 335
 Met Ala Leu Cys Val Arg Trp Tyr Met Ser Gln Ser Ser Arg Trp Ala
 340 345 350
 Thr Glu Ala Thr Ser Ala Arg Lys Ala Asp Tyr Gln Ile Trp Cys Gly
 355 360 365
 Pro Ala Ile Gly Ser Phe Asn Asp Phe Val Arg Gly Thr Lys Leu Asp
 370 375 380
 Ala Thr Ala Gly Thr Gly Glu Phe Pro Arg Val Val Asp Ile Asn Gln
 385 390 395 400

His

<210> 37
 <211> 1350
 <212> ADN

<213> *Schizochytrium* sp.

<400> 37

```

atgacatcat cgaagaagac tcccgtgtgg gagatgagca aggaggagct gctggacggc      60
aagacggtgg tcttcgacta caacgagctg ctccaattcg ccgagggcga cgtgggccaa    120
gtgttcggac ccgagttcga catcatcgac aagtaccggc gtcgcgtgcg gctgccggcg    180
cgcgagtacc tgctcgtgtc gcgcgtgacg ctgatggacg ccgaggtgaa caacttccgc    240
gtcgggtcgc gcatgggtgac cgagtagcgc gtgcccggtga acggggagct gtcggagggc    300
ggggacgtgc cgtggggcggg gctggtggag tccggggcagt gcgacctgat gctcatctcg    360
tacaatggga tcgacttcca gtgcaagggc gaccgcgtgt accgcctgct caacacatcg    420
ctcaccttct tcgggggtggc gcacgagggc gagacgctgg tgtacgacat ccgcgtcacg    480
gggttcgcca agggcgcggg cggggagatc tcgatgttct tcttcgagta cgactgcttc    540
gtggacggcc gcctgctgat cgagatgcgc gacgggtgcg ccgggttctt cacggacgcc    600
gagctggccg ccggcaaggg cgtgcttaag accaaggcgg agctggcggc gcgcgcgcag    660
atccagaagc aggacatcgc gccctttgcg ccggcgccgt gctcgacaaa gacctcgtg    720
gacgcgcgcg agatgcggct gctcgtggac cgccagtggg cgcgcgtctt cggcagcggc    780
atggcgggca tcgactacaa gttgtgcgct cgcaagatgc tcatgatcga ccgcgtcacg    840
cacctcgacc cgcgcggcgg cgcgcacggc ctccggctgc tgatcgggga gaagggtgctg    900
gagcgcgacc actggtactt cccctgccac tttgtgcgcg acgaggtgat ggcggggtcg    960
ctggtcagcg acggctgctc gcagctcctc aagggtgtaca tgctgtggct cggcctgcac   1020
acgaccgtgg gcgcgttcga ctttcgtccc gtgagcgggc acgccaacaa ggtgcgggtgc   1080
cgcgggcaga tctcaccgca caagggcaag ctctgtgtacg tgatggagat caaggaaatg   1140
ggctttgacg cgaagacggg cgatccgttt gcgatcgcgg acgtggacat catcgacgtc   1200
aacttcgagg agggacaggg gtttgcgga gtggaagacc tgcacagcta cggccagggc   1260
gacctccgca agaagatcgt cgtcgacttc aagggcacgc cgctctccct gcagaagcgg   1320
aaggagcagc agaaggaaag catgaccgtg                                     1350

```

5

<210> 38

<211> 450

<212> PRT

10 <213> *Schizochytrium* sp.

<400> 38

ES 2 623 610 T3

Met Thr Ser Ser Lys Lys Thr Pro Val Trp Glu Met Ser Lys Glu Glu
1 5 10 15

Leu Leu Asp Gly Lys Thr Val Val Phe Asp Tyr Asn Glu Leu Leu Glu
20 25 30

Phe Ala Glu Gly Asp Val Gly Gln Val Phe Gly Pro Glu Phe Asp Ile
35 40 45

ES 2 623 610 T3

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

Trp Tyr Phe Pro Cys His Phe Val Arg Asp Glu Val Met Ala Gly Ser
305 310 315 320

Leu Val Ser Asp Gly Cys Ser Gln Leu Leu Lys Val Tyr Met Leu Trp
325 330 335

Leu Gly Leu His Thr Thr Val Gly Ala Phe Asp Phe Arg Pro Val Ser
340 345 350

Gly His Ala Asn Lys Val Arg Cys Arg Gly Gln Ile Ser Pro His Lys
355 360 365

Gly Lys Leu Val Tyr Val Met Glu Ile Lys Glu Met Gly Phe Asp Ala
370 375 380

Lys Thr Gly Asp Pro Phe Ala Ile Ala Asp Val Asp Ile Ile Asp Val
385 390 395 400

Asn Phe Glu Glu Gly Gln Ala Phe Ala Gly Val Glu Asp Leu His Ser
405 410 415

Tyr Gly Gln Gly Asp Leu Arg Lys Lys Ile Val Val Asp Phe Lys Gly
420 425 430

Ile Ala Leu Ser Leu Gln Lys Arg Lys Glu Gln Gln Lys Glu Ser Met
435 440 445

Thr Val
450

<210> 39

<211> 1200

<212> ADN

<213> *Schizochytrium* sp.

<400> 39

taccgcgcgc	gggcgggtgtg	cttctcgccg	ttccccaaca	acccgcttga	caacgaccac	60
acgcgcggcc	agatgcggtt	gacctggttc	aacatgtccg	aattcatgtg	cggcaaagtg	120
tccaactgcc	tgggccccga	gtttgcgcgc	ttcgacgcga	gcaagacgag	ccgcagcccc	180
gcctttgacc	tggcgctcgt	gacgcgggtg	acgagcgtgg	cggacatgga	gcacggggccg	240
ttctacaacg	tggacgtcaa	cccgggccag	ggcacgatgg	tgggcgagtt	cgactgtccc	300
gcggacgcgt	ggtttcttcg	cgctcgcgac	cgcgacgacc	acatgccgta	ctcgatcctg	360
atggagatcg	cgtgcgagac	gtcggggcgtc	ctcacctcgg	tgctcaaggc	gccgctgacg	420
atggacaagg	acgacatcct	cttcgcgaac	ctcgacgcag	acgccgagct	cgtggggcgac	480
gccatgccgg	acgtgcgcgg	caagacgatc	cgcaacttca	ccaagtgcac	aggctacagc	540
atgctcggca	agatgggcat	ccaccgcttc	acctttgagc	tcagcgtcga	cggcgccgtc	600
ttctacaagg	gcagcacctc	gtttggctgg	ttcgtccccg	aggtcttcga	gtcgcagacc	660
ggtctcgaca	acggcaagcc	gcgcctgcct	tggtaaccgc	agaacaacgt	cgcctgcgac	720
acgctctccg	cgcccgccct	cgcttcctcc	gcgcaaggtc	agctgcagct	gcagcgacgc	780
gggtcgcagg	cgcagttcct	ggacacaatc	cacctggcgg	gcagcggcgc	cggcgtgcac	840
ggccagggct	acgcgcacgg	ggagaaggcc	gtgaacaagc	aagattggtt	cttctcgtgc	900
cacttctggt	tcgaccccg	gatgcccggg	tccttgggca	tcgagtcgat	gttccagctc	960
gtcgaggcgt	ggtgcgtgaa	gcagggactc	gcggcgcggc	acggcatcgc	tcacccagtg	1020
ttcgcgcacg	cgcccggggc	cacgagctgg	aagtaccgcg	ggcagctaac	ccccaagaac	1080
gaccgcatgg	acagcgaggt	gcacatcaag	tcggtggcgg	ccttctcctc	ctgggtcgac	1140
gtcgtcgcgg	acgggttcct	cttcgtcgac	ggcctccgcg	tctactcggc	agacaacctc	1200

<210> 40
 <211> 400
 <212> PRT
 <213> *Schizochytrium* sp.

<400> 40

ES 2 623 610 T3

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

ES 2 623 610 T3

Asp 145	Ile	Leu	Phe	Arg	Asn 150	Leu	Asp	Ala	Asp	Ala 155	Glu	Leu	Val	Gly	Asp 160
Ala	Met	Pro	Asp	Val 165	Arg	Gly	Lys	Thr	Ile 170	Arg	Asn	Phe	Thr	Lys 175	Cys
Thr	Gly	Tyr	Ser 180	Met	Leu	Gly	Lys	Met 185	Gly	Ile	His	Arg	Phe 190	Thr	Phe
Glu	Leu	Ser 195	Val	Asp	Gly	Ala	Val 200	Phe	Tyr	Lys	Gly	Ser 205	Thr	Ser	Phe
Gly 210	Trp	Phe	Val	Pro	Glu	Val 215	Phe	Glu	Ser	Gln	Thr 220	Gly	Leu	Asp	Asn
Gly 225	Lys	Pro	Arg	Leu	Pro 230	Trp	Tyr	Arg	Glu	Asn 235	Asn	Val	Ala	Val	Asp 240
Thr	Leu	Ser	Ala 245	Pro	Ala	Ser	Ala	Ser	Ser 250	Ala	Gln	Gly	Gln	Leu 255	Gln
Leu	Gln	Arg	Arg 260	Gly	Ser	Gln	Ala	Gln 265	Phe	Leu	Asp	Thr	Ile 270	His	Leu
Ala	Gly	Ser 275	Gly	Ala	Gly	Val	His 280	Gly	Gln	Gly	Tyr	Ala 285	His	Gly	Glu
Lys 290	Ala	Val	Asn	Lys	Gln	Asp 295	Trp	Phe	Phe	Ser	Cys 300	His	Phe	Trp	Phe
Asp 305	Pro	Val	Met	Pro	Gly 310	Ser	Leu	Gly	Ile	Glu 315	Ser	Met	Phe	Gln	Leu 320
Val	Glu	Ala	Trp	Cys 325	Val	Lys	Gln	Gly	Leu 330	Ala	Ala	Arg	His	Gly 335	Ile
Ala	His	Pro	Val 340	Phe	Ala	His	Ala	Pro 345	Gly	Ala	Thr	Ser	Trp 350	Lys	Tyr
Arg	Gly	Gln 355	Leu	Thr	Pro	Lys	Asn 360	Asp	Arg	Met	Asp	Ser 365	Glu	Val	His
Ile 370	Lys	Ser	Val	Ala	Ala	Phe 375	Ser	Ser	Trp	Val	Asp 380	Val	Val	Ala	Asp
Gly 385	Phe	Leu	Phe	Val	Asp 390	Gly	Leu	Arg	Val	Tyr 395	Ser	Ala	Asp	Asn	Leu 400

ES 2 623 610 T3

<210> 41
 <211> 1353
 <212> ADN
 <213> *Schizochytrium* sp.

5

<400> 41

```

cagctggacg cggggagcga ggtgcccgcc tgcgcggtga gcgacctggg cgataggggc      60
ttcatggaga cgtacgggggt ggtggcgccg ctgtacagcg gggcgatggc caagggcatc      120
gcgtcggcgg acctgggtgat cgcgatgggc cagcgcaaga tgctggggtc gtttggcgcg      180
ggcgggctcc cgatgcacgt cgtgcgcgcg gggattgaga agatccaggc agcgtgcca      240
gcggggccat acgcgggtcaa cctgattcac tgccttttg acgccaacct ggagaagggc      300
aacgtggacc tcttcctgga gaaggcgctg cgcgtcgtgg aggcgtcggc ctcatggag      360
ctcacgcccc aggtggtgcg ctaccgcgcg acgggcctct ctgcgcacgc gcgcggcggc      420
tccgtgcgca cggcccacaa gatcatcggc aaggtcagcc gcaccgagct ggccgagatg      480
tttatcgggc ccgcgcgcga agccattctc gacaagcttg tggcgtccgg cgagatcacc      540
cccgagcagg cggcgcctggc gctcgagggtg cccatggcgg acgacatcgc cgtcgaggcc      600
gattcggggc ggacacccga caaccgcccc atccacgtca tcttgcacct catctcagc      660
ctgcgcaacc gcctccagcg cgagctcaag taccctgcgc gacaccgcgt gcgcgtcggc      720
gccggggggc gcatcgggtg cccgcaagcg gctctgggcg ccttccacat gggcgccgcg      780
tttgtggtga cgggcacggt caaccagctg agccggcagg ccgggacatg cgacaatgtg      840
cggcggcagc tgtcgcgcgc gacgtactcg gacatcacga tggcgcgggc ggcgacatg      900
ttcgagcagg gcgtcgagct gcagggtgctc aagaaggga cgatgtttcc ctgcgcgcgc      960
aagaagctgt tcgagctgtt tcacaagtac gactcgttcg aggcgatgcc ggcgacgag      1020
ctggcgcgcg tcgagaagcg catcttcagc aagtcactcg ccgaggtgtg ggccgagacc      1080
aaggacttct acatcacgcg gctcaacaac ccggagaaga tccgcaaggc ggagaacgag      1140
gaccccaagc tcaagatgtc actctgcttc cgctgggtacc tcgggctcag ctogttctgg      1200
gccaacaacg gcatcgcgga ccgcacgatg gactaccaga tctggtgcgg ccctgccatc      1260
ggcgccctca acgacttcat cgcgcgactcg tacctcgacg tggccgtctc gggcgagttc      1320
cccgacgtcg tgcagatcaa cctgcagatc ctg      1353
  
```

10 <210> 42
 <211> 451
 <212> PRT
 <213> *Schizochytrium* sp.

15 <400> 42

ES 2 623 610 T3

Gln Leu Asp Ala Gly Ser Glu Val Pro Ala Cys Ala Val Ser Asp Leu
 1 5 10 15

Gly Asp Arg Gly Phe Met Glu Thr Tyr Gly Val Val Ala Pro Leu Tyr

ES 2 623 610 T3

20	25	30
Ser Gly Ala Met Ala Lys Gly Ile Ala Ser Ala Asp Leu Val Ile Ala		
35	40	45
Met Gly Gln Arg Lys Met Leu Gly Ser Phe Gly Ala Gly Gly Leu Pro		
50	55	60
Met His Val Val Arg Ala Gly Ile Glu Lys Ile Gln Ala Ala Leu Pro		
65	70	75
Ala Gly Pro Tyr Ala Val Asn Leu Ile His Ser Pro Phe Asp Ala Asn		
85	90	95
Leu Glu Lys Gly Asn Val Asp Leu Phe Leu Glu Lys Gly Val Arg Val		
100	105	110
Val Glu Ala Ser Ala Phe Met Glu Leu Thr Pro Gln Val Val Arg Tyr		
115	120	125
Arg Ala Thr Gly Leu Ser Arg Asp Ala Arg Gly Gly Ser Val Arg Thr		
130	135	140
Ala His Lys Ile Ile Gly Lys Val Ser Arg Thr Glu Leu Ala Glu Met		
145	150	155
Phe Ile Arg Pro Ala Pro Gln Ala Ile Leu Asp Lys Leu Val Ala Ser		
165	170	175
Gly Glu Ile Thr Pro Glu Gln Ala Ala Leu Ala Leu Glu Val Pro Met		
180	185	190
Ala Asp Asp Ile Ala Val Glu Ala Asp Ser Gly Gly His Thr Asp Asn		
195	200	205
Arg Pro Ile His Val Ile Leu Pro Leu Ile Leu Ser Leu Arg Asn Arg		
210	215	220
Leu Gln Arg Glu Leu Lys Tyr Pro Ala Arg His Arg Val Arg Val Gly		
225	230	235
Ala Gly Gly Gly Ile Gly Cys Pro Gln Ala Ala Leu Gly Ala Phe His		
245	250	255
Met Gly Ala Ala Phe Val Val Thr Gly Thr Val Asn Gln Leu Ser Arg		
260	265	270
Gln Ala Gly Thr Cys Asp Asn Val Arg Arg Gln Leu Ser Arg Ala Thr		
275	280	285

Tyr Ser Asp Ile Thr Met Ala Pro Ala Ala Asp Met Phe Glu Gln Gly
290 295 300

Val Glu Leu Gln Val Leu Lys Lys Gly Thr Met Phe Pro Ser Arg Ala
305 310 315 320

Lys Lys Leu Phe Glu Leu Phe His Lys Tyr Asp Ser Phe Glu Ala Met
325 330 335

Pro Ala Asp Glu Leu Ala Arg Val Glu Lys Arg Ile Phe Ser Lys Ser
340 345 350

Leu Ala Glu Val Trp Ala Glu Thr Lys Asp Phe Tyr Ile Thr Arg Leu
355 360 365

Asn Asn Pro Glu Lys Ile Arg Lys Ala Glu Asn Glu Asp Pro Lys Leu
370 375 380

Lys Met Ser Leu Cys Phe Arg Trp Tyr Leu Gly Leu Ser Ser Phe Trp
385 390 395 400

Ala Asn Asn Gly Ile Ala Asp Arg Thr Met Asp Tyr Gln Ile Trp Cys
405 410 415

Gly Pro Ala Ile Gly Ala Phe Asn Asp Phe Ile Ala Asp Ser Tyr Leu
420 425 430

Asp Val Ala Val Ser Gly Glu Phe Pro Asp Val Val Gln Ile Asn Leu
435 440 445

Gln Ile Leu
450

<210> 43

<211>4

5 <212> PRT

<213> *Schizochytrium* sp.

<220>

10 <221> misc_feature

<222> (2)..(2)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 43

15

Asp Xaa Ala Cys
1

<210> 44

<211>4

<212> PRT
 <213> *Schizochytrium* sp.
 <400> 44
 5
 Gly Phe Gly Gly
 1
 <210> 45
 <211> 12
 10 <212> ADN
 <213> *Schizochytrium* sp.
 <400> 45
 gggtttggcg gt 12
 15
 <210> 46
 <211> 5
 <212> PRT
 <213> *Schizochytrium* sp.
 20
 <400> 46
 Gly Phe Ser Leu Gly
 1 5
 25 <210> 47
 <211> 5
 <212> PRT
 <213> *Schizochytrium* sp.
 30 <400> 47
 Leu Gly Ile Asp Ser
 1 5
 35 <210> 48
 <211> 903
 <212> ADN
 <213> *Schizochytrium* sp.
 <400> 48
 40

ES 2 623 610 T3

```

ctgtcagacc cgcacgtggc cgtgcgcgag tctggttact ccgcctcggg ccagcgcctgc 60
acgacaaacta cgaagtcgct gactacgggc aagccgcacc agccgatctc ctcgctggac 120
ctctttcttg tgcggggcgg cgcgcgcggc atcacccgc tgtgcgtgcg cgagctggcg 180
cagcgcgtgg gggcgggcac gtacgtgctc atcggccgct cggagctgcc cagcagggag 240
cctgcctggg cggtcggcgt ggagtctggc aagccgctgg agaaggccgc gctggcgctc 300
ctgaaggcgg agtttgcagc gggccgcggg gccaaagcca cgcgatgct gcacaagaag 360
ctcgtagggc cgtggtcgg agcgcgcgag gtgcgagcct cgctcgcca gatcaactga 420
cagggcgcca cggctgtgta cgagtcgtgc gacgtgagct ctgccgcaa ggtgcgtgag 480
atggtagagc gcgtgcagca gcagggcggg cggcgcggtg cgggcgtgtt ccacgcgtcg 540
ggcgtgctgc gcgacaagct cgtggagaac aagtcgctgg cggacttcag cgcctgtac 600

gacaccaagg tgggcggcct catcaacctg ctggcctgcg tggacctggc gcagctgcgt 660
cacctcgtgc tcttcagctc gtcgcggggc ttccacggca acgtcgggca gtcggactac 720
gcaatggcca acgagggcgt caacaagctg gcggcgccacc tgtcggcggt gcaccgcgag 780
ctgtgcgcgc gctcgatctg cttcggaccg tgggacggcg gcatggtgac ccccgcgctc 840
aaggccaact tcatecgcat gggcatocag atcatccgc gccaaaggcg cgcgcagacc 900
gtc 903

```

<210> 49
 <211> 301
 <212> PRT
 <213> *Schizochytrium* sp.
 <400> 49

5

ES 2 623 610 T3

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

Phe His Ala Ser Gly Val Leu Arg Asp Lys Leu Val Glu Asn Lys Ser
180 185 190

Leu Ala Asp Phe Ser Ala Val Tyr Asp Thr Lys Val Gly Gly Leu Ile
195 200 205

Asn Leu Leu Ala Cys Val Asp Leu Ala Gln Leu Arg His Leu Val Leu
210 215 220

Phe Ser Ser Leu Ala Gly Phe His Gly Asn Val Gly Gln Ser Asp Tyr
225 230 235 240

Ala Met Ala Asn Glu Ala Leu Asn Lys Leu Ala Ala His Leu Ser Ala
245 250 255

Val His Pro Gln Leu Cys Ala Arg Ser Ile Cys Phe Gly Pro Trp Asp
260 265 270

Gly Gly Met Val Thr Pro Ala Leu Lys Ala Asn Phe Ile Arg Met Gly
275 280 285

Ile Gln Ile Ile Pro Arg Gln Gly Gly Ala Gln Thr Val
290 295 300

<210> 50
<211> 13
<212> PRT
5 <213> *Schizochytrium* sp.

<220>
<221> misc_feature
10 <222> (2)..(3)
<223> Xaa puede ser cualquier aminoácido de origen natural

<220>
<221> misc_feature
15 <222> (5)..(7)
<223> Xaa puede ser cualquier aminoácido de origen natural

<220>
<221> misc_feature
20 <222> (9)..(12)
<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 50

Leu Xaa Xaa His Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Pro
1 5 10

25 <210> 51
<211> 12
<212> ADN
30 <213> *Schizochytrium* sp.

<400>51
ggctttggtg ga 12

5 <210> 52
<211>5
<212> PRT
<213> *Schizochytrium* sp.

10 <220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa puede ser cualquier aminoácido de origen natural

15 <220>
<221> misc_feature
<222> (4)..(4)
<223> Xaa puede ser cualquier aminoácido de origen natural

20 <400> 52

Gly Xaa Ser Xaa Gly
1 5

25 <210> 53
<211>5
<212> PRT
<213> *Schizochytrium* sp.

30 <220>
<221> misc_feature
<222> (2)..(3)
<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 53

Phe Xaa Xaa His Phe
1 5

35 <210> 54
<211> 2910
<212> PRT
<213> *Schizochytrium* sp.

40 <400> 54

Met Ala Ala Arg Leu Gln Glu Gln Lys Gly Gly Glu Met Asp Thr Arg
1 5 10 15

Ile Ala Ile Ile Gly Met Ser Ala Ile Leu Pro Cys Gly Thr Thr Val
20 25 30

Arg Glu Ser Trp Glu Thr Ile Arg Ala Gly Ile Asp Cys Leu Ser Asp
35 40 45

Leu Pro Glu Asp Arg Val Asp Val Thr Ala Tyr Phe Asp Pro Val Lys
50 55 60

ES 2 623 610 T3

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

325	330	335
Ala Cys Val Asp Pro Ala Thr Val Thr Leu Val Glu Gly His Gly Thr 340 345 350		
Gly Thr Pro Val Gly Asp Arg Ile Glu Leu Thr Ala Leu Arg Asn Leu 355 360 365		
Phe Asp Lys Ala Tyr Gly Glu Gly Asn Thr Glu Lys Val Ala Val Gly 370 375 380		
Ser Ile Lys Ser Ser Ile Gly His Leu Lys Ala Val Ala Gly Leu Ala 385 390 395 400		
Gly Met Ile Lys Val Ile Met Ala Leu Lys His Lys Thr Leu Pro Gly 405 410 415		
Thr Ile Asn Val Asp Asn Pro Pro Asn Leu Tyr Asp Asn Thr Pro Ile 420 425 430		
Asn Glu Ser Ser Leu Tyr Ile Asn Thr Met Asn Arg Pro Trp Phe Pro 435 440 445		
Pro Pro Gly Val Pro Arg Arg Ala Gly Ile Ser Ser Phe Gly Phe Gly 450 455 460		
Gly Ala Asn Tyr His Ala Val Leu Glu Glu Ala Glu Pro Glu His Thr 465 470 475 480		
Thr Ala Tyr Arg Leu Asn Lys Arg Pro Gln Pro Val Leu Met Met Ala 485 490 495		
Ala Thr Pro Ala Ala Leu Gln Ser Leu Cys Glu Ala Gln Leu Lys Glu 500 505 510		
Phe Glu Ala Ala Ile Lys Glu Asn Glu Thr Val Lys Asn Thr Ala Tyr 515 520 525		
Ile Lys Cys Val Lys Phe Gly Glu Gln Phe Lys Phe Pro Gly Ser Ile 530 535 540		
Pro Ala Thr Asn Ala Arg Leu Gly Phe Leu Val Lys Asp Ala Glu Asp 545 550 555 560		
Ala Cys Ser Thr Leu Arg Ala Ile Cys Ala Gln Phe Ala Lys Asp Val 565 570 575		
Thr Lys Glu Ala Trp Arg Leu Pro Arg Glu Gly Val Ser Phe Arg Ala 580 585 590		

Lys Gly Ile Ala Thr Asn Gly Ala Val Ala Ala Leu Phe Ser Gly Gln
 595 600 605
 Gly Ala Gln Tyr Thr His Met Phe Ser Glu Val Ala Met Asn Trp Pro
 610 615 620
 Gln Phe Arg Gln Ser Ile Ala Ala Met Asp Ala Ala Gln Ser Lys Val
 625 630 635 640
 Ala Gly Ser Asp Lys Asp Phe Glu Arg Val Ser Gln Val Leu Tyr Pro
 645 650 655
 Arg Lys Pro Tyr Glu Arg Glu Pro Glu Gln Asn Pro Lys Lys Ile Ser
 660 665 670
 Leu Thr Ala Tyr Ser Gln Pro Ser Thr Leu Ala Cys Ala Leu Gly Ala
 675 680 685
 Phe Glu Ile Phe Lys Glu Ala Gly Phe Thr Pro Asp Phe Ala Ala Gly
 690 695 700
 His Ser Leu Gly Glu Phe Ala Ala Leu Tyr Ala Ala Gly Cys Val Asp
 705 710 715 720
 Arg Asp Glu Leu Phe Glu Leu Val Cys Arg Arg Ala Arg Ile Met Gly
 725 730 735
 Gly Lys Asp Ala Pro Ala Thr Pro Lys Gly Cys Met Ala Ala Val Ile
 740 745 750
 Gly Pro Asn Ala Glu Asn Ile Lys Val Gln Ala Ala Asn Val Trp Leu
 755 760 765
 Gly Asn Ser Asn Ser Pro Ser Gln Thr Val Ile Thr Gly Ser Val Glu
 770 775 780
 Gly Ile Gln Ala Glu Ser Ala Arg Leu Gln Lys Glu Gly Phe Arg Val
 785 790 795 800
 Val Pro Leu Ala Cys Glu Ser Ala Phe His Ser Pro Gln Met Glu Asn
 805 810 815
 Ala Ser Ser Ala Phe Lys Asp Val Ile Ser Lys Val Ser Phe Arg Thr
 820 825 830
 Pro Lys Ala Glu Thr Lys Leu Phe Ser Asn Val Ser Gly Glu Thr Tyr
 835 840 845

Pro Thr Asp Ala Arg Glu Met Leu Thr Gln His Met Thr Ser Ser Val
850 855 860

Lys Phe Leu Thr Gln Val Arg Asn Met His Gln Ala Gly Ala Arg Ile
865 870 875 880

Phe Val Glu Phe Gly Pro Lys Gln Val Leu Ser Lys Leu Val Ser Glu
885 890 895

Thr Leu Lys Asp Asp Pro Ser Val Val Thr Val Ser Val Asn Pro Ala
900 905 910

Ser Gly Thr Asp Ser Asp Ile Gln Leu Arg Asp Ala Ala Val Gln Leu
915 920 925

Val Val Ala Gly Val Asn Leu Gln Gly Phe Asp Lys Trp Asp Ala Pro
930 935 940

Asp Ala Thr Arg Met Gln Ala Ile Lys Lys Lys Arg Thr Thr Leu Arg
945 950 955 960

Leu Ser Ala Ala Thr Tyr Val Ser Asp Lys Thr Lys Lys Val Arg Asp
965 970 975

Ala Ala Met Asn Asp Gly Arg Cys Val Thr Tyr Leu Lys Gly Ala Ala
980 985 990

Pro Leu Ile Lys Ala Pro Glu Pro Val Val Asp Glu Ala Ala Lys Arg
995 1000 1005

Glu Ala Glu Arg Leu Gln Lys Glu Leu Gln Asp Ala Gln Arg Gln
1010 1015 1020

Leu Asp Asp Ala Lys Arg Ala Ala Ala Glu Ala Asn Ser Lys Leu
1025 1030 1035

Ala Ala Ala Lys Glu Glu Ala Lys Thr Ala Ala Ala Ser Ala Lys
1040 1045 1050

Pro Ala Val Asp Thr Ala Val Val Glu Lys His Arg Ala Ile Leu
1055 1060 1065

Lys Ser Met Leu Ala Glu Leu Asp Gly Tyr Gly Ser Val Asp Ala
1070 1075 1080

Ser Ser Leu Gln Gln Gln Gln Gln Gln Gln Thr Ala Pro Ala Pro
1085 1090 1095

Val	Lys	Ala	Ala	Ala	Pro	Ala	Ala	Pro	Val	Ala	Ser	Ala	Pro	Ala
	1100					1105					1110			
Pro	Ala	Val	Ser	Asn	Glu	Leu	Leu	Glu	Lys	Ala	Glu	Thr	Val	Val
	1115					1120					1125			
Met	Glu	Val	Leu	Ala	Ala	Lys	Thr	Gly	Tyr	Glu	Thr	Asp	Met	Ile
	1130					1135					1140			
Glu	Ala	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile
	1145					1150					1155			
Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Met	Leu	Asn	Val
	1160					1165					1170			
Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly
	1175					1180					1185			
Glu	Val	Val	Asn	Ala	Met	Lys	Ala	Glu	Ile	Ala	Gly	Ser	Ser	Ala
	1190					1195					1200			
Pro	Ala	Pro	Ala	Ala	Ala	Ala	Pro	Ala	Pro	Ala	Lys	Ala	Ala	Pro
	1205					1210					1215			
Ala	Ala	Ala	Ala	Pro	Ala	Val	Ser	Asn	Glu	Leu	Leu	Glu	Lys	Ala
	1220					1225					1230			
Glu	Thr	Val	Val	Met	Glu	Val	Leu	Ala	Ala	Lys	Thr	Gly	Tyr	Glu
	1235					1240					1245			
Thr	Asp	Met	Ile	Glu	Ser	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly
	1250					1255					1260			
Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala
	1265					1270					1275			
Met	Leu	Asn	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr
	1280					1285					1290			
Arg	Thr	Val	Gly	Glu	Val	Val	Asn	Ala	Met	Lys	Ala	Glu	Ile	Ala
	1295					1300					1305			
Gly	Gly	Ser	Ala	Pro	Ala	Pro	Ala	Ala	Ala	Ala	Pro	Gly	Pro	Ala
	1310					1315					1320			
Ala	Ala	Ala	Pro	Ala	Pro	Ala	Ala	Ala	Ala	Pro	Ala	Val	Ser	Asn
	1325					1330					1335			
Glu	Leu	Leu	Glu	Lys	Ala	Glu	Thr	Val	Val	Met	Glu	Val	Leu	Ala

ES 2 623 610 T3

1340	1345	1350
Ala Lys Thr Gly Tyr Glu Thr	Asp Met Ile Glu Ser	Asp Met Glu
1355	1360	1365
Leu Glu Thr Glu Leu Gly Ile	Asp Ser Ile Lys Arg	Val Glu Ile
1370	1375	1380
Leu Ser Glu Val Gln Ala Met	Leu Asn Val Glu Ala	Lys Asp Val
1385	1390	1395
Asp Ala Leu Ser Arg Thr Arg	Thr Val Gly Glu Val	Val Asp Ala
1400	1405	1410
Met Lys Ala Glu Ile Ala Gly	Gly Ser Ala Pro Ala	Pro Ala Ala
1415	1420	1425
Ala Ala Pro Ala Pro Ala Ala	Ala Ala Pro Ala Pro	Ala Ala Pro
1430	1435	1440
Ala Pro Ala Val Ser Ser Glu	Leu Leu Glu Lys Ala	Glu Thr Val
1445	1450	1455
Val Met Glu Val Leu Ala Ala	Lys Thr Gly Tyr Glu	Thr Asp Met
1460	1465	1470
Ile Glu Ser Asp Met Glu Leu	Glu Thr Glu Leu Gly	Ile Asp Ser
1475	1480	1485
Ile Lys Arg Val Glu Ile Leu	Ser Glu Val Gln Ala	Met Leu Asn
1490	1495	1500
Val Glu Ala Lys Asp Val Asp	Ala Leu Ser Arg Thr	Arg Thr Val
1505	1510	1515
Gly Glu Val Val Asp Ala Met	Lys Ala Glu Ile Ala	Gly Gly Ser
1520	1525	1530
Ala Pro Ala Pro Ala Ala Ala	Ala Pro Ala Pro Ala	Ala Ala Ala
1535	1540	1545
Pro Ala Pro Ala Ala Pro Ala	Pro Ala Ala Pro Ala	Pro Ala Val
1550	1555	1560
Ser Ser Glu Leu Leu Glu Lys	Ala Glu Thr Val Val	Met Glu Val
1565	1570	1575
Leu Ala Ala Lys Thr Gly Tyr	Glu Thr Asp Met Ile	Glu Ser Asp
1580	1585	1590

Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val
1595						1600					1605			
Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Met	Leu	Asn	Val	Glu	Ala	Lys
1610						1615					1620			
Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Val
1625						1630					1635			
Asp	Ala	Met	Lys	Ala	Glu	Ile	Ala	Gly	Ser	Ser	Ala	Ser	Ala	Pro
1640						1645					1650			
Ala	Ala	Ala	Ala	Pro	Ala	Pro	Ala	Ala	Ala	Ala	Pro	Ala	Pro	Ala
1655						1660					1665			
Ala	Ala	Ala	Pro	Ala	Val	Ser	Asn	Glu	Leu	Leu	Glu	Lys	Ala	Glu
1670						1675					1680			
Thr	Val	Val	Met	Glu	Val	Leu	Ala	Ala	Lys	Thr	Gly	Tyr	Glu	Thr
1685						1690					1695			
Asp	Met	Ile	Glu	Ser	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile
1700						1705					1710			
Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Met
1715						1720					1725			
Leu	Asn	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg
1730						1735					1740			
Thr	Val	Gly	Glu	Val	Val	Asp	Ala	Met	Lys	Ala	Glu	Ile	Ala	Gly
1745						1750					1755			
Gly	Ser	Ala	Pro	Ala	Pro	Ala	Ala	Ala	Ala	Pro	Ala	Pro	Ala	Ala
1760						1765					1770			
Ala	Ala	Pro	Ala	Val	Ser	Asn	Glu	Leu	Leu	Glu	Lys	Ala	Glu	Thr
1775						1780					1785			
Val	Val	Met	Glu	Val	Leu	Ala	Ala	Lys	Thr	Gly	Tyr	Glu	Thr	Asp
1790						1795					1800			
Met	Ile	Glu	Ser	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp
1805						1810					1815			
Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Met	Leu
1820						1825					1830			

ES 2 623 610 T3

Asn Val 1835	Glu Ala Lys Asp Val 1840	Asp Ala Leu Ser Arg Thr Arg Thr 1845
Val Gly 1850	Glu Val Val Asp Ala 1855	Met Lys Ala Glu Ile Ala Gly Ser 1860
Ser Ala 1865	Pro Ala Pro Ala Ala 1870	Ala Ala Pro Ala Pro Ala Ala Ala 1875
Ala Pro 1880	Ala Pro Ala Ala Ala 1885	Ala Pro Ala Val Ser Ser Glu Leu 1890
Leu Glu 1895	Lys Ala Glu Thr Val 1900	Val Met Glu Val Leu Ala Ala Lys 1905
Thr Gly 1910	Tyr Glu Thr Asp Met 1915	Ile Glu Ser Asp Met Glu Leu Glu 1920
Thr Glu 1925	Leu Gly Ile Asp Ser 1930	Ile Lys Arg Val Glu Ile Leu Ser 1935
Glu Val 1940	Gln Ala Met Leu Asn 1945	Val Glu Ala Lys Asp Val Asp Ala 1950
Leu Ser 1955	Arg Thr Arg Thr Val 1960	Gly Glu Val Val Asp Ala Met Lys 1965
Ala Glu 1970	Ile Ala Gly Gly Ser 1975	Ala Pro Ala Pro Ala Ala Ala Ala 1980
Pro Ala 1985	Pro Ala Ala Ala Ala 1990	Pro Ala Val Ser Asn Glu Leu Leu 1995
Glu Lys 2000	Ala Glu Thr Val Val 2005	Met Glu Val Leu Ala Ala Lys Thr 2010
Gly Tyr 2015	Glu Thr Asp Met Ile 2020	Glu Ser Asp Met Glu Leu Glu Thr 2025
Glu Leu 2030	Gly Ile Asp Ser Ile 2035	Lys Arg Val Glu Ile Leu Ser Glu 2040
Val Gln 2045	Ala Met Leu Asn Val 2050	Glu Ala Lys Asp Val Asp Ala Leu 2055
Ser Arg 2060	Thr Arg Thr Val Gly 2065	Glu Val Val Asp Ala Met Lys Ala 2070

Glu Ile	Ala Gly Gly Ser	Ala	Pro Ala Pro Ala	Ala	Ala Ala Pro	
2075		2080		2085		
Ala Ser	Ala Gly Ala Ala	Pro	Ala Val Lys Ile	Asp	Ser Val His	
2090		2095		2100		
Gly Ala	Asp Cys Asp Asp	Leu	Ser Leu Met His	Ala	Lys Val Val	
2105		2110		2115		
Asp Ile	Arg Arg Pro Asp	Glu	Leu Ile Leu Glu	Arg	Pro Glu Asn	
2120		2125		2130		
Arg Pro	Val Leu Val Val	Asp	Asp Gly Ser Glu	Leu	Thr Leu Ala	
2135		2140		2145		
Leu Val	Arg Val Leu Gly	Ala	Cys Ala Val Val	Leu	Thr Phe Glu	
2150		2155		2160		
Gly Leu	Gln Leu Ala Gln	Arg	Ala Gly Ala Ala	Ala	Ile Arg His	
2165		2170		2175		
Lys Ala	Ile Lys Glu Ala	Glu	Gln Arg Phe Gly	Ala	Leu Gly Gly	
2180		2185		2190		
Phe Ile	Ser Gln Gln Ala	Glu	Arg Phe Glu Pro	Ala	Glu Ile Leu	
2195		2200		2205		
Gly Phe	Thr Leu Met Cys	Ala	Lys Phe Ala Lys	Ala	Ser Leu Cys	
2210		2215		2220		
Thr Ala	Val Ala Gly Val	Leu	Ala Lys Asp Leu	Ser	Ala Glu Ser	
2225		2230		2235		
Ala Glu	Gly Arg Pro Ala	Phe	Ile Gly Val Ala	Arg	Leu Asp Gly	
2240		2245		2250		
Arg Leu	Gly Phe Thr Ser	Gln	Gly Thr Ser Asp	Ala	Leu Lys Arg	
2255		2260		2265		
Ala Gln	Arg Gly Ala Ile	Phe	Gly Leu Cys Lys	Thr	Ile Gly Leu	
2270		2275		2280		
Glu Trp	Ser Glu Ser Asp	Val	Phe Ser Arg Gly	Val	Asp Ile Ala	
2285		2290		2295		
Gln Gly	Met His Pro Glu	Asp	Ala Ala Val Ala	Ile	Val Arg Glu	
2300		2305		2310		
Met Ala	Cys Ala Asp Ile	Arg	Ile Arg Glu Val	Gly	Ile Gly Ala	

ES 2 623 610 T3

2315	2320	2325
Asn Gln Gln Arg Cys Thr Ile 2330	Arg Ala Ala Lys 2335	Leu Glu Thr Gly 2340
Asn Pro Gln Arg Gln Ile Ala 2345	Lys Asp Asp Val 2350	Leu Leu Val Ser 2355
Gly Gly Ala Arg Gly Ile Thr 2360	Pro Leu Cys Ile 2365	Arg Glu Ile Thr 2370
Arg Gln Ile Ala Gly Gly Lys 2375	Tyr Ile Leu Leu 2380	Gly Arg Ser Lys 2385
Val Ser Ala Ser Glu Pro Ala 2390	Trp Cys Ala Gly 2395	Ile Thr Asp Glu 2400
Lys Ala Val Gln Lys Ala Ala 2405	Thr Gln Glu Leu 2410	Lys Arg Ala Phe 2415
Ser Ala Gly Glu Gly Pro Lys 2420	Pro Thr Pro Arg 2425	Ala Val Thr Lys 2430
Leu Val Gly Ser Val Leu Gly 2435	Ala Arg Glu Val 2440	Arg Ser Ser Ile 2445
Ala Ala Ile Glu Ala Leu Gly 2450	Gly Lys Ala Ile 2455	Tyr Ser Ser Cys 2460
Asp Val Asn Ser Ala Ala Asp 2465	Val Ala Lys Ala 2470	Val Arg Asp Ala 2475
Glu Ser Gln Leu Gly Ala Arg 2480	Val Ser Gly Ile 2485	Val His Ala Ser 2490
Gly Val Leu Arg Asp Arg 2495	Leu Ile Glu Lys Lys 2500	Leu Pro Asp Glu 2505
Phe Asp Ala Val Phe Gly Thr 2510	Lys Val Thr Gly 2515	Leu Glu Asn Leu 2520
Leu Ala Ala Val Asp Arg Ala 2525	Asn Leu Lys His 2530	Met Val Leu Phe 2535
Ser Ser Leu Ala Gly Phe His 2540	Gly Asn Val Gly 2545	Gln Ser Asp Tyr 2550
Ala Met Ala Asn Glu Ala Leu 2555	Asn Lys Met Gly 2560	Leu Glu Leu Ala 2565

Lys Asp 2570	Val Ser Val Lys Ser 2575	Ile Cys Phe Gly Pro 2580	Trp Asp Gly
Gly Met 2585	Val Thr Pro Gln Leu 2590	Lys Lys Gln Phe Gln 2595	Glu Met Gly
Val Gln 2600	Ile Ile Pro Arg Glu 2605	Gly Gly Ala Asp Thr 2610	Val Ala Arg
Ile Val 2615	Leu Gly Ser Ser Pro 2620	Ala Glu Ile Leu Val 2625	Gly Asn Trp
Arg Thr 2630	Pro Ser Lys Lys Val 2635	Gly Ser Asp Thr Ile 2640	Thr Leu His
Arg Lys 2645	Ile Ser Ala Lys Ser 2650	Asn Pro Phe Leu Glu 2655	Asp His Val
Ile Gln 2660	Gly Arg Arg Val Leu 2665	Pro Met Thr Leu Ala 2670	Ile Gly Ser
Leu Ala 2675	Glu Thr Cys Leu Gly 2680	Leu Phe Pro Gly Tyr 2685	Ser Leu Trp
Ala Ile 2690	Asp Asp Ala Gln Leu 2695	Phe Lys Gly Val Thr 2700	Val Asp Gly
Asp Val 2705	Asn Cys Glu Val Thr 2710	Leu Thr Pro Ser Thr 2715	Ala Pro Ser
Gly Arg 2720	Val Asn Val Gln Ala 2725	Thr Leu Lys Thr Phe 2730	Ser Ser Gly
Lys Leu 2735	Val Pro Ala Tyr Arg 2740	Ala Val Ile Val Leu 2745	Ser Asn Gln
Gly Ala 2750	Pro Pro Ala Asn Ala 2755	Thr Met Gln Pro Pro 2760	Ser Leu Asp
Ala Asp 2765	Pro Ala Leu Gln Gly 2770	Ser Val Tyr Asp Gly 2775	Lys Thr Leu
Phe His 2780	Gly Pro Ala Phe Arg 2785	Gly Ile Asp Asp Val 2790	Leu Ser Cys
Thr Lys 2795	Ser Gln Leu Val Ala 2800	Lys Cys Ser Ala Val 2805	Pro Gly Ser

Asp Ala Ala Arg Gly Glu Phe Ala Thr Asp Thr Asp Ala His Asp
2810 2815 2820

Pro Phe Val Asn Asp Leu Ala Phe Gln Ala Met Leu Val Trp Val
2825 2830 2835

Arg Arg Thr Leu Gly Gln Ala Ala Leu Pro Asn Ser Ile Gln Arg
2840 2845 2850

Ile Val Gln His Arg Pro Val Pro Gln Asp Lys Pro Phe Tyr Ile
2855 2860 2865

Thr Leu Arg Ser Asn Gln Ser Gly Gly His Ser Gln His Lys His
2870 2875 2880

Ala Leu Gln Phe His Asn Glu Gln Gly Asp Leu Phe Ile Asp Val
2885 2890 2895

Gln Ala Ser Val Ile Ala Thr Asp Ser Leu Ala Phe
2900 2905 2910

<210> 55

<211> 2895

<212> PRT

<213> *Thraustochytrium aureum*

<400> 55

Arg Lys Cys Ile Arg Pro Ser Leu Gly His His Trp Ala Ile Ile Gly
1 5 10 15

Val Leu Gly Arg Ala Leu Arg Ile Val Arg Pro Ile Arg Tyr Glu Ala
20 25 30

Thr Asn Leu Arg Arg Leu Pro Arg Ser Gly Trp Leu Val Ala Leu Gly
35 40 45

Leu Phe Cys Asp Leu Ser Ser Cys Ala Gly Lys Leu Asp Leu Gln Thr
50 55 60

Arg Asp Thr Ala Lys Asp Pro Cys Cys Lys Arg Lys Trp Ser Ala Ser
65 70 75 80

Arg Ala Pro Pro Arg Pro Arg Ala Glu Ala Asp Lys Ala Ser Asn Glu
85 90 95

Met Glu Thr Lys Asp Asp Arg Val Ala Ile Val Gly Met Ser Ala Ile
100 105 110

Leu Pro Cys Gly Glu Ser Val Arg Glu Ser Trp Glu Ala Ile Arg Glu
 115 120 125
 Gly Leu Asp Cys Leu Gln Asp Leu Pro Ala Asp Arg Val Asp Ile Thr
 130 135 140
 Ala Tyr Tyr Asp Pro Asn Lys Thr Thr Lys Asp Lys Ile Tyr Cys Lys
 145 150 155 160
 Arg Gly Gly Phe Ile Pro Glu Tyr Asp Phe Asp Ala Arg Glu Phe Gly
 165 170 175
 Leu Asn Met Phe Gln Met Glu Asp Ser Asp Ala Asn Gln Thr Val Thr
 180 185 190
 Leu Leu Lys Val Lys Glu Ala Leu Glu Asp Ala Gly Val Glu Pro Phe
 195 200 205
 Thr Lys Lys Lys Lys Asn Ile Gly Cys Val Leu Gly Ile Gly Gly Gly
 210 215 220
 Gln Lys Ala Ser His Glu Phe Tyr Ser Arg Leu Asn Tyr Val Val Val
 225 230 235 240
 Glu Lys Val Leu Arg Lys Met Asn Leu Pro Asp Glu Val Val Glu Ala
 245 250 255
 Ala Val Glu Lys Tyr Lys Ala Asn Phe Pro Glu Trp Arg Leu Asp Ser
 260 265 270
 Phe Pro Gly Phe Leu Gly Asn Val Thr Ala Gly Arg Cys Ser Asn Val
 275 280 285
 Phe Asn Met Glu Gly Met Asn Cys Val Val Asp Ala Ala Cys Ala Ser
 290 295 300
 Ser Leu Ile Ala Ile Lys Val Ala Ile Asp Glu Leu Leu His Gly Asp
 305 310 315 320
 Cys Asp Thr Met Ile Ala Gly Ala Thr Cys Thr Asp Asn Ser Ile Gly
 325 330 335
 Met Tyr Met Ala Phe Ser Lys Thr Pro Val Phe Ser Thr Asp Gln Ser
 340 345 350
 Val Lys Ala Tyr Asp Ala Lys Thr Lys Gly Met Leu Ile Gly Glu Gly
 355 360 365
 Ser Ala Met Val Val Leu Lys Arg Tyr Ala Asp Ala Val Arg Asp Gly

370		375		380
Asp Glu Ile His Ala Val Ile Arg Ala Cys Ala Ser Ser Ser Asp Gly				
385		390		400
Lys Ala Ala Gly Ile Tyr Ala Pro Thr Val Ser Gly Gln Glu Glu Ala				
	405		410	415
Leu Arg Arg Ala Tyr Ala Arg Ala Gly Val Asp Pro Ser Thr Val Thr				
	420		425	430
Leu Val Glu Gly His Gly Thr Gly Thr Pro Val Gly Asp Arg Ile Glu				
	435		440	445
Leu Thr Ala Leu Arg Asn Val Phe Asp Ala Ala Asn Lys Gly Arg Lys				
	450		455	460
Glu Thr Val Ala Val Gly Ser Ile Lys Ser Gln Ile Gly His Leu Lys				
465		470		480
Ala Val Ala Gly Phe Ala Gly Leu Val Lys Val Val Met Ala Leu Lys				
	485		490	495
His Lys Thr Leu Pro Gln Thr Ile Asn Val His Asp Pro Pro Ala Leu				
	500		505	510
His Asp Gly Ser Pro Ile Gln Asp Ser Ser Leu Tyr Ile Asn Thr Met				
	515		520	525
Asn Arg Pro Trp Phe Thr Ala Pro Gly Val Pro Arg Arg Ala Gly Ile				
	530		535	540
Ser Ser Phe Gly Phe Gly Gly Ala Asn Tyr His Ala Val Leu Glu Glu				
545		550		560
Ala Glu Pro Glu His Ala Lys Pro Tyr Arg Met Asn Gln Val Pro Gln				
	565		570	575
Pro Val Leu Leu His Ala Ser Ser Ala Ser Ala Leu Ala Ser Ile Cys				
	580		585	590
Asp Ala Gln Ala Asp Ala Leu Gln Ala Ala Val Ser Pro Glu Ala Ser				
	595		600	605
Lys His Ala Asp Tyr Arg Ala Ile Val Ala Phe His Glu Ala Phe Lys				
	610		615	620
Leu Arg Ala Gly Val Pro Ala Gly His Ala Arg Ile Gly Phe Val Ser				
625		630		640

Gly Ser Ala Ala Ala Thr Leu Ala Val Leu Arg Ala Ala Ser Ala Lys
 645 650 655

Leu Lys Gln Ser Ser Ala Thr Leu Glu Trp Thr Leu Leu Arg Glu Gly
 660 665 670

Val Thr Tyr Arg Ser Ala Ala Met His Thr Pro Gly Ser Val Ala Ala
 675 680 685

Leu Phe Ala Gly Gln Gly Ala Gln Tyr Thr His Met Phe Ala Asp Val
 690 695 700

Ala Met Asn Trp Pro Pro Phe Arg Ser Ala Val Gln Glu Met Asp Ala
 705 710 715 720

Ala Gln Val Thr Ala Ala Ala Pro Lys Arg Leu Ser Glu Val Leu Tyr
 725 730 735

Pro Arg Lys Pro Tyr Ala Ala Glu Pro Glu Gln Asp Asn Lys Ala Ile
 740 745 750

Ser Met Thr Ile Asn Ser Gln Pro Ala Leu Met Ala Cys Ala Ala Gly
 755 760 765

Ala Phe Glu Val Phe Arg Gln Ala Gly Leu Ala Pro Asp His Val Ala
 770 775 780

Gly His Ser Leu Gly Glu Phe Gly Ala Leu Leu Ala Ala Gly Cys Ala
 785 790 795 800

Ser Arg Glu Glu Leu Phe Arg Leu Val Cys Ser Arg Ala Lys Ala Met
 805 810 815

Gln Asp Val Pro Gln Gly Asp Gly Ala Trp Leu Ala Asn Cys Asn Ser
 820 825 830

Pro Ser Gln Val Val Ile Ser Gly Asp Lys Thr Ala Val Glu Arg Glu
 835 840 845

Ser Ser Arg Leu Ala Gly Leu Gly Phe Arg Ile Ile Pro Leu Ala Cys
 850 855 860

Glu Gly Ala Phe His Ser Pro His Met Thr Ala Ala Gln Ala Thr Phe
 865 870 875 880

Gln Ala Ala Leu Asp Ser Leu Lys Ile Ser Thr Pro Thr Asn Gly Ala
 885 890 895

Arg Leu Tyr Asn Asn Val Ser Gly Lys Thr Cys Arg Ser Leu Gly Glu
 900 905 910

Leu Arg Asp Cys Leu Gly Lys His Met Thr Ser Pro Val Leu Phe Gln
 915 920 925

Ala Gln Val Glu Asn Met Tyr Ala Ala Gly Ala Arg Ile Phe Val Glu
 930 935 940

Phe Gly Pro Lys Gln Val Leu Ser Lys Leu Val Gly Glu Ile Leu Ala
 945 950 955 960

Asp Lys Ser Asp Phe Val Thr Val Ala Val Asn Ser Ser Ser Ser Lys
 965 970 975

Asp Ser Asp Val Gln Leu Arg Glu Ala Ala Ala Lys Leu Ala Val Leu
 980 985 990

Gly Val Pro Leu Ala Asn Phe Asp Pro Trp Glu Leu Cys Asp Ala Arg
 995 1000 1005

Arg Leu Arg Glu Cys Pro Arg Ser Lys Thr Thr Leu Arg Leu Ser
 1010 1015 1020

Ala Ala Thr Tyr Val Ser Asn Lys Thr Leu Ala Ala Arg Glu Lys
 1025 1030 1035

Val Met Glu Asp Asn Cys Asp Phe Ser Ser Leu Phe Ala Ser Gly
 1040 1045 1050

Pro Ala Ser Gln Glu Met Glu Arg Glu Ile Ala Asn Leu Arg Ala
 1055 1060 1065

Glu Leu Glu Ala Ala Gln Arg Gln Leu Asp Thr Ala Lys Thr Gln
 1070 1075 1080

Leu Ala Arg Lys Gln Val Gln Asp Pro Thr Ala Asp Arg Gln Arg
 1085 1090 1095

Asp Met Ile Ala Lys His Arg Ser Thr Leu Ala Ala Met Val Lys
 1100 1105 1110

Glu Phe Glu Ala Leu Ala Ser Gly Ser Pro Cys Ala Val Pro Phe
 1115 1120 1125

Ala Pro Val Val Asp Thr Ala Val Glu Asp Val Pro Phe Ala Asp
 1130 1135 1140

Lys Val 1145	Ser Thr Pro Pro Pro	Gln Val Thr Ser Ala 1155	Pro Ile Ala
Glu Leu 1160	Ala Arg Ala Glu Ala	Val Val Met Glu Val 1170	Leu Ala Ala
Lys Thr 1175	Gly Tyr Glu Val Asp	Met Ile Glu Ala Asp 1185	Met Leu Leu
Asp Ala 1190	Glu Leu Gly Ile Asp	Ser Val Lys Arg Ile 1200	Glu Ile Leu
Ala Ala 1205	Val Gln Ala Gln Leu	Gly Val Glu Ala Lys 1215	Asp Val Asp
Ala Leu 1220	Ser Arg Thr Arg Thr	Val Gly Glu Val Val 1230	Asp Ala Met
Lys Ala 1235	Glu Ile Gly Gly Gln	Ala Thr Ser Ala Pro 1245	Ser Pro Met
Ala Gln 1250	Pro Gln Ala Ser Ala	Pro Ser Pro Ser Pro 1260	Thr Ala Ser
Val Leu 1265	Pro Lys Pro Val Ala	Leu Pro Ala Ser Val 1275	Asp Pro Ala
Lys Leu 1280	Ala Arg Ala Glu Ala	Val Val Met Glu Val 1290	Leu Ala Ala
Lys Thr 1295	Gly Tyr Glu Val Asp	Met Ile Glu Ala Asp 1305	Met Leu Leu
Asp Ala 1310	Glu Leu Gly Ile Asp	Ser Val Lys Arg Ile 1320	Glu Ile Leu
Ala Ala 1325	Val Gln Ala Gln Leu	Gly Val Glu Ala Lys 1335	Asp Val Asp
Ala Leu 1340	Ser Arg Thr Arg Thr	Val Gly Glu Val Val 1350	Asp Ala Met
Lys Ala 1355	Glu Ile Gly Gly Gln	Ala Thr Ser Ala Pro 1365	Ala Ser Val
Ala Gln 1370	Pro Gln Ala Ser Ala	Pro Ser Pro Ser Ala 1380	Thr Thr Ala
Ser Val	Leu Pro Lys Pro Val	Ala Ala Pro Thr Ser	Ala Asp Pro

ES 2 623 610 T3

1385	1390	1395
Ala Lys Leu Ala Arg Ala Glu	Ala Val Val Met Glu	Val Leu Ala
1400	1405	1410
Ala Lys Thr Gly Tyr Glu Val	Asp Met Ile Glu Ala	Asp Met Leu
1415	1420	1425
Leu Asp Ala Glu Leu Gly Ile	Asp Ser Val Lys Arg	Ile Glu Ile
1430	1435	1440
Leu Ala Ala Val Gln Ala Gln	Leu Gly Val Glu Ala	Lys Asp Val
1445	1450	1455
Asp Ala Leu Ser Arg Thr Arg	Thr Val Gly Glu Val	Val Glu Ala
1460	1465	1470
Met Lys Ala Glu Ile Gly Gly	Gln Ala Thr Ser Ala	Pro Ala Ser
1475	1480	1485
Val Ala Gln Pro Gln Ile Ser	Val Ser Pro Thr Pro	Leu Ala Ala
1490	1495	1500
Ser Pro Ser Ala Asp Pro Ala	Lys Leu Ala Arg Ala	Glu Ala Val
1505	1510	1515
Val Met Glu Val Leu Ala Ala	Lys Thr Gly Tyr Glu	Val Asp Met
1520	1525	1530
Ile Glu Ala Asp Met Leu Leu	Asp Ala Glu Leu Gly	Ile Asp Ser
1535	1540	1545
Val Lys Arg Ile Glu Ile Leu	Ala Ala Val Gln Ala	Gln Leu Gly
1550	1555	1560
Val Glu Ala Lys Asp Val Asp	Ala Leu Ser Arg Thr	Arg Thr Val
1565	1570	1575
Gly Glu Val Val Asp Ala Met	Lys Ala Glu Ile Gly	Gly Gln Ala
1580	1585	1590
Thr Ser Ala Pro Ala Ser Val	Ala Gln Pro Gln Ala	Ser Ala Pro
1595	1600	1605
Ser Pro Ser Ala Thr Ala Ser	Val Leu Pro Lys Pro	Val Ala Ala
1610	1615	1620
Pro Thr Ser Ala Asp Pro Ala	Lys Leu Ala Arg Ala	Glu Ala Val
1625	1630	1635

Val Met 1640	Glu Val Leu Ala Ala 1645	Lys Thr Gly Tyr Glu 1650	Val Asp Met
Ile Glu 1655	Ala Asp Met Leu Leu 1660	Asp Ala Glu Leu Gly 1665	Ile Asp Ser
Val Lys 1670	Arg Ile Glu Ile Leu 1675	Ala Ala Val Gln Ala 1680	Gln Leu Gly
Val Glu 1685	Ala Lys Asp Val Asp 1690	Ala Leu Ser Arg Thr 1695	Arg Thr Val
Gly Glu 1700	Val Val Glu Ala Met 1705	Lys Ala Glu Ile Gly 1710	Gly Gln Ala
Thr Ser 1715	Ala Pro Ala Ser Met 1720	Ala Gln Pro Gln Ile 1725	Ser Val Ser
Pro Thr 1730	Pro Leu Ala Ala Ser 1735	Pro Ser Ala Asp Pro 1740	Ala Lys Leu
Ala Arg 1745	Ala Glu Ala Val Val 1750	Met Glu Val Leu Ala 1755	Ala Lys Thr
Gly Tyr 1760	Glu Val Asp Met Ile 1765	Glu Ala Asp Met Leu 1770	Leu Asp Ala
Glu Leu 1775	Gly Ile Asp Ser Val 1780	Lys Arg Ile Glu Ile 1785	Leu Ala Ala
Val Gln 1790	Ala Gln Leu Gly Val 1795	Glu Ala Lys Asp Val 1800	Asp Ala Leu
Ser Arg 1805	Thr Arg Thr Val Gly 1810	Glu Val Val Asp Ala 1815	Met Lys Ala
Glu Ile 1820	Gly Gly Gln Ala Thr 1825	Ser Ala Pro Ala Ser 1830	Val Ala Gln
Pro Gln 1835	Ala Ser Ala Pro Ser 1840	Pro Ser Ala Thr Ala 1845	Ser Ala Pro
Val Thr 1850	Pro Leu Ala Ala Pro 1855	Ala Ser Val Asp Pro 1860	Ala Lys Leu
Ala Arg 1865	Ala Glu Ala Val Val 1870	Met Glu Val Leu Ala 1875	Ala Lys Thr

Gly Tyr 1880	Glu Val Asp Met 1885	Ile Glu Ala Asp Met 1890	Leu Leu Asp Ala
Glu Leu 1895	Gly Ile Asp Ser 1900	Lys Arg Ile Glu 1905	Ile Leu Ala Ala
Val Gln 1910	Ala Gln Leu Gly 1915	Glu Ala Lys Asp 1920	Val Asp Ala Leu
Ser Arg 1925	Thr Arg Thr Val 1930	Gly Glu Val Val Asp 1935	Ala Met Lys Ala
Glu Ile 1940	Gly Gly Gln Ala 1945	Thr Ser Ala Pro Ala 1950	Ser Val Ala Gln
Pro Gln 1955	Ala Ser Ala Pro 1960	Ser Pro Ser Ala Thr 1965	Ala Ser Val Leu
Pro Lys 1970	Pro Val Ala Ser 1975	Pro Ala Ser Val Asp 1980	Pro Ala Lys Leu
Ala Arg 1985	Ala Glu Ala Val 1990	Val Met Glu Val Leu 1995	Ala Ala Lys Thr
Gly Tyr 2000	Glu Val Asp Met 2005	Ile Asp Ala Asp Met 2010	Leu Leu Asp Ala
Glu Leu 2015	Gly Ile Asp Ser 2020	Val Lys Arg Ile Glu 2025	Ile Leu Ala Ala
Val Gln 2030	Ala Gln Leu Gly 2035	Val Glu Ala Lys Asp 2040	Val Asp Ala Leu
Ser Arg 2045	Thr Arg Thr Val 2050	Gly Glu Val Val Glu 2055	Ala Met Lys Ala
Glu Ile 2060	Gly Ala Ala Gly 2065	Pro Asn Asp Ala Gln 2070	Ala Ala Ser Gly
His Leu 2075	Phe Gly Thr Gly 2080	Cys Glu Asp Leu Ser 2085	Leu Cys Ser Ala
Ser Val 2090	Val Glu Ile Ala 2095	Arg Cys Ser Glu Leu 2100	Ala Leu Glu Arg
Pro Met 2105	Asp Arg Pro Ile 2110	Leu Ile Val Ser Asp 2115	Gly Ser Ala Leu

Pro	Ala	Ala	Leu	Ala	Ser	Arg	Leu	Gly	Ser	Cys	Ala	Val	Ile	Leu
2120						2125					2130			
Thr	Thr	Ala	Gly	Glu	Thr	Asp	Gln	Ser	Val	Arg	Ser	Thr	Lys	His
2135						2140					2145			
Val	Asp	Met	Glu	Gly	Trp	Gly	Glu	Ala	Asp	Leu	Val	Arg	Ala	Leu
2150						2155					2160			
Glu	Ala	Val	Glu	Ser	Arg	Phe	Gly	Val	Pro	Gly	Gly	Val	Val	Val
2165						2170					2175			
Leu	Glu	Arg	Ala	Ser	Glu	Thr	Ala	Arg	Asp	Gln	Leu	Gly	Phe	Ala
2180						2185					2190			
Leu	Leu	Leu	Ala	Lys	His	Ser	Ser	Lys	Ala	Leu	Asn	Gln	Gln	Ile
2195						2200					2205			
Pro	Gly	Gly	Arg	Ala	Cys	Phe	Val	Gly	Val	Ser	Arg	Ile	Asp	Gly
2210						2215					2220			
Lys	Leu	Gly	Leu	Ser	Gly	Ala	Cys	Ala	Lys	Gly	Lys	Gly	Trp	Ala
2225						2230					2235			
Glu	Ala	Ala	Glu	Ile	Ala	Gln	Gln	Gly	Ala	Val	Ala	Gly	Leu	Cys
2240						2245					2250			
Lys	Thr	Leu	Asp	Leu	Glu	Trp	Pro	His	Val	Phe	Ala	Arg	Ser	Ile
2255						2260					2265			
Asp	Ile	Glu	Leu	Gly	Ala	Asn	Glu	Glu	Thr	Ala	Ala	Gln	Ala	Ile
2270						2275					2280			
Phe	Glu	Glu	Leu	Ser	Cys	Pro	Asp	Leu	Thr	Val	Arg	Glu	Ala	Gly
2285						2290					2295			
Tyr	Thr	Lys	Asp	Gly	Lys	Arg	Trp	Thr	Thr	Glu	Ala	Arg	Pro	Val
2300						2305					2310			
Gly	Leu	Gly	Lys	Pro	Lys	Gln	Ala	Leu	Arg	Ser	Ser	Asp	Val	Phe
2315						2320					2325			
Leu	Val	Ser	Gly	Gly	Ala	Arg	Gly	Ile	Thr	Pro	Val	Cys	Val	Arg
2330						2335					2340			
Glu	Leu	Ala	Lys	Ser	Ile	Ser	Gly	Gly	Thr	Phe	Val	Leu	Leu	Gly
2345						2350					2355			
Arg	Ser	Pro	Leu	Ala	Asp	Asp	Pro	Ala	Trp	Ala	Cys	Gly	Val	Glu

ES 2 623 610 T3

2360		2365		2370
Glu Ala Asn Ile Gly Thr	Ala Ala Met Ala His	Leu Lys Ala Glu		
2375	2380	2385		
Phe Ala Ala Gly Arg Gly	Pro Lys Pro Thr Pro	Lys Ala His Lys		
2390	2395	2400		
Ala Leu Val Gly Ser Val	Leu Gly Ala Arg Glu	Val Leu Gly Ser		
2405	2410	2415		
Leu Glu Ser Ile Arg Ala	Gln Gly Ala Arg Ala	Glu Tyr Val Ser		
2420	2425	2430		
Cys Asp Val Ser Cys Ala	Glu Arg Val Lys Ala	Val Val Asp Asp		
2435	2440	2445		
Leu Glu Arg Arg Val Gly	Ala Val Thr Gly Val	Val His Ala Ser		
2450	2455	2460		
Gly Val Leu Arg Asp Lys	Ser Val Glu Arg Leu	Glu Leu Ala Asp		
2465	2470	2475		
Phe Glu Val Val Tyr Gly	Thr Lys Val Asp Gly	Leu Leu Asn Leu		
2480	2485	2490		
Leu Gln Ala Val Asp Arg	Pro Lys Leu Arg His	Leu Val Leu Phe		
2495	2500	2505		
Ser Ser Leu Ala Gly Phe	His Gly Asn Thr Gly	Gln Ala Val Tyr		
2510	2515	2520		
Ala Met Ala Asn Glu Ala	Leu Asn Lys Met Ala	Phe His Leu Glu		
2525	2530	2535		
Thr Ala Met Pro Gly Leu	Ser Val Lys Thr Ile	Gly Phe Gly Pro		
2540	2545	2550		
Trp Asp Gly Gly Met Val	Asn Asp Ala Leu Lys	Ala His Phe Ala		
2555	2560	2565		
Ser Met Gly Val Gln Ile	Ile Pro Leu Asp Gly	Gly Ala Glu Thr		
2570	2575	2580		
Val Ser Arg Ile Ile Gly	Ala Cys Ser Pro Thr	Gln Val Leu Val		
2585	2590	2595		
Gly Asn Trp Gly Leu Pro	Pro Val Val Pro Asn	Ala Ser Val His		
2600	2605	2610		

Lys	Ile	Thr	Val	Arg	Leu	Gly	Gly	Glu	Ser	Ala	Asn	Pro	Phe	Leu
2615						2620					2625			
Ser	Ser	His	Thr	Ile	Gln	Gly	Arg	Lys	Val	Leu	Pro	Met	Thr	Val
2630						2635					2640			
Ala	Leu	Gly	Leu	Leu	Ala	Glu	Ala	Ala	Arg	Gly	Leu	Tyr	Val	Gly
2645						2650					2655			
His	Gln	Val	Val	Gly	Ile	Glu	Asp	Ala	Gln	Val	Phe	Gln	Gly	Val
2660						2665					2670			
Val	Leu	Asp	Lys	Gly	Ala	Thr	Cys	Glu	Val	Gln	Leu	Arg	Arg	Glu
2675						2680					2685			
Ser	Ser	Thr	Ala	Ser	Pro	Ser	Glu	Val	Val	Leu	Ser	Ala	Ser	Leu
2690						2695					2700			
Asn	Val	Phe	Ala	Ala	Gly	Lys	Val	Val	Pro	Ala	Tyr	Arg	Ala	His
2705						2710					2715			
Val	Val	Leu	Gly	Ala	Ser	Gly	Pro	Arg	Thr	Gly	Gly	Val	Gln	Leu
2720						2725					2730			
Glu	Leu	Lys	Asp	Leu	Gly	Val	Asp	Ala	Asp	Pro	Ala	Cys	Ser	Val
2735						2740					2745			
Gly	Lys	Gly	Ala	Leu	Tyr	Asp	Gly	Arg	Thr	Leu	Phe	His	Gly	Pro
2750						2755					2760			
Ala	Phe	Gln	Tyr	Met	Asp	Glu	Val	Leu	Arg	Cys	Ser	Pro	Ala	Glu
2765						2770					2775			
Leu	Ala	Val	Arg	Cys	Arg	Val	Val	Pro	Ser	Ala	Ala	Gln	Asp	Arg
2780						2785					2790			
Gly	Gln	Phe	Val	Ser	Arg	Gly	Val	Leu	Tyr	Asp	Pro	Phe	Leu	Asn
2795						2800					2805			
Asp	Thr	Val	Phe	Gln	Ala	Leu	Leu	Val	Trp	Ala	Arg	Leu	Val	Arg
2810						2815					2820			
Asp	Ser	Ala	Ser	Leu	Pro	Ser	Asn	Val	Glu	Arg	Ile	Ser	Phe	His
2825						2830					2835			
Gly	Gln	Pro	Pro	Ser	Glu	Gly	Glu	Val	Phe	Tyr	Thr	Thr	Leu	Lys
2840						2845					2850			

Leu Asp Ser Ala Ala Ser Gly Pro Leu Asp Pro Ile Ala Lys Ala
2855 2860 2865

Gln Phe Phe Leu His Arg Ala Cys Gly Ala Val Phe Ala Ser Gly
2870 2875 2880

Arg Ala Ser Val Val Leu Asn Lys Ala Leu Ser Phe
2885 2890 2895

<210> 56

<211> 2811

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 56

Met Lys Asp Met Glu Asp Arg Arg Val Ala Ile Val Gly Met Ser Ala
1 5 10 15

His Leu Pro Cys Gly Thr Asp Val Lys Glu Ser Trp Gln Ala Ile Arg
20 25 30

Asp Gly Ile Asp Cys Leu Ser Asp Leu Pro Ala Asp Arg Leu Asp Val
35 40 45

Thr Ala Tyr Tyr Asn Pro Asn Lys Ala Thr Lys Asp Lys Ile Tyr Cys
50 55 60

Lys Arg Gly Gly Phe Ile Pro Asn Tyr Asp Phe Asp Pro Arg Glu Phe
65 70 75 80

Gly Leu Asn Met Phe Gln Met Glu Asp Ser Asp Ala Asn Gln Thr Leu
85 90 95

Thr Leu Leu Lys Val Lys Gln Ala Leu Glu Asp Ala Ser Ile Glu Pro
100 105 110

Phe Thr Lys Glu Lys Lys Asn Ile Gly Cys Val Leu Gly Ile Gly Gly
115 120 125

Gly Gln Lys Ala Ser His Glu Phe Tyr Ser Arg Leu Asn Tyr Val Val
130 135 140

Val Glu Lys Val Leu Arg Lys Met Gly Leu Pro Asp Ala Asp Val Glu
145 150 155 160

Glu Ala Val Glu Lys Tyr Lys Ala Asn Phe Pro Glu Trp Arg Leu Asp
165 170 175

Ser Phe Pro Gly Phe Leu Gly Asn Val Thr Ala Gly Arg Cys Ser Asn
 180 185 190
 Thr Phe Asn Met Glu Gly Met Asn Cys Val Val Asp Ala Ala Cys Ala
 195 200 205
 Ser Ser Leu Ile Ala Ile Lys Val Ala Val Glu Glu Leu Leu Phe Gly
 210 215 220
 Asp Cys Asp Thr Met Ile Ala Gly Ala Thr Cys Thr Asp Asn Ser Leu
 225 230 235 240
 Gly Met Tyr Met Ala Phe Ser Lys Thr Pro Val Phe Ser Thr Asp Pro
 245 250 255
 Ser Val Arg Ala Tyr Asp Glu Lys Thr Lys Gly Met Leu Ile Gly Glu
 260 265 270
 Gly Ser Ala Met Phe Val Leu Lys Arg Tyr Ala Asp Ala Val Arg Asp
 275 280 285
 Gly Asp Thr Ile His Ala Val Leu Arg Ser Cys Ser Ser Ser Ser Asp
 290 295 300
 Gly Lys Ala Ala Gly Ile Tyr Thr Pro Thr Ile Ser Gly Gln Glu Glu
 305 310 315 320
 Ala Leu Arg Arg Ala Tyr Ala Arg Ala Gly Val Cys Pro Ser Thr Ile
 325 330 335
 Gly Leu Val Glu Gly His Gly Thr Gly Thr Pro Val Gly Asp Arg Ile
 340 345 350
 Glu Leu Thr Ala Leu Arg Asn Leu Phe Asp Lys Ala Phe Gly Ser Lys
 355 360 365
 Lys Glu Gln Ile Ala Val Gly Ser Ile Lys Ser Gln Ile Gly His Leu
 370 375 380
 Lys Ser Val Ala Gly Phe Ala Gly Leu Val Lys Ala Val Leu Ala Leu
 385 390 395 400
 Lys His Lys Thr Leu Pro Gly Ser Ile Asn Val Asp Gln Pro Pro Leu
 405 410 415
 Leu Tyr Asp Gly Thr Gln Ile Gln Asp Ser Ser Leu Tyr Ile Asn Lys
 420 425 430
 Thr Asn Arg Pro Trp Phe Thr Gln Asn Lys Leu Pro Arg Arg Ala Gly

ES 2 623 610 T3

435		440		445
Val Ser Ser Phe Gly Phe Gly Gly Ala Asn Tyr His Ala Val Leu Glu	450	455	460	
Glu Phe Glu Pro Glu His Glu Lys Pro Tyr Arg Leu Asn Thr Val Gly	465	470	475	480
His Pro Val Leu Leu Tyr Ala Pro Ser Val Glu Ala Leu Lys Val Leu	485	490	495	
Cys Asn Asp Gln Leu Ala Glu Leu Thr Ile Ala Leu Glu Glu Ala Lys	500	505	510	
Thr His Lys Asn Val Asp Lys Val Cys Gly Tyr Lys Phe Ile Asp Glu	515	520	525	
Phe Gln Leu Gln Gly Ser Cys Pro Pro Glu Asn Pro Arg Val Gly Phe	530	535	540	
Leu Ala Thr Leu Pro Thr Ser Asn Ile Ile Val Ala Leu Lys Ala Ile	545	550	555	560
Leu Ala Gln Leu Asp Ala Lys Pro Asp Ala Lys Lys Trp Asp Leu Pro	565	570	575	
His Lys Lys Ala Phe Gly Ala Thr Phe Ala Ser Ser Ser Val Lys Gly	580	585	590	
Ser Val Ala Ala Leu Phe Ala Gly Gln Gly Thr Gln Tyr Leu Asn Met	595	600	605	
Phe Ser Asp Val Ala Met Asn Trp Pro Pro Phe Arg Asp Ser Ile Val	610	615	620	
Ala Met Glu Glu Ala Gln Thr Glu Val Phe Glu Gly Gln Val Glu Pro	625	630	635	640
Ile Ser Lys Val Leu Phe Pro Arg Glu Arg Tyr Ala Ser Glu Ser Glu	645	650	655	
Gln Gly Asn Glu Leu Leu Cys Leu Thr Glu Tyr Ser Gln Pro Thr Thr	660	665	670	
Ile Ala Ala Ala Val Gly Ala Phe Asp Ile Phe Lys Ala Ala Gly Phe	675	680	685	
Lys Pro Asp Met Val Gly Gly His Ser Leu Gly Glu Phe Ala Ala Leu	690	695	700	

ES 2 623 610 T3

Tyr	Ala	Ala	Gly	Ser	Ile	Ser	Arg	Asp	Asp	Leu	Tyr	Lys	Leu	Val	Cys	705	710	715	720
Lys	Arg	Ala	Lys	Ala	Met	Ala	Asn	Ala	Ser	Asp	Gly	Ala	Met	Ala	Ala	725	730	735	
Val	Ile	Gly	Pro	Asp	Ala	Arg	Leu	Val	Thr	Pro	Gln	Asn	Ser	Asp	Val	740	745	750	
Tyr	Val	Ala	Asn	Phe	Asn	Ser	Ala	Thr	Gln	Val	Val	Ile	Ser	Gly	Thr	755	760	765	
Val	Gln	Gly	Val	Lys	Glu	Glu	Ser	Lys	Leu	Leu	Ile	Ser	Lys	Gly	Phe	770	775	780	
Arg	Val	Leu	Pro	Leu	Lys	Cys	Gln	Gly	Ala	Phe	His	Ser	Pro	Leu	Met	785	790	795	800
Gly	Pro	Ser	Glu	Asp	Ser	Phe	Lys	Ser	Leu	Val	Glu	Thr	Cys	Thr	Ile	805	810	815	
Ser	Pro	Pro	Lys	Asn	Val	Lys	Phe	Phe	Cys	Asn	Val	Ser	Gly	Lys	Glu	820	825	830	
Ser	Pro	Asn	Pro	Lys	Gln	Thr	Leu	Lys	Ser	His	Met	Thr	Ser	Ser	Val	835	840	845	
Gln	Phe	Glu	Glu	Gln	Ile	Arg	Asn	Met	Tyr	Asp	Ala	Gly	Ala	Arg	Val	850	855	860	
Phe	Leu	Glu	Phe	Gly	Pro	Arg	Gln	Val	Leu	Ala	Lys	Leu	Ile	Ala	Glu	865	870	875	880
Met	Phe	Pro	Ser	Cys	Thr	Ala	Ile	Ser	Val	Asn	Pro	Ala	Ser	Ser	Gly	885	890	895	
Asp	Ser	Asp	Val	Gln	Leu	Arg	Leu	Ala	Ala	Val	Lys	Phe	Ala	Val	Ser	900	905	910	
Gly	Ala	Ala	Leu	Ser	Thr	Phe	Asp	Pro	Trp	Glu	Tyr	Arg	Lys	Pro	Gln	915	920	925	
Asp	Leu	Leu	Ile	Arg	Lys	Pro	Arg	Lys	Thr	Ala	Leu	Val	Leu	Ser	Ala	930	935	940	
Ala	Thr	Tyr	Val	Ser	Pro	Lys	Thr	Leu	Ala	Glu	Arg	Lys	Lys	Ala	Met	945	950	955	960

ES 2 623 610 T3

Glu	Asp	Ile	Lys	Leu	Val	Ser	Ile	Thr	Pro	Arg	Asp	Ser	Met	Val	Ser	965	970	975
Ile	Gly	Lys	Ile	Ala	Gln	Glu	Val	Arg	Thr	Ala	Lys	Gln	Pro	Leu	Glu	980	985	990
Thr	Glu	Ile	Arg	Arg	Leu	Asn	Lys	Glu	Leu	Glu	His	Leu	Lys	Arg	Glu	995	1000	1005
Leu	Ala	Ala	Ala	Lys	Ala	Ser	Val	Lys	Ser	Ala	Ser	Lys	Ser	Ser		1010	1015	1020
Lys	Glu	Arg	Ser	Val	Leu	Ser	Lys	His	Arg	Ala	Leu	Leu	Gln	Asn		1025	1030	1035
Ile	Leu	Gln	Asp	Tyr	Asp	Asp	Leu	Arg	Val	Val	Pro	Phe	Ala	Val		1040	1045	1050
Arg	Ser	Val	Ala	Val	Asp	Asn	Thr	Ala	Pro	Tyr	Ala	Asp	Gln	Val		1055	1060	1065
Ser	Thr	Pro	Ala	Ser	Glu	Arg	Ser	Ala	Ser	Pro	Leu	Phe	Glu	Lys		1070	1075	1080
Arg	Ser	Ser	Val	Ser	Ser	Ala	Arg	Leu	Ala	Glu	Ala	Glu	Ala	Ala		1085	1090	1095
Val	Leu	Ser	Val	Leu	Ala	Asp	Lys	Thr	Gly	Tyr	Asp	Ser	Ser	Met		1100	1105	1110
Ile	Glu	Met	Asp	Met	Asp	Leu	Glu	Ser	Glu	Leu	Gly	Val	Asp	Ser		1115	1120	1125
Ile	Lys	Arg	Val	Glu	Ile	Met	Ser	Glu	Val	Gln	Thr	Leu	Leu	Ser		1130	1135	1140
Val	Glu	Val	Ser	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Lys	Thr	Val		1145	1150	1155
Gly	Asp	Val	Ile	Glu	Ala	Met	Lys	Leu	Glu	Leu	Gly	Gly	Pro	Gln		1160	1165	1170
Gly	Gln	Thr	Leu	Thr	Ala	Glu	Ser	Ile	Arg	Gln	Pro	Pro	Val	Ser		1175	1180	1185
Glu	Pro	Ala	Val	Pro	Thr	Ser	Ser	Ser	Ser	Ser	Ile	Ala	Asn	Val		1190	1195	1200

ES 2 623 610 T3

Ser Ser	Ala Arg	Leu Ala	Glu	Ala Glu	Ala Ala	Val	Leu Ser	Val
1205			1210			1215		
Leu Ala	Asp Lys	Thr Gly	Tyr	Asp Ser	Ser Met	Ile	Glu Met	Asp
1220			1225			1230		
Met Asp	Leu Glu	Ser Glu	Leu	Gly Val	Asp Ser	Ile	Lys Arg	Val
1235			1240			1245		
Glu Ile	Met Ser	Glu Val	Gln	Thr Leu	Leu Ser	Val	Glu Val	Ser
1250			1255			1260		
Asp Val	Asp Ala	Leu Ser	Arg	Thr Lys	Thr Val	Gly	Asp Val	Ile
1265			1270			1275		
Glu Ala	Met Lys	Leu Glu	Leu	Gly Gly	Pro Gln	Gly	Gln Thr	Leu
1280			1285			1290		
Thr Ala	Glu Ser	Ile Arg	Gln	Pro Pro	Val Ser	Glu	Pro Ala	Val
1295			1300			1305		
Pro Thr	Ser Ser	Ser Ser	Ser	Ile Ala	Asn Val	Ser	Ser Ala	Arg
1310			1315			1320		
Leu Ala	Glu Ala	Glu Ala	Ala	Val Leu	Ser Val	Leu	Ala Asp	Lys
1325			1330			1335		
Thr Gly	Tyr Asp	Ser Ser	Met	Ile Glu	Met Asp	Met	Asp Leu	Glu
1340			1345			1350		
Ser Glu	Leu Gly	Val Asp	Ser	Ile Lys	Arg Val	Glu	Ile Met	Ser
1355			1360			1365		
Glu Val	Gln Thr	Leu Leu	Ser	Val Glu	Val Ser	Asp	Val Asp	Ala
1370			1375			1380		
Leu Ser	Arg Thr	Lys Thr	Val	Gly Asp	Val Ile	Glu	Ala Met	Lys
1385			1390			1395		
Leu Glu	Leu Gly	Gly Pro	Gln	Gly Gln	Thr Leu	Thr	Ala Glu	Ser
1400			1405			1410		
Ile Arg	Gln Pro	Pro Val	Ser	Glu Pro	Ala Val	Pro	Thr Ser	Ser
1415			1420			1425		
Ser Ser	Ser Ile	Ala Asn	Val	Leu Ser	Ala Arg	Leu	Ala Glu	Ala
1430			1435			1440		
Glu Ala	Ala Val	Leu Ser	Val	Leu Ala	Asp Lys	Thr	Gly Tyr	Asp

ES 2 623 610 T3

1445	1450	1455
Ser Ser Met Ile Glu Met Asp Met Asp Leu Glu Ser Glu Leu Gly	1465	1470
1460		
Val Asp Ser Ile Lys Arg Val Glu Ile Met Ser Glu Val Gln Thr	1480	1485
1475		
Leu Leu Ser Val Glu Val Ser Asp Val Asp Ala Leu Ser Arg Thr	1495	1500
1490		
Lys Thr Val Gly Asp Val Ile Glu Ala Met Lys Leu Glu Leu Gly	1510	1515
1505		
Gly Pro Gln Gly Gln Thr Leu Thr Ala Glu Ser Ile Arg Gln Pro	1525	1530
1520		
Pro Val Ser Glu Pro Ala Val Pro Thr Ser Ser Ser Ser Ser Ile	1540	1545
1535		
Ala Asn Val Ser Ser Ala Arg Leu Ala Glu Ala Glu Ala Ala Val	1555	1560
1550		
Leu Ser Val Leu Ala Asp Lys Thr Gly Tyr Asp Ser Ser Met Ile	1570	1575
1565		
Glu Met Asp Met Asp Leu Glu Ser Glu Leu Gly Val Asp Ser Ile	1585	1590
1580		
Lys Arg Val Glu Ile Met Ser Glu Val Gln Thr Leu Leu Ser Val	1600	1605
1595		
Glu Val Ser Asp Val Asp Ala Leu Ser Arg Thr Lys Thr Val Gly	1615	1620
1610		
Asp Val Ile Glu Ala Met Lys Leu Glu Leu Gly Gly Pro Gln Gly	1630	1635
1625		
Gln Thr Leu Thr Ser Glu Pro Ile His Gln Pro Pro Val Ser Glu	1645	1650
1640		
Pro Ala Val Pro Thr Ser Ser Ser Ser Ser Ser Ile Ala Asn Val Ser	1660	1665
1655		
Ser Ala Arg Leu Ala Glu Ala Glu Ala Ala Val Leu Ser Val Leu	1675	1680
1670		
Ala Asp Lys Thr Gly Tyr Asp Ser Ser Met Ile Glu Met Asp Met	1690	1695
1685		

Asp 1700	Leu	Glu	Ser	Glu	Leu	Gly 1705	Val	Asp	Ser	Ile	Lys 1710	Arg	Val	Glu
Ile 1715	Met	Ser	Glu	Val	Gln	Thr 1720	Leu	Leu	Ser	Val	Glu 1725	Val	Ser	Asp
Val 1730	Asp	Ala	Leu	Ser	Arg	Thr 1735	Lys	Thr	Val	Gly	Asp 1740	Val	Ile	Glu
Ala 1745	Met	Lys	Met	Glu	Leu	Gly 1750	Gly	Pro	Gln	Gly	Gln 1755	Thr	Leu	Thr
Ala 1760	Glu	Ser	Ile	Arg	Gln	Pro 1765	Pro	Val	Ser	Glu	Pro 1770	Ala	Val	Pro
Thr 1775	Ser	Ser	Ser	Ser	Ser	Ile 1780	Ala	Asn	Val	Ser	Ser 1785	Ala	Arg	Leu
Ala 1790	Glu	Ala	Glu	Ala	Ala	Val 1795	Leu	Ser	Val	Leu	Ala 1800	Asp	Lys	Thr
Gly 1805	Tyr	Asp	Ser	Ser	Met	Ile 1810	Glu	Met	Asp	Met	Asp 1815	Leu	Glu	Ser
Glu 1820	Leu	Gly	Val	Asp	Ser	Ile 1825	Lys	Arg	Val	Glu	Ile 1830	Met	Ser	Glu
Val 1835	Gln	Ala	Leu	Leu	Ser	Val 1840	Glu	Val	Ser	Asp	Val 1845	Asp	Ala	Leu
Ser 1850	Arg	Thr	Lys	Thr	Val	Gly 1855	Asp	Val	Ile	Glu	Ala 1860	Met	Lys	Met
Glu 1865	Leu	Gly	Gly	Pro	Gln	Gly 1870	Gln	Thr	Leu	Thr	Ala 1875	Glu	Ser	Ile
Arg 1880	Glu	Pro	Pro	Val	Ser	Glu 1885	Pro	Ala	Val	Pro	Thr 1890	Ser	Ser	Ser
Ser 1895	Ser	Ile	Ala	Asn	Val	Ser 1900	Ser	Ala	Arg	Leu	Ala 1905	Glu	Ala	Glu
Ala 1910	Ala	Val	Leu	Ser	Val	Leu 1915	Ala	Asp	Lys	Thr	Gly 1920	Tyr	Asp	Ser
Ser 1925	Met	Ile	Glu	Met	Asp	Met 1930	Asp	Leu	Glu	Ser	Glu 1935	Leu	Gly	Val

ES 2 623 610 T3

Asp Ser 1940	Ile Lys Arg Val 1945	Glu Ile Met Ser Glu Val 1950	Gln Thr Leu
Leu Ser 1955	Val Glu Val Ser 1960	Asp Val Asp Ala Leu Ser 1965	Arg Thr Lys
Thr Val 1970	Gly Asp Val Ile 1975	Glu Ala Met Lys Leu Glu 1980	Leu Gly Glu
Ser Ser 1985	Ser Ile Glu Thr 1990	Leu Asn Cys Thr Glu Val 1995	Glu His Thr
Ser Tyr 2000	Lys Ser Val Lys 2005	Ala Ser Gly Cys Glu Asn 2010	Val Asp Thr
Arg Phe 2015	Ala Lys Val Val 2020	Gln Ile Ser Leu Pro Ser 2025	Lys Leu Lys
Ser Thr 2030	Val Ser His Asp 2035	Arg Pro Val Ile Val Val 2040	Asp Asp Gly
Thr Pro 2045	Leu Thr Thr Glu 2050	Leu Cys Lys Ile Leu Gly 2055	Gly Asn Ile
Val Val 2060	Leu Ser Tyr Gln 2065	Gly Lys Pro Ala Gly Pro 2070	Arg Gly Val
Glu Val 2075	Pro Asp Leu Ser 2080	Glu Glu Ala Leu Ile Gln 2085	Ala Leu Ala
Leu Ile 2090	Arg Ser Thr Tyr 2095	Gly Val Pro Ile Gly Phe 2100	Ile Cys Gln
Gln Val 2105	Ser Asn Val Ser 2110	Thr Lys Ala Gln Leu Cys 2115	Trp Ala Leu
Leu Ala 2120	Ala Lys His Leu 2125	Lys Lys Asp Leu Asn Ala 2130	Val Leu Pro
Asp Ser 2135	Arg Ser Phe Phe 2140	Val Gly Val Val Arg Leu 2145	Asn Gly Lys
Leu Gly 2150	Thr Phe Glu Asn 2155	Ile Ser Asp Phe Ser Lys 2160	Phe Asp Leu
Thr Lys 2165	Ala Leu Asp Tyr 2170	Gly Gln Arg Gly Ser Leu 2175	Leu Gly Leu

ES 2 623 610 T3

Cys Lys Ser Leu Asp Leu Glu Trp Glu Gln Val Phe Cys Arg Gly	2180	2185	2190
Ile Asp Leu Ala Cys Asp Leu Met Pro Leu Gln Ala Ala Arg Ile	2195	2200	2205
Leu Arg Asn Glu Leu Gln Cys Pro Asn Met Arg Leu Arg Glu Val	2210	2215	2220
Gly Tyr Asp Ile Ser Gly Ala Arg Tyr Thr Ile Ser Thr Asp Asp	2225	2230	2235
Leu Leu Cys Gly Pro Ser Lys Ala Lys Val Glu Ala Ala Asp Leu	2240	2245	2250
Phe Leu Val Thr Gly Gly Ala Arg Gly Ile Thr Pro His Cys Val	2255	2260	2265
Arg Glu Ile Ala Ser Arg Ser Pro Gly Thr Thr Phe Val Leu Val	2270	2275	2280
Gly Arg Ser Glu Met Ser Asp Glu Pro Asp Trp Ala Val Gly His	2285	2290	2295
Tyr Asn Lys Asp Leu Asp Gln Ser Thr Met Lys His Leu Lys Ala	2300	2305	2310
Thr His Ala Ala Gly Gly Val Lys Pro Thr Pro Lys Ala His Arg	2315	2320	2325
Ala Leu Val Asn Arg Val Thr Gly Ser Arg Glu Val Arg Glu Ser	2330	2335	2340
Leu Arg Ala Ile Gln Glu Ala Gly Ala Asn Val Glu Tyr Ile Ala	2345	2350	2355
Cys Asp Val Ser Asp Glu Asn Lys Val Arg Gln Leu Val Gln Arg	2360	2365	2370
Val Glu Gln Lys Tyr Gly Cys Glu Ile Thr Gly Ile Trp His Ala	2375	2380	2385
Ser Gly Val Leu Arg Asp Lys Leu Val Glu Gln Lys Thr Thr Asp	2390	2395	2400
Asp Phe Glu Ala Val Phe Gly Thr Lys Val Thr Gly Leu Val Asn	2405	2410	2415
Ile Val Ser Gln Val Asn Met Ser Lys Leu Arg His Phe Ile Leu			

ES 2 623 610 T3

2420		2425		2430
Phe Ser Ser Leu Ala Gly	Phe His Gly Asn Lys Gly	Gln Thr Asp		
2435	2440	2445		
Tyr Ala Ile Ala Asn Glu	Ala Leu Asn Lys Ile Ala	His Thr Leu		
2450	2455	2460		
Ser Ala Phe Leu Pro Lys	Leu Asn Ala Lys Val Leu	Asp Phe Gly		
2465	2470	2475		
Pro Trp Val Gly Ser Gly	Met Val Thr Glu Thr Leu	Glu Lys His		
2480	2485	2490		
Phe Lys Ala Met Gly Val	Gln Thr Ile Pro Leu Glu	Pro Gly Ala		
2495	2500	2505		
Arg Thr Val Ala Gln Ile	Ile Leu Ala Ser Ser Pro	Pro Gln Ser		
2510	2515	2520		
Leu Leu Gly Asn Trp Gly	Phe Pro Ala Thr Lys Pro	Leu Gln Arg		
2525	2530	2535		
Ser Asn Val Val Thr Gly	Thr Leu Ser Pro Glu Glu	Ile Glu Phe		
2540	2545	2550		
Ile Ala Asp His Lys Ile	Gln Gly Arg Lys Val Leu	Pro Met Met		
2555	2560	2565		
Ala Ala Ile Gly Phe Met	Ala Ser Ile Ala Glu Gly	Leu Tyr Pro		
2570	2575	2580		
Gly Tyr Asn Leu Gln Gly	Val Glu Asn Ala Gln Leu	Phe Gln Gly		
2585	2590	2595		
Leu Thr Ile Asn Gln Glu	Thr Lys Phe Gln Ile Thr	Leu Ile Glu		
2600	2605	2610		
Glu His Asn Ser Glu Glu	Asn Leu Asp Val Leu Thr	Ser Leu Gly		
2615	2620	2625		
Val Met Leu Glu Ser Gly	Lys Val Leu Pro Ala Tyr	Arg Cys Val		
2630	2635	2640		
Val Cys Leu Asn Thr Thr	Gln Gln Gln Pro Lys Leu	Ser Pro Lys		
2645	2650	2655		
Ile Leu Asn Leu Glu Val	Asp Pro Ala Cys Glu Val	Asn Pro Tyr		
2660	2665	2670		

Asp Gly Lys Ser Leu Phe His Gly Pro Leu Leu Gln Phe Val Gln
2675 2680 2685

Gln Val Leu His Ser Ser Thr Lys Gly Leu Val Ala Lys Cys Arg
2690 2695 2700

Ala Leu Pro Ile Lys Glu Ala Ile Arg Gly Pro Phe Ile Lys Gln
2705 2710 2715

Thr Leu His Asp Pro Ile Leu Asp Asp Val Ile Phe Gln Leu Met
2720 2725 2730

Leu Val Trp Cys Arg Asn Ala Leu Gly Ser Ala Ser Leu Pro Asn
2735 2740 2745

Arg Ile Glu Lys Met Ser Tyr Phe Gly Asn Val Ser Glu Gly Ser
2750 2755 2760

Thr Phe Phe Ala Ser Val Thr Pro Val Gly Pro Arg Val Pro Lys
2765 2770 2775

Asp Pro Val Ile Lys Met Gln Phe Leu Leu Gln Asp Glu Ser Gly
2780 2785 2790

Asn Thr Phe Ser Ser Gly Glu Gly Ser Val Val Leu Ser Asp Glu
2795 2800 2805

Leu Val Phe
2810

<210> 57

<211> 2059

5 <212> PRT

<213> *Schizochytrium* sp.

<400> 57

Met Ala Ala Arg Asn Val Ser Ala Ala His Glu Met His Asp Glu Lys
1 5 10 15

Arg Ile Ala Val Val Gly Met Ala Val Gln Tyr Ala Gly Cys Lys Thr
20 25 30

Lys Asp Glu Phe Trp Glu Val Leu Met Asn Gly Lys Val Glu Ser Lys
35 40 45

Val Ile Ser Asp Lys Arg Leu Gly Ser Asn Tyr Arg Ala Glu His Tyr
50 55 60

ES 2 623 610 T3

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

ES 2 623 610 T3

				325						330				335	
Thr	Pro	Gln	Gly	Asp	Arg	Val	Glu	Ile	Asp	Ala	Val	Lys	Ala	Cys	Phe
			340					345					350		
Glu	Gly	Lys	Val	Pro	Arg	Phe	Gly	Thr	Thr	Lys	Gly	Asn	Phe	Gly	His
		355					360					365			
Thr	Leu	Val	Ala	Ala	Gly	Phe	Ala	Gly	Met	Cys	Lys	Val	Leu	Leu	Ser
	370					375					380				
Met	Lys	His	Gly	Ile	Ile	Pro	Pro	Thr	Pro	Gly	Ile	Asp	Asp	Glu	Thr
385					390					395				400	
Lys	Met	Asp	Pro	Leu	Val	Val	Ser	Gly	Glu	Ala	Ile	Pro	Trp	Pro	Glu
				405					410					415	
Thr	Asn	Gly	Glu	Pro	Lys	Arg	Ala	Gly	Leu	Ser	Ala	Phe	Gly	Phe	Gly
			420					425					430		
Gly	Thr	Asn	Ala	His	Ala	Val	Phe	Glu	Glu	His	Asp	Pro	Ser	Asn	Ala
		435					440					445			
Ala	Cys	Thr	Gly	His	Asp	Ser	Ile	Ser	Ala	Leu	Ser	Ala	Arg	Cys	Gly
	450					455					460				
Gly	Glu	Ser	Asn	Met	Arg	Ile	Ala	Ile	Thr	Gly	Met	Asp	Ala	Thr	Phe
465					470					475				480	
Gly	Ala	Leu	Lys	Gly	Leu	Asp	Ala	Phe	Glu	Arg	Ala	Ile	Tyr	Thr	Gly
				485					490					495	
Ala	His	Gly	Ala	Ile	Pro	Leu	Pro	Glu	Lys	Arg	Trp	Arg	Phe	Leu	Gly
			500					505					510		
Lys	Asp	Lys	Asp	Phe	Leu	Asp	Leu	Cys	Gly	Val	Lys	Ala	Thr	Pro	His
		515					520					525			
Gly	Cys	Tyr	Ile	Glu	Asp	Val	Glu	Val	Asp	Phe	Gln	Arg	Leu	Arg	Thr
	530					535					540				
Pro	Met	Thr	Pro	Glu	Asp	Met	Leu	Leu	Pro	Gln	Gln	Leu	Leu	Ala	Val
545					550					555				560	
Thr	Thr	Ile	Asp	Arg	Ala	Ile	Leu	Asp	Ser	Gly	Met	Lys	Lys	Gly	Gly
				565					570					575	
Asn	Val	Ala	Val	Phe	Val	Gly	Leu	Gly	Thr	Asp	Leu	Glu	Leu	Tyr	Arg
			580					585					590		

ES 2 623 610 T3

His Arg Ala Arg Val Ala Leu Lys Glu Arg Val Arg Pro Glu Ala Ser
 595 600 605
 Lys Lys Leu Asn Asp Met Met Gln Tyr Ile Asn Asp Cys Gly Thr Ser
 610 615 620
 Thr Ser Tyr Thr Ser Tyr Ile Gly Asn Leu Val Ala Thr Arg Val Ser
 625 630 635 640
 Ser Gln Trp Gly Phe Thr Gly Pro Ser Phe Thr Ile Thr Glu Gly Asn
 645 650 655
 Asn Ser Val Tyr Arg Cys Ala Glu Leu Gly Lys Tyr Leu Leu Glu Thr
 660 665 670
 Gly Glu Val Asp Gly Val Val Val Ala Gly Val Asp Leu Cys Gly Ser
 675 680 685
 Ala Glu Asn Leu Tyr Val Lys Ser Arg Arg Phe Lys Val Ser Thr Ser
 690 695 700
 Asp Thr Pro Arg Ala Ser Phe Asp Ala Ala Ala Asp Gly Tyr Phe Val
 705 710 715 720
 Gly Glu Gly Cys Gly Ala Phe Val Leu Lys Arg Glu Thr Ser Cys Thr
 725 730 735
 Lys Asp Asp Arg Ile Tyr Ala Cys Met Asp Ala Ile Val Pro Gly Asn
 740 745 750
 Val Pro Ser Ala Cys Leu Arg Glu Ala Leu Asp Gln Ala Arg Val Lys
 755 760 765
 Pro Gly Asp Ile Glu Met Leu Glu Leu Ser Ala Asp Ser Ala Arg His
 770 775 780
 Leu Lys Asp Pro Ser Val Leu Pro Lys Glu Leu Thr Ala Glu Glu Glu
 785 790 795 800
 Ile Gly Gly Leu Gln Thr Ile Leu Arg Asp Asp Asp Lys Leu Pro Arg
 805 810 815
 Asn Val Ala Thr Gly Ser Val Lys Ala Thr Val Gly Asp Thr Gly Tyr
 820 825 830
 Ala Ser Gly Ala Ala Ser Leu Ile Lys Ala Ala Leu Cys Ile Tyr Asn
 835 840 845

ES 2 623 610 T3

Arg Tyr Leu Pro Ser Asn Gly Asp Asp Trp Asp Glu Pro Ala Pro Glu
 850 855 860
 Ala Pro Trp Asp Ser Thr Leu Phe Ala Cys Gln Thr Ser Arg Ala Trp
 865 870 875 880
 Leu Lys Asn Pro Gly Glu Arg Arg Tyr Ala Ala Val Ser Gly Val Ser
 885 890 895
 Glu Thr Arg Ser Cys Tyr Ser Val Leu Leu Ser Glu Ala Glu Gly His
 900 905 910
 Tyr Glu Arg Glu Asn Arg Ile Ser Leu Asp Glu Glu Ala Pro Lys Leu
 915 920 925
 Ile Val Leu Arg Ala Asp Ser His Glu Glu Ile Leu Gly Arg Leu Asp
 930 935 940
 Lys Ile Arg Glu Arg Phe Leu Gln Pro Thr Gly Ala Ala Pro Arg Glu
 945 950 955 960
 Ser Glu Leu Lys Ala Gln Ala Arg Arg Ile Phe Leu Glu Leu Leu Gly
 965 970 975
 Glu Thr Leu Ala Gln Asp Ala Ala Ser Ser Gly Ser Gln Lys Pro Leu
 980 985 990
 Ala Leu Ser Leu Val Ser Thr Pro Ser Lys Leu Gln Arg Glu Val Glu
 995 1000 1005
 Leu Ala Ala Lys Gly Ile Pro Arg Cys Leu Lys Met Arg Arg Asp
 1010 1015 1020
 Trp Ser Ser Pro Ala Gly Ser Arg Tyr Ala Pro Glu Pro Leu Ala
 1025 1030 1035
 Ser Asp Arg Val Ala Phe Met Tyr Gly Glu Gly Arg Ser Pro Tyr
 1040 1045 1050
 Tyr Gly Ile Thr Gln Asp Ile His Arg Ile Trp Pro Glu Leu His
 1055 1060 1065
 Glu Val Ile Asn Glu Lys Thr Asn Arg Leu Trp Ala Glu Gly Asp
 1070 1075 1080
 Arg Trp Val Met Pro Arg Ala Ser Phe Lys Ser Glu Leu Glu Ser
 1085 1090 1095

ES 2 623 610 T3

Gln Gln	Gln Glu Phe Asp Arg	Asn Met Ile Glu Met	Phe Arg Leu
1100	1105	1110	
Gly Ile	Leu Thr Ser Ile Ala	Phe Thr Asn Leu Ala	Arg Asp Val
1115	1120	1125	
Leu Asn	Ile Thr Pro Lys Ala	Ala Phe Gly Leu Ser	Leu Gly Glu
1130	1135	1140	
Ile Ser	Met Ile Phe Ala Phe	Ser Lys Lys Asn Gly	Leu Ile Ser
1145	1150	1155	
Asp Gln	Leu Thr Lys Asp Leu	Arg Glu Ser Asp Val	Trp Asn Lys
1160	1165	1170	
Ala Leu	Ala Val Glu Phe Asn	Ala Leu Arg Glu Ala	Trp Gly Ile
1175	1180	1185	
Pro Gln	Ser Val Pro Lys Asp	Glu Phe Trp Gln Gly	Tyr Ile Val
1190	1195	1200	
Arg Gly	Thr Lys Gln Asp Ile	Glu Ala Ala Ile Ala	Pro Asp Ser
1205	1210	1215	
Lys Tyr	Val Arg Leu Thr Ile	Ile Asn Asp Ala Asn	Thr Ala Leu
1220	1225	1230	
Ile Ser	Gly Lys Pro Asp Ala	Cys Lys Ala Ala Ile	Ala Arg Leu
1235	1240	1245	
Gly Gly	Asn Ile Pro Ala Leu	Pro Val Thr Gln Gly	Met Cys Gly
1250	1255	1260	
His Cys	Pro Glu Val Gly Pro	Tyr Thr Lys Asp Ile	Ala Lys Ile
1265	1270	1275	
His Ala	Asn Leu Glu Phe Pro	Val Val Asp Gly Leu	Asp Leu Trp
1280	1285	1290	
Thr Thr	Ile Asn Gln Lys Arg	Leu Val Pro Arg Ala	Thr Gly Ala
1295	1300	1305	
Lys Asp	Glu Trp Ala Pro Ser	Ser Phe Gly Glu Tyr	Ala Gly Gln
1310	1315	1320	
Leu Tyr	Glu Lys Gln Ala Asn	Phe Pro Gln Ile Val	Glu Thr Ile
1325	1330	1335	
Tyr Lys	Gln Asn Tyr Asp Val	Phe Val Glu Val Gly	Pro Asn Asn

ES 2 623 610 T3

1340		1345		1350
His Arg Ser Thr Ala Val	Arg Thr Thr Leu Gly Pro	Gln Arg Asn		
1355	1360	1365		
His Leu Ala Gly Ala Ile	Asp Lys Gln Asn Glu Asp	Ala Trp Thr		
1370	1375	1380		
Thr Ile Val Lys Leu Val	Ala Ser Leu Lys Ala His	Leu Val Pro		
1385	1390	1395		
Gly Val Thr Ile Ser Pro	Leu Tyr His Ser Lys Leu	Val Ala Glu		
1400	1405	1410		
Ala Gln Ala Cys Tyr Ala	Ala Leu Cys Lys Gly Glu	Lys Pro Lys		
1415	1420	1425		
Lys Asn Lys Phe Val Arg	Lys Ile Gln Leu Asn Gly	Arg Phe Asn		
1430	1435	1440		
Ser Lys Ala Asp Pro Ile	Ser Ser Ala Asp Leu Ala	Ser Phe Pro		
1445	1450	1455		
Pro Ala Asp Pro Ala Ile	Glu Ala Ala Ile Ser Ser	Arg Ile Met		
1460	1465	1470		
Lys Pro Val Ala Pro Lys	Phe Tyr Ala Arg Leu Asn	Ile Asp Glu		
1475	1480	1485		
Gln Asp Glu Thr Arg Asp	Pro Ile Leu Asn Lys Asp	Asn Ala Pro		
1490	1495	1500		
Ser Ser Ser Ser Ser	Ser Ser Ser Ser Ser	Ser Ser Ser		
1505	1510	1515		
Pro Ser Pro Ala Pro Ser	Ala Pro Val Gln Lys Lys	Ala Ala Pro		
1520	1525	1530		
Ala Ala Glu Thr Lys Ala	Val Ala Ser Ala Asp Ala	Leu Arg Ser		
1535	1540	1545		
Ala Leu Leu Asp Leu Asp	Ser Met Leu Ala Leu Ser	Ser Ala Ser		
1550	1555	1560		
Ala Ser Gly Asn Leu Val	Glu Thr Ala Pro Ser Asp	Ala Ser Val		
1565	1570	1575		
Ile Val Pro Pro Cys Asn	Ile Ala Asp Leu Gly Ser	Arg Ala Phe		
1580	1585	1590		

ES 2 623 610 T3

Met	Lys	Thr	Tyr	Gly	Val	Ser	Ala	Pro	Leu	Tyr	Thr	Gly	Ala	Met
1595						1600					1605			
Ala	Lys	Gly	Ile	Ala	Ser	Ala	Asp	Leu	Val	Ile	Ala	Ala	Gly	Arg
1610						1615					1620			
Gln	Gly	Ile	Leu	Ala	Ser	Phe	Gly	Ala	Gly	Gly	Leu	Pro	Met	Gln
1625						1630					1635			
Val	Val	Arg	Glu	Ser	Ile	Glu	Lys	Ile	Gln	Ala	Ala	Leu	Pro	Asn
1640						1645					1650			
Gly	Pro	Tyr	Ala	Val	Asn	Leu	Ile	His	Ser	Pro	Phe	Asp	Ser	Asn
1655						1660					1665			
Leu	Glu	Lys	Gly	Asn	Val	Asp	Leu	Phe	Leu	Glu	Lys	Gly	Val	Thr
1670						1675					1680			
Phe	Val	Glu	Ala	Ser	Ala	Phe	Met	Thr	Leu	Thr	Pro	Gln	Val	Val
1685						1690					1695			
Arg	Tyr	Arg	Ala	Ala	Gly	Leu	Thr	Arg	Asn	Ala	Asp	Gly	Ser	Val
1700						1705					1710			
Asn	Ile	Arg	Asn	Arg	Ile	Ile	Gly	Lys	Val	Ser	Arg	Thr	Glu	Leu
1715						1720					1725			
Ala	Glu	Met	Phe	Met	Arg	Pro	Ala	Pro	Glu	His	Leu	Leu	Gln	Lys
1730						1735					1740			
Leu	Ile	Ala	Ser	Gly	Glu	Ile	Asn	Gln	Glu	Gln	Ala	Glu	Leu	Ala
1745						1750					1755			
Arg	Arg	Val	Pro	Val	Ala	Asp	Asp	Ile	Ala	Val	Glu	Ala	Asp	Ser
1760						1765					1770			
Gly	Gly	His	Thr	Asp	Asn	Arg	Pro	Ile	His	Val	Ile	Leu	Pro	Leu
1775						1780					1785			
Ile	Ile	Asn	Leu	Arg	Asp	Arg	Leu	His	Arg	Glu	Cys	Gly	Tyr	Pro
1790						1795					1800			
Ala	Asn	Leu	Arg	Val	Arg	Val	Gly	Ala	Gly	Gly	Gly	Ile	Gly	Cys
1805						1810					1815			
Pro	Gln	Ala	Ala	Leu	Ala	Thr	Phe	Asn	Met	Gly	Ala	Ser	Phe	Ile
1820						1825					1830			

ES 2 623 610 T3

Val Thr Gly Thr Val Asn Gln Val Ala Lys Gln Ser Gly Thr Cys 1835 1840 1845
Asp Asn Val Arg Lys Gln Leu Ala Lys Ala Thr Tyr Ser Asp Val 1850 1855 1860
Cys Met Ala Pro Ala Ala Asp Met Phe Glu Glu Gly Val Lys Leu 1865 1870 1875
Gln Val Leu Lys Lys Gly Thr Met Phe Pro Ser Arg Ala Asn Lys 1880 1885 1890
Leu Tyr Glu Leu Phe Cys Lys Tyr Asp Ser Phe Glu Ser Met Pro 1895 1900 1905
Pro Ala Glu Leu Ala Arg Val Glu Lys Arg Ile Phe Ser Arg Ala 1910 1915 1920
Leu Glu Glu Val Trp Asp Glu Thr Lys Asn Phe Tyr Ile Asn Arg 1925 1930 1935
Leu His Asn Pro Glu Lys Ile Gln Arg Ala Glu Arg Asp Pro Lys 1940 1945 1950
Leu Lys Met Ser Leu Cys Phe Arg Trp Tyr Leu Ser Leu Ala Ser 1955 1960 1965
Arg Trp Ala Asn Thr Gly Ala Ser Asp Arg Val Met Asp Tyr Gln 1970 1975 1980
Val Trp Cys Gly Pro Ala Ile Gly Ser Phe Asn Asp Phe Ile Lys 1985 1990 1995
Gly Thr Tyr Leu Asp Pro Ala Val Ala Asn Glu Tyr Pro Cys Val 2000 2005 2010
Val Gln Ile Asn Lys Gln Ile Leu Arg Gly Ala Cys Phe Leu Arg 2015 2020 2025
Arg Leu Glu Ile Leu Arg Asn Ala Arg Leu Ser Asp Gly Ala Ala 2030 2035 2040
Ala Leu Val Ala Ser Ile Asp Asp Thr Tyr Val Pro Ala Glu Lys 2045 2050 2055
Leu

<210> 58
<211> 1935
<212> PRT
<213> *Thraustochytrium* sp.

5

<400> 58

ES 2 623 610 T3

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

Ser Thr Phe Gln Ala Met Pro Val Gly Ala Asp Gly Val Ser Leu Pro
 245 250 255
 Leu His Lys Thr Ser Ala Gly Leu Thr Pro Gly Glu Gly Gly Ser Ile
 260 265 270
 Met Val Leu Lys Arg Leu Lys Asp Ala Ile Arg Asp Gly Asn His Ile
 275 280 285
 Tyr Gly Val Leu Leu Glu Ala Asn Leu Ser Asn Ala Gly Cys Gly Leu
 290 295 300
 Pro Leu Ser Pro His Leu Pro Ser Glu Glu Ser Cys Ile Arg Asp Thr
 305 310 315 320
 Tyr Arg Arg Ala Gly Val Ala Ala Asp Gln Ser Ile Gln Tyr Ile Glu
 325 330 335
 Cys His Ala Thr Gly Thr Pro Arg Gly Asp Val Val Glu Ile Glu Ala
 340 345 350
 Val Glu Arg Val Phe Lys Lys Asn Val Pro Arg Leu Gly Ser Thr Lys
 355 360 365
 Gly Asn Phe Gly His Ser Leu Val Ala Ala Gly Phe Ala Gly Met Ala
 370 375 380
 Lys Leu Leu Leu Ala Met Glu His Gly Val Ile Pro Pro Thr Pro Gly
 385 390 395 400
 Leu Asp Ala Ser Asn Gln Ala Ser Glu His Val Val Thr Lys Ala Ile
 405 410 415
 Thr Trp Pro Glu Thr His Gly Ala Pro Lys Arg Ala Gly Leu Ser Ala
 420 425 430
 Phe Gly Phe Gly Gly Thr Asn Ala His Ala Leu Phe Glu Glu Phe Asn
 435 440 445
 Ala Glu Gly Ile Ser Tyr Arg Pro Gly Lys Pro Pro Val Glu Ser Asn
 450 455 460
 Thr Arg Pro Ser Val Val Ile Thr Gly Met Asp Cys Thr Phe Gly Ser
 465 470 475 480
 Leu Glu Gly Ile Asp Ala Phe Glu Thr Ala Leu Tyr Glu Gly Arg Asp
 485 490 495

ES 2 623 610 T3

Ala	Ala	Arg	Asp	Leu	Pro	Ala	Lys	Arg	Trp	Arg	Phe	Leu	Gly	Glu	Asp	500	505	510
Leu	Glu	Phe	Leu	Arg	Ala	Ile	Arg	Leu	Lys	Glu	Lys	Pro	Arg	Gly	Cys	515	520	525
Phe	Val	Glu	Ser	Val	Asp	Val	Asn	Phe	Arg	Arg	Leu	Lys	Thr	Pro	Leu	530	535	540
Thr	Pro	Glu	Asp	Met	Leu	Arg	Pro	Gln	Gln	Leu	Leu	Ala	Val	Ser	Thr	545	550	555
Met	Asp	Arg	Ala	Ile	Ile	Asp	Ala	Gly	Leu	Lys	Lys	Gly	Gln	His	Val	565	570	575
Ala	Val	Leu	Val	Gly	Leu	Gly	Thr	Asp	Leu	Glu	Leu	Tyr	Arg	His	Arg	580	585	590
Ala	Arg	Val	Ala	Leu	Lys	Glu	Val	Leu	His	Pro	Ser	Leu	Lys	Ser	Asp	595	600	605
Thr	Ala	Ile	Leu	Gln	Lys	Ile	Met	Gln	Tyr	Val	Asn	Asp	Ala	Gly	Thr	610	615	620
Ser	Thr	Ser	Tyr	Thr	Ser	Tyr	Ile	Gly	Asn	Leu	Val	Ala	Thr	Arg	Ile	625	630	635
Ser	Ser	Gln	Trp	Gly	Phe	Thr	Gly	Pro	Ser	Phe	Thr	Val	Thr	Glu	Gly	645	650	655
Asn	Asn	Ser	Val	Tyr	Arg	Cys	Ala	Gln	Leu	Ala	Lys	Asp	Met	Leu	Gln	660	665	670
Val	Asn	Arg	Val	Asp	Ala	Val	Val	Ile	Ala	Gly	Val	Asp	Leu	Asn	Gly	675	680	685
Ser	Ala	Glu	Ser	Phe	Phe	Val	Arg	Ala	Asn	Arg	Gln	Lys	Ile	Ser	Lys	690	695	700
Leu	Ser	His	Pro	Cys	Ala	Ser	Phe	Asp	Arg	Asp	Ala	Asp	Gly	Phe	Phe	705	710	715
Ala	Gly	Glu	Gly	Cys	Gly	Ala	Leu	Val	Phe	Lys	Arg	Leu	Glu	Asp	Cys	725	730	735
Ala	Pro	Gln	Glu	Lys	Ile	Tyr	Ala	Ser	Ile	Asp	Ser	Ile	Ala	Ile	Asp	740	745	750
Lys	Glu	Pro	Thr	Ser	Ser	Ala	Val	Lys	Ala	Val	Tyr	Gln	Ser	Asp	Ser			

ES 2 623 610 T3

755		760		765
Ser Leu Ser Asp Ile Glu	Leu Leu Glu Ile Ser Gly Asp Ser Lys Arg			
770	775	780		
Phe Ala Ala Phe Glu Gly Ala Val Glu Ile Gln Ser Ser Val Glu Ala				
785	790	795		800
Gln Leu Lys Gly Leu Ser Lys Val Leu Glu Pro Ala Lys Gly Gln Gly				
	805	810		815
Val Ala Val Gly Ser Thr Arg Ala Thr Val Gly Asp Ile Gly Tyr Ala				
	820	825		830
Thr Gly Ala Ala Ser Leu Ile Lys Thr Ala Leu Cys Leu Tyr Asn Arg				
	835	840		845
Tyr Leu Pro Ala Leu Ala Asn Trp Ser Gly Pro Cys Glu Gln Ser Ala				
	850	855		860
Trp Gly Ser Asn Met Phe Val Cys His Glu Thr Arg Pro Trp Met Lys				
	865	870		875
Asn Gln Asn Glu Lys Arg Cys Ala Leu Ile Ser Gly Thr Asp Pro Ser				
	885	890		895
His Thr Cys Phe Ser Leu Val Leu Ser Asp Thr Gly Cys Tyr Glu Glu				
	900	905		910
His Asn Arg Thr Cys Phe Asp Val Gln Ala Pro Gln Leu Val Leu Ile				
	915	920		925
His Gly Phe Asp Gly Lys Thr Ile Val Arg Arg Leu Glu Gly Tyr Leu				
	930	935		940
Leu Glu Leu Val Glu Gly His Ala Ser Pro Ser Glu Tyr Phe His Lys				
	945	950		955
Leu Ile Gly Gln Ser Leu Leu Glu Asn Ser Lys Glu Ser Lys Leu Thr				
	965	970		975
Leu Ser Leu Val Cys Asn Pro Asn Gln Leu Gln Lys Glu Leu Met Leu				
	980	985		990
Ala Ile Lys Gly Val Gln Arg Ser Met Leu Thr Gly Lys Asp Trp Val				
	995	1000		1005
Ser Pro Ser Gly Ser Cys Phe Ala Pro Asn Pro Leu Ser Ser Ala				
	1010	1015		1020

ES 2 623 610 T3

Lys Val	Ala Phe Met Tyr Gly	Glu Gly Arg Ser Pro	Tyr Cys Gly
1025	1030	1035	
Val Gly	Leu Gly Leu His Arg	Leu Trp Pro Gly Leu	His Glu Asn
1040	1045	1050	
Val Asn	Asn Lys Thr Val Asp	Leu Trp Thr Glu Gly	Asp Gly Trp
1055	1060	1065	
Leu Tyr	Pro Arg Thr Leu Thr	Arg Glu Glu His Thr	Lys Ala Ile
1070	1075	1080	
Glu Ser	Phe Asn Ala Asn Gln	Ile Glu Met Phe Arg	Ala Gly Ile
1085	1090	1095	
Phe Ile	Ser Met Cys Gln Thr	Asp Tyr Val Met Asn	Val Leu Gly
1100	1105	1110	
Val Gln	Pro Lys Ala Gly Phe	Gly Leu Ser Leu Gly	Glu Ile Ser
1115	1120	1125	
Met Leu	Phe Ala Met Ser Lys	Glu Asn Cys Arg Gln	Ser Gln Glu
1130	1135	1140	
Met Thr	Asn Arg Leu Arg Gly	Ser Pro Val Trp Ser	Asn Glu Leu
1145	1150	1155	
Ala Ile	Asn Phe Asn Ala Ile	Arg Lys Leu Trp Lys	Ile Pro Arg
1160	1165	1170	
Gly Ala	Pro Leu Glu Ser Phe	Trp Gln Gly Tyr Leu	Val His Gly
1175	1180	1185	
Thr Arg	Glu Glu Val Glu His	Ala Ile Gly Leu Ser	Glu Pro Tyr
1190	1195	1200	
Val Arg	Leu Leu Ile Val Asn	Asp Ser Arg Ser Ala	Leu Ile Ala
1205	1210	1215	
Gly Lys	Pro Asp Ala Cys Gln	Ala Val Ile Ser Arg	Leu Asn Ser
1220	1225	1230	
Lys Phe	Pro Ser Leu Pro Val	Lys Gln Gly Met Ile	Gly His Cys
1235	1240	1245	
Pro Glu	Val Arg Ala Phe Ile	Lys Asp Ile Gly Tyr	Ile His Glu
1250	1255	1260	

Thr	Leu	Arg	Ile	Ser	Asn	Asp	Tyr	Ser	Asp	Cys	Gln	Leu	Phe	Ser
1265						1270					1275			
Ala	Val	Thr	Lys	Gly	Ala	Leu	Asp	Ser	Ser	Thr	Met	Glu	Ile	Lys
1280						1285					1290			
His	Phe	Val	Gly	Glu	Val	Tyr	Ser	Arg	Ile	Ala	Asp	Phe	Pro	Gln
1295						1300					1305			
Ile	Val	Asn	Thr	Val	His	Ser	Ala	Gly	Tyr	Asp	Val	Phe	Leu	Glu
1310						1315					1320			
Leu	Gly	Cys	Asp	Ala	Ser	Arg	Ser	Ala	Ala	Val	Gln	Asn	Ile	Leu
1325						1330					1335			
Gly	Gly	Gln	Gly	Lys	Phe	Leu	Ser	Thr	Ala	Ile	Asp	Lys	Lys	Gly
1340						1345					1350			
His	Ser	Ala	Trp	Ser	Gln	Val	Leu	Arg	Ala	Thr	Ala	Ser	Leu	Ala
1355						1360					1365			
Ala	His	Arg	Val	Pro	Gly	Ile	Ser	Ile	Leu	Asp	Leu	Phe	His	Pro
1370						1375					1380			
Asn	Phe	Arg	Glu	Met	Cys	Cys	Thr	Met	Ala	Thr	Thr	Pro	Lys	Val
1385						1390					1395			
Glu	Asp	Lys	Phe	Leu	Arg	Thr	Ile	Gln	Ile	Asn	Gly	Arg	Phe	Glu
1400						1405					1410			
Lys	Glu	Met	Ile	His	Leu	Glu	Asp	Thr	Thr	Leu	Ser	Cys	Leu	Pro
1415						1420					1425			
Ala	Pro	Ser	Glu	Ala	Asn	Ile	Ala	Ala	Ile	Gln	Ser	Arg	Ser	Ile
1430						1435					1440			
Arg	Ser	Ala	Ala	Ala	Arg	Ser	Gly	Gln	Ser	His	Asp	Cys	Ala	Ser
1445						1450					1455			
His	Ser	His	Glu	Glu	Asn	Lys	Asp	Ser	Cys	Pro	Glu	Lys	Leu	Lys
1460						1465					1470			
Leu	Asp	Ser	Val	Ser	Val	Ala	Ile	Asn	Phe	Asp	Asn	Asp	Asp	Arg
1475						1480					1485			
Ile	Gln	Leu	Gly	His	Ala	Gly	Phe	Arg	Glu	Met	Tyr	Asn	Thr	Arg
1490						1495					1500			

ES 2 623 610 T3

Tyr	Ser	Leu	Tyr	Thr	Gly	Ala	Met	Ala	Lys	Gly	Ile	Ala	Ser	Ala	
1505						1510					1515				
Asp	Leu	Val	Ile	Ala	Ala	Gly	Lys	Glu	Gly	Ile	Leu	Ala	Ser	Tyr	
1520						1525					1530				
Gly	Ala	Gly	Gly	Leu	Pro	Leu	Ala	Thr	Val	Arg	Lys	Gly	Ile	Asp	
1535						1540					1545				
Lys	Ile	Gln	Gln	Ala	Leu	Pro	Ser	Gly	Pro	Tyr	Ala	Val	Asn	Leu	
1550						1555					1560				
Ile	His	Ser	Pro	Phe	Asp	Gly	Asn	Leu	Glu	Gln	Gly	Asn	Val	Asp	
1565						1570					1575				
Leu	Phe	Leu	Glu	Lys	Asn	Val	Arg	Val	Ala	Glu	Cys	Ser	Ala	Phe	
1580						1585					1590				
Thr	Thr	Leu	Thr	Val	Pro	Val	Val	His	Tyr	Arg	Ala	Ala	Gly	Leu	
1595						1600					1605				
Val	Arg	Arg	Gln	Asp	Gly	Ser	Ile	Leu	Ile	Lys	Asn	Arg	Ile	Ile	
1610						1615					1620				
Ala	Lys	Val	Ser	Arg	Thr	Glu	Leu	Ala	Glu	Met	Phe	Leu	Arg	Pro	
1625						1630					1635				
Ala	Pro	Gln	Ile	Ile	Leu	Glu	Lys	Leu	Val	Ala	Ala	Glu	Ile	Ile	
1640						1645					1650				
Ser	Ser	Asp	Gln	Ala	Arg	Met	Ala	Ala	Lys	Val	Pro	Met	Ala	Asp	
1655						1660					1665				
Asp	Ile	Ala	Val	Glu	Ala	Asp	Ser	Gly	Gly	His	Thr	Asp	Asn	Arg	
1670						1675					1680				
Pro	Met	His	Val	Ile	Leu	Pro	Leu	Ile	Ile	Gln	Leu	Arg	Asn	Thr	
1685						1690					1695				
Ile	Leu	Ala	Glu	Tyr	Gly	Cys	Ala	Thr	Ala	Phe	Arg	Thr	Arg	Ile	
1700						1705					1710				
Gly	Ala	Gly	Gly	Gly	Ile	Gly	Cys	Pro	Ser	Ala	Ala	Leu	Ala	Ala	
1715						1720					1725				
Phe	Asp	Met	Gly	Ala	Ser	Phe	Val	Val	Thr	Gly	Ser	Ile	Asn	Gln	
1730						1735					1740				
Ile	Cys	Arg	Glu	Ala	Gly	Thr	Cys	Asp	Thr	Val	Arg	Glu	Leu	Leu	

1745		1750		1755
Ala Asn Ser Ser Tyr Ser Asp Val Thr Met Ala Pro Ala Ala Asp				
1760		1765		1770
Met Phe Asp Gln Gly Val Lys Leu Gln Val Leu Lys Arg Gly Thr				
1775		1780		1785
Met Phe Pro Ser Arg Ala Asn Lys Leu Arg Lys Leu Phe Val Asn				
1790		1795		1800
Tyr Glu Ser Leu Glu Thr Leu Pro Ser Lys Glu Leu Lys Tyr Leu				
1805		1810		1815
Glu Asn Ile Ile Phe Lys Gln Ala Val Asp Gln Val Trp Glu Glu				
1820		1825		1830
Thr Lys Arg Phe Tyr Cys Glu Lys Leu Asn Asn Pro Asp Lys Ile				
1835		1840		1845
Ala Arg Ala Met Lys Asp Pro Lys Leu Lys Met Ser Leu Cys Phe				
1850		1855		1860
Arg Trp Tyr Leu Ser Lys Ser Ser Gly Trp Ala Asn Ala Gly Ile				
1865		1870		1875
Lys Ser Arg Ala Leu Asp Tyr Gln Ile Trp Cys Gly Pro Ala Met				
1880		1885		1890
Gly Ser Phe Asn Asn Phe Ala Ser Gly Thr Ser Leu Asp Trp Lys				
1895		1900		1905
Val Thr Gly Val Phe Pro Gly Val Ala Glu Val Asn Met Ala Ile				
1910		1915		1920
Leu Asp Gly Ala Arg Glu Leu Ala Ala Lys Arg Asn				
1925		1930		1935

<210> 59

<211> 2037

<212> PRT

<213> *Thraustochytrium aureum*

<400> 59

ES 2 623 610 T3

Gln Ala Ile Gly His Arg Ala Ala Arg Trp Ser Cys Arg Ser Lys Ser
 1 5 10 15

Lys Ala Arg Gly His Lys Ala Gln Lys Glu Met Asn Gln Gly Gly Arg
 20 25 30

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

Val Ser Met Pro Phe His Arg Asp His Leu Leu Ser Gly Glu Ala Asp
 290 295 300

Leu Met Leu Cys Gly Ala Thr Cys Phe Pro Glu Pro Phe Phe Ile Leu
 305 310 315 320

Thr Gly Phe Ser Thr Phe His Ala Met Pro Val Gly Glu Asn Gly Val
 325 330 335

Ser Met Pro Phe His Arg Gln Pro Ser Glu Glu Ala Cys Leu Lys Ala
 340 345 350

Thr Tyr Glu Leu Val Gly Val Pro Pro Arg Asp Val Gln Tyr Val Glu
 355 360 365

Cys His Ala Thr Gly Thr Pro Gln Gly Asp Thr Val Glu Leu Gln Ala
 370 375 380

Val Lys Ala Cys Phe Glu Gly Ala Ser Pro Arg Ile Gly Ser Thr Lys
 385 390 395 400

Gly Asn Phe Gly His Thr Leu Val Ala Ala Gly Phe Ala Gly Met Cys
 405 410 415

Lys Val Leu Leu Ala Met Glu Arg Gly Val Ile Pro Pro Thr Pro Gly
 420 425 430

Val Asp Ser Gly Thr Gln Ile Asp Pro Leu Val Val Thr Ala Ala Leu
 435 440 445

Pro Trp Pro Asp Thr Arg Gly Gly Pro Lys Arg Ala Gly Leu Ser Ala
 450 455 460

Phe Gly Phe Gly Gly Thr Asn Ala His Ala Val Phe Glu Glu His Ile
 465 470 475 480

Pro Ser Arg Ala Pro Pro Ala Val Leu Cys Gln Pro Arg Leu Gly Ser
 485 490 495

Gly Pro Asn Arg Lys Leu Ala Ile Val Gly Met Asp Ala Thr Phe Gly
 500 505 510

Ser Leu Lys Gly Leu Ser Ala Leu Glu Ala Ala Leu Tyr Glu Ala Arg
 515 520 525

His Ala Ala Arg Pro Leu Pro Ala Lys Arg Trp Arg Phe Leu Gly Gly
 530 535 540

Asp Glu Ser Phe Leu His Glu Ile Gly Leu Glu Cys Ser Pro His Gly
 545 550 555 560
 Cys Tyr Ile Glu Asp Val Asp Val Asp Phe Lys Arg Leu Arg Thr Pro
 565 570 575
 Met Val Pro Glu Asp Leu Leu Arg Pro Gln Gln Leu Leu Ala Val Ser
 580 585 590
 Thr Ile Asp Lys Ala Ile Leu Asp Ser Gly Leu Ala Lys Gly Gly Asn
 595 600 605
 Val Ala Val Leu Val Gly Leu Gly Thr Asp Leu Glu Leu Tyr Arg His
 610 615 620
 Arg Ala Arg Val Ala Leu Lys Glu Arg Leu Gln Gly Leu Val Arg Ser
 625 630 635 640
 Ala Glu Gly Gly Ala Leu Thr Ser Arg Leu Met Asn Tyr Ile Asn Asp
 645 650 655
 Ser Gly Thr Ser Thr Ser Tyr Thr Ser Tyr Ile Gly Asn Leu Val Ala
 660 665 670
 Thr Arg Val Ser Ser Gln Trp Gly Phe Thr Gly Pro Ser Phe Thr Val
 675 680 685
 Thr Glu Gly Ala Asn Ser Val His Arg Cys Ala Gln Leu Ala Lys Tyr
 690 695 700
 Met Leu Asp Arg Gly Glu Val Asp Ala Val Val Val Ala Gly Val Asp
 705 710 715 720
 Leu Cys Gly Ser Ala Glu Ala Phe Phe Val Arg Ser Arg Arg Met Gln
 725 730 735
 Ile Ser Lys Ser Gln Arg Pro Ala Ala Pro Phe Asp Arg Ala Ala Asp
 740 745 750
 Gly Phe Phe Ala Gly Glu Gly Cys Gly Ala Leu Val Phe Lys Arg Leu
 755 760 765
 Thr Asp Cys Val Ser Gly Glu Arg Ile Tyr Ala Ser Leu Asp Ser Val
 770 775 780
 Val Val Ala Thr Thr Pro Arg Ala Ala Leu Arg Ala Ala Ala Gly Ser
 785 790 795 800
 Ala Arg Val Asp Pro Ala Ser Ile Asp Met Val Glu Leu Ser Ala Asp

ES 2 623 610 T3

805										810					815				
Ser	His	Arg	Phe	Val	Arg	Ala	Pro	Gly	Thr	Val	Ala	Gln	Pro	Leu	Thr				
			820					825					830						
Ala	Glu	Val	Glu	Val	Gly	Ala	Val	Arg	Glu	Val	Ile	Gly	Thr	Ala	Gly				
		835					840					845							
Arg	Gly	Ser	Arg	Ser	Val	Ala	Val	Gly	Ser	Val	Arg	Ala	Asn	Val	Gly				
	850					855					860								
Asp	Ala	Gly	Phe	Ala	Ser	Gly	Ala	Ala	Ala	Leu	Val	Lys	Thr	Ala	Leu				
865					870					875					880				
Cys	Leu	His	Asn	Arg	Tyr	Leu	Ala	Ala	Thr	Pro	Gly	Trp	Asp	Ala	Pro				
				885					890					895					
Ala	Ala	Gly	Val	Asp	Phe	Gly	Ala	Glu	Leu	Tyr	Val	Cys	Arg	Glu	Ser				
			900					905						910					
Arg	Ala	Trp	Val	Lys	Asn	Ala	Gly	Val	Ala	Arg	His	Ala	Ala	Ile	Ser				
		915					920						925						
Gly	Val	Asp	Glu	Gly	Gly	Ser	Cys	Tyr	Gly	Leu	Val	Leu	Ser	Asp	Val				
	930					935					940								
Pro	Gly	Gln	Tyr	Glu	Thr	Gly	Asn	Arg	Ile	Ser	Leu	Gln	Ala	Glu	Ser				
945					950					955					960				
Pro	Lys	Leu	Leu	Leu	Leu	Ser	Ala	Pro	Asp	His	Ala	Ala	Leu	Leu	Asp				
				965					970					975					
Lys	Val	Ala	Ala	Glu	Leu	Ala	Ala	Leu	Glu	Gln	Ala	Asp	Gly	Leu	Ser				
		980					985						990						
Ala	Ala	Ala	Ala	Ala	Val	Asp	Arg	Leu	Leu	Gly	Glu	Ser	Leu	Val	Gly				
		995					1000						1005						
Cys	Ala	Ala	Gly	Ser	Gly	Gly	Leu	Thr	Leu	Cys	Leu	Val	Ala	Ser					
	1010					1015					1020								
Pro	Ala	Ser	Leu	His	Lys	Glu	Leu	Ala	Leu	Ala	His	Arg	Gly	Ile					
	1025					1030					1035								
Pro	Arg	Cys	Ile	Lys	Ala	Arg	Arg	Asp	Trp	Ala	Ser	Pro	Ala	Gly					
	1040					1045					1050								
Ser	Tyr	Phe	Ala	Pro	Glu	Pro	Ile	Ala	Ser	Asp	Arg	Val	Ala	Phe					
	1055					1060					1065								

Met	Tyr	Gly	Glu	Gly	Arg	Ser	Pro	Tyr	Cys	Gly	Val	Gly	Arg	Asp
1070						1075					1080			
Leu	His	Arg	Ile	Trp	Pro	Ala	Leu	His	Glu	Arg	Val	Asn	Ala	Lys
1085						1090					1095			
Thr	Val	Asn	Leu	Trp	Gly	Asp	Gly	Asp	Ala	Trp	Leu	Leu	Pro	Arg
1100						1105					1110			
Ala	Thr	Ser	Ala	Glu	Glu	Glu	Glu	Gln	Leu	Cys	Arg	Asn	Phe	Asp
1115						1120					1125			
Ser	Asn	Gln	Val	Glu	Met	Phe	Arg	Thr	Gly	Val	Tyr	Ile	Ser	Met
1130						1135					1140			
Cys	Leu	Thr	Asp	Leu	Ala	Arg	Ser	Leu	Ile	Gly	Leu	Gly	Pro	Lys
1145						1150					1155			
Ala	Ser	Phe	Gly	Leu	Ser	Leu	Gly	Glu	Val	Ser	Met	Leu	Phe	Ala
1160						1165					1170			
Leu	Ser	Glu	Ser	Asn	Cys	Arg	Leu	Ser	Glu	Glu	Met	Thr	Arg	Arg
1175						1180					1185			
Leu	Arg	Ala	Ser	Pro	Val	Trp	Asn	Ser	Glu	Leu	Ala	Val	Glu	Phe
1190						1195					1200			
Asn	Ala	Leu	Arg	Lys	Leu	Trp	Gly	Val	Ala	Pro	Gly	Ala	Pro	Val
1205						1210					1215			
Asp	Ser	Phe	Trp	Gln	Gly	Tyr	Val	Val	Arg	Ala	Thr	Arg	Ala	Gln
1220						1225					1230			
Val	Glu	Gln	Ala	Ile	Gly	Glu	Asp	Asn	Gln	Phe	Val	Arg	Leu	Leu
1235						1240					1245			
Ile	Val	Asn	Asp	Ser	Gln	Ser	Val	Leu	Ile	Ala	Gly	Lys	Pro	Ala
1250						1255					1260			
Ala	Cys	Glu	Ala	Val	Ile	Ala	Arg	Ile	Gly	Ser	Ile	Leu	Pro	Pro
1265						1270					1275			
Leu	Gln	Val	Ser	Gln	Gly	Met	Val	Gly	His	Cys	Ala	Glu	Val	Leu
1280						1285					1290			
Pro	Tyr	Thr	Ser	Glu	Ile	Gly	Arg	Ile	His	Asn	Met	Leu	Arg	Phe
1295						1300					1305			

ES 2 623 610 T3

Pro	Ser	Gln	Asp	Glu	Thr	Gly	Gly	Cys	Lys	Met	Tyr	Ser	Ser	Val
1310						1315					1320			
Ser	Asn	Ser	Arg	Ile	Gly	Pro	Val	Glu	Glu	Ser	Gln	Met	Gly	Pro
1325						1330					1335			
Gly	Thr	Glu	Leu	Val	Phe	Ser	Pro	Ser	Met	Glu	Asp	Phe	Val	Ala
1340						1345					1350			
Gln	Leu	Tyr	Ser	Arg	Val	Ala	Asp	Phe	Pro	Ala	Ile	Thr	Glu	Ala
1355						1360					1365			
Val	Tyr	Gln	Gln	Gly	His	Asp	Val	Phe	Val	Glu	Val	Gly	Pro	Asp
1370						1375					1380			
His	Ser	Arg	Ser	Ala	Ala	Val	Arg	Ser	Thr	Leu	Gly	Pro	Thr	Arg
1385						1390					1395			
Arg	His	Ile	Ala	Val	Ala	Met	Asp	Arg	Lys	Gly	Glu	Ser	Ala	Trp
1400						1405					1410			
Ser	Gln	Leu	Leu	Lys	Met	Leu	Ala	Thr	Leu	Ala	Ser	His	Arg	Val
1415						1420					1425			
Pro	Gly	Leu	Asp	Leu	Ser	Ser	Met	Tyr	His	Pro	Ala	Val	Val	Glu
1430						1435					1440			
Arg	Cys	Arg	Leu	Ala	Leu	Ala	Ala	Gln	Arg	Ser	Gly	Gln	Pro	Glu
1445						1450					1455			
Gln	Arg	Asn	Lys	Phe	Leu	Arg	Thr	Ile	Glu	Val	Asn	Gly	Phe	Tyr
1460						1465					1470			
Asp	Pro	Ala	Asp	Ala	Thr	Ile	Pro	Glu	Ala	Val	Ala	Thr	Ile	Leu
1475						1480					1485			
Pro	Ala	Thr	Ala	Ala	Ile	Ser	Pro	Pro	Lys	Leu	Gly	Ala	Pro	His
1490						1495					1500			
Asp	Ser	Gln	Pro	Glu	Ala	Glu	Ala	Arg	Pro	Val	Gly	Glu	Ala	Ser
1505						1510					1515			
Val	Pro	Arg	Arg	Ala	Thr	Ser	Ser	Ser	Lys	Leu	Ala	Arg	Thr	Leu
1520						1525					1530			
Ala	Ile	Asp	Ala	Cys	Asp	Ser	Asp	Val	Arg	Ala	Ala	Leu	Leu	Asp
1535						1540					1545			

ES 2 623 610 T3

Leu Asp 1550	Ala Pro Ile Ala Val 1555	Gly Gly Ser Ser Arg 1560	Ala Gln Val
Pro Pro 1565	Cys Pro Val Ser Ala 1570	Leu Gly Ser Ala Ala 1575	Phe Arg Ala
Ala His 1580	Gly Val Asp Tyr Ala 1585	Leu Tyr Met Gly Ala 1590	Met Ala Lys
Gly Val 1595	Ala Ser Ala Glu Met 1600	Val Ile Ala Ala Gly 1605	Lys Ala Arg
Met Leu 1610	Ala Ser Phe Gly Ala 1615	Gly Gly Leu Pro Leu 1620	Gly Glu Val
Glu Glu 1625	Ala Leu Asp Lys Ile 1630	Gln Ala Ala Leu Pro 1635	Glu Gly Pro
Phe Ala 1640	Val Asn Leu Ile His 1645	Ser Pro Phe Asp Pro 1650	Asn Leu Glu
Glu Gly 1655	Asn Val Glu Leu Phe 1660	Leu Arg Arg Gly Ile 1665	Arg Leu Val
Glu Ala 1670	Ser Ala Phe Met Ser 1675	Val Thr Pro Ser Leu 1680	Val Arg Tyr
Arg Val 1685	Ala Gly Leu Glu Arg 1690	Gly Pro Gly Gly Thr 1695	Ala Arg Val
Leu Asn 1700	Arg Val Ile Gly Lys 1705	Val Ser Arg Ala Glu 1710	Leu Ala Glu
Met Phe 1715	Met Arg Pro Pro Pro 1720	Ala Ala Ile Val Ser 1725	Lys Leu Leu
Ala Gln 1730	Gly Leu Val Thr Glu 1735	Glu Gln Ala Ser Leu 1740	Ala Glu Ile
Val Pro 1745	Leu Val Asp Asp Val 1750	Ala Ile Glu Ala Asp 1755	Ser Gly Gly
His Thr 1760	Asp Asn Arg Pro Ile 1765	His Val Val Leu Pro 1770	Val Val Leu
Ala Leu 1775	Arg Asp Arg Val Met 1780	Arg Glu Cys Lys Tyr 1785	Pro Ala Ala
Asn Arg	Val Arg Val Gly Ala	Gly Gly Gly Ile Gly	Cys Pro Ala

ES 2 623 610 T3

1790		1795		1800
Ala Ala Arg Ala Ala Phe Asp Met Gly Ala Ala Phe Val Leu Thr	1805	1810		1815
Gly Ser Ile Asn Gln Leu Thr Arg Gln Ala Gly Thr Ser Asp Ser	1820	1825		1830
Val Arg Ala Ala Leu Ala Arg Ala Thr Tyr Ser Asp Val Thr Met	1835	1840		1845
Ala Pro Ala Ala Asp Met Phe Asp Gln Gly Val Lys Leu Gln Val	1850	1855		1860
Leu Lys Arg Gly Thr Met Phe Pro Ala Arg Ala Asn Lys Leu Tyr	1865	1870		1875
Glu Leu Phe Thr Thr Tyr Gln Ser Leu Asp Ala Ile Pro Arg Ala	1880	1885		1890
Glu Leu Ala Arg Leu Glu Lys Arg Val Phe Arg Met Ser Ile Asp	1895	1900		1905
Glu Val Trp Asn Glu Thr Lys Gln Phe Tyr Glu Thr Arg Leu Asn	1910	1915		1920
Asn Pro Ala Lys Val Ala Arg Ala Glu Arg Asp Pro Lys Leu Lys	1925	1930		1935
Met Ser Leu Cys Phe Arg Trp Tyr Leu Ser Lys Ser Ser Lys Trp	1940	1945		1950
Ala Ser Thr Gly Gln Val Gly Arg Glu Leu Asp Tyr Gln Val Trp	1955	1960		1965
Cys Gly Pro Thr Ile Gly Ala Phe Asn Glu Phe Val Lys Gly Ser	1970	1975		1980
Ser Leu Asp Ala Glu Ala Cys Gly Gly Arg Phe Pro Cys Val Val	1985	1990		1995
Arg Val Asn Gln Glu Ile Leu Cys Gly Ala Ala Tyr Glu Gln Arg	2000	2005		2010
Leu Ala Arg Phe Met Leu Leu Ala Gly Arg Glu Ser Ala Asp Ala	2015	2020		2025
Leu Ala Tyr Thr Val Ala Glu Ala Arg	2030	2035		

<210> 60
 <211> 1470
 <212> PRT
 <213> *Thraustochytrium* sp.

5

<400> 60

```

Met Gly Pro Arg Val Ala Ser Gly Lys Val Pro Ala Trp Glu Met Ser
 1           5           10           15

Lys Ser Glu Leu Cys Asp Asp Arg Thr Val Val Phe Asp Tyr Glu Glu
          20           25           30

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

Phe Lys Val Val Asp Gly Phe Arg Arg Arg Val Arg Leu Pro Ala Arg
 50           55           60

Glu Tyr Leu Leu Val Thr Arg Val Thr Leu Met Asp Ala Glu Val Gly
65           70           75           80

Asn Phe Arg Val Gly Ala Arg Met Val Thr Glu Tyr Asp Val Pro Val
          85           90           95

Asn Gly Glu Leu Ser Glu Gly Gly Asp Val Pro Trp Ala Val Leu Val
      100           105           110

Glu Ala Gly Gln Cys Asp Leu Leu Leu Ile Ser Tyr Met Gly Ile Asp
      115           120           125

Phe Gln Cys Lys Gly Glu Arg Val Tyr Arg Leu Leu Asn Thr Thr Leu
      130           135           140

Thr Phe Phe Gly Val Ala Lys Glu Gly Glu Thr Leu Val Tyr Asp Ile
145           150           155           160

Arg Val Thr Gly Phe Ala Lys Arg Pro Asp Gly Asp Ile Ser Met Phe
          165           170           175

Phe Phe Glu Tyr Asp Cys Tyr Cys Asn Gly Lys Leu Leu Ile Glu Met
          180           185           190

Arg Asp Gly Ser Ala Gly Phe Phe Thr Asp Glu Glu Leu Ala Ala Gly
      195           200           205

Lys Gly Val Val Val Thr Arg Ala Gln Gln Asn Met Arg Asp Lys Ile
      210           215           220

```

ES 2 623 610 T3

Val	Arg	Gln	Ser	Ile	Glu	Pro	Phe	Ala	Leu	Ala	Ala	Cys	Thr	His	Lys
225					230					235					240
Thr	Thr	Leu	Asn	Glu	Ser	Asp	Met	Gln	Ser	Leu	Val	Glu	Arg	Asn	Trp
			245					250						255	
Ala	Asn	Val	Phe	Gly	Thr	Ser	Asn	Lys	Met	Ala	Glu	Leu	Asn	Tyr	Lys
		260						265					270		
Ile	Cys	Ala	Arg	Lys	Met	Leu	Met	Ile	Asp	Arg	Val	Thr	His	Ile	Asp
	275						280					285			
His	His	Gly	Gly	Ala	Tyr	Gly	Leu	Gly	Leu	Leu	Val	Gly	Glu	Lys	Ile
	290					295					300				
Leu	Asp	Arg	Asn	His	Trp	Tyr	Phe	Pro	Cys	His	Phe	Val	Asn	Asp	Gln
305					310					315					320
Val	Met	Ala	Gly	Ser	Leu	Val	Ser	Asp	Gly	Cys	Ser	Gln	Leu	Leu	Lys
				325					330					335	
Leu	Tyr	Met	Ile	Trp	Leu	Gly	Leu	His	Leu	Lys	Met	Glu	Glu	Phe	Asp
			340					345					350		
Phe	Leu	Pro	Val	Ser	Gly	His	Lys	Asn	Lys	Val	Arg	Cys	Arg	Gly	Gln
		355					360					365			
Ile	Ser	Pro	His	Lys	Gly	Lys	Leu	Val	Tyr	Val	Met	Glu	Ile	Lys	Lys
	370					375					380				
Met	Gly	Tyr	Asp	Gln	Ala	Ser	Gly	Ser	Pro	Tyr	Ala	Ile	Ala	Asp	Val
385					390					395					400
Asp	Ile	Ile	Asp	Val	Asn	Glu	Glu	Leu	Gly	Gln	Ser	Phe	Asp	Ile	Asn
				405					410					415	
Asp	Leu	Ala	Ser	Tyr	Gly	Lys	Gly	Asp	Leu	Ser	Lys	Lys	Ile	Val	Val
			420					425					430		
Asp	Phe	Lys	Gly	Ile	Ala	Leu	Gln	Leu	Lys	Gly	Arg	Ala	Phe	Ser	Arg
		435					440					445			
Met	Ser	Ser	Ser	Ser	Ser	Leu	Asn	Glu	Gly	Trp	Gln	Cys	Val	Pro	Lys
	450					455					460				
Pro	Ser	Gln	Arg	Met	Glu	His	Glu	Gln	Pro	Pro	Ala	His	Cys	Leu	Ala
465					470					475					480
Ser	Asp	Pro	Glu	Ala	Pro	Ser	Thr	Val	Thr	Trp	His	Pro	Met	Ser	Lys

ES 2 623 610 T3

				485				490				495			
Leu	Pro	Gly	Asn	Pro	Thr	Pro	Phe	Phe	Ser	Pro	Ser	Ser	Tyr	Pro	Pro
			500					505					510		
Arg	Ala	Ile	Cys	Phe	Ile	Pro	Phe	Pro	Gly	Asn	Pro	Leu	Asp	Asn	Asn
		515					520					525			
Cys	Lys	Ala	Gly	Glu	Met	Pro	Leu	Asn	Trp	Tyr	Asn	Met	Ser	Glu	Phe
	530					535					540				
Met	Cys	Gly	Lys	Val	Ser	Asn	Cys	Leu	Gly	Pro	Glu	Phe	Ala	Arg	Phe
545					550					555					560
Asp	Lys	Ser	Asn	Thr	Ser	Arg	Ser	Pro	Ala	Phe	Asp	Leu	Ala	Leu	Val
				565					570					575	
Thr	Arg	Val	Val	Glu	Val	Thr	Asn	Met	Glu	His	Gly	Lys	Phe	Leu	Asn
			580					585					590		
Val	Asp	Cys	Asn	Pro	Ser	Lys	Gly	Thr	Met	Val	Gly	Glu	Phe	Asp	Cys
		595					600					605			
Pro	Gln	Asp	Ala	Trp	Phe	Phe	Asp	Gly	Ser	Cys	Asn	Asp	Gly	His	Met
	610					615					620				
Pro	Tyr	Ser	Ile	Ile	Met	Glu	Ile	Gly	Leu	Gln	Thr	Ser	Gly	Val	Leu
625					630					635					640
Thr	Ser	Val	Leu	Lys	Ala	Pro	Leu	Thr	Met	Asp	Lys	Asp	Asp	Ile	Leu
			645						650					655	
Phe	Arg	Asn	Leu	Asp	Ala	Ser	Ala	Glu	Met	Val	Arg	Pro	Asp	Val	Asp
			660					665					670		
Val	Arg	Gly	Lys	Thr	Ile	Arg	Asn	Val	Thr	Lys	Cys	Thr	Gly	Tyr	Ala
		675					680					685			
Met	Leu	Gly	Lys	Met	Gly	Ile	His	Arg	Phe	Thr	Phe	Glu	Leu	Ser	Val
	690					695					700				
Asp	Gly	Val	Val	Phe	Tyr	Lys	Gly	Ser	Thr	Ser	Phe	Gly	Trp	Phe	Thr
705					710					715					720
Pro	Glu	Val	Phe	Ala	Gln	Gln	Ala	Gly	Leu	Asp	Asn	Gly	Lys	Lys	Thr
			725						730					735	
Glu	Pro	Trp	Cys	Lys	Thr	Asn	Asn	Thr	Ser	Val	Arg	Arg	Val	Glu	Ile
			740					745					750		

ES 2 623 610 T3

Ala	Ser	Ala	Lys	Gly	Lys	Glu	Gln	Leu	Thr	Glu	Lys	Leu	Pro	Asp	Ala
		755					760					765			
Thr	Asn	Ala	Gln	Val	Leu	Arg	Arg	Ser	Glu	Gln	Cys	Glu	Tyr	Leu	Asp
	770					775					780				
Tyr	Leu	Asn	Ile	Ala	Pro	Asp	Ser	Gly	Leu	His	Gly	Lys	Gly	Tyr	Ala
785					790					795					800
His	Gly	His	Lys	Asp	Val	Asn	Pro	Gln	Asp	Trp	Phe	Phe	Ser	Cys	His
				805					810					815	
Phe	Trp	Phe	Asp	Pro	Val	Met	Pro	Gly	Ser	Leu	Gly	Ile	Glu	Ser	Met
			820					825					830		
Phe	Gln	Leu	Ile	Glu	Ala	Phe	Ala	Val	Asp	Gln	Asn	Ile	Pro	Gly	Glu
		835					840					845			
Tyr	Asn	Val	Ser	Asn	Pro	Thr	Phe	Ala	His	Ala	Pro	Gly	Lys	Thr	Ala
	850					855					860				
Trp	Lys	Tyr	Arg	Gly	Gln	Leu	Thr	Pro	Lys	Asn	Arg	Ala	Met	Asp	Cys
865					870					875					880
Glu	Val	His	Ile	Val	Ser	Ile	Thr	Ala	Ser	Pro	Glu	Asn	Gly	Gly	Tyr
				885					890					895	
Val	Asp	Ile	Val	Ala	Asp	Gly	Ala	Leu	Trp	Val	Asp	Gly	Leu	Arg	Val
			900					905					910		
Tyr	Glu	Ala	Lys	Glu	Leu	Arg	Val	Arg	Val	Val	Ser	Ala	Lys	Pro	Gln
		915					920					925			
Ala	Ile	Pro	Asp	Val	Gln	Gln	Gln	Pro	Pro	Ser	Ala	Lys	Ala	Asp	Pro
	930					935					940				
Gly	Lys	Thr	Gly	Val	Ala	Leu	Ser	Pro	Thr	Gln	Leu	Arg	Asp	Val	Leu
945					950					955					960
Leu	Glu	Val	Asp	Asn	Pro	Leu	Tyr	Leu	Gly	Val	Glu	Asn	Ser	Asn	Leu
				965					970					975	
Val	Gln	Phe	Glu	Ser	Lys	Pro	Ala	Thr	Ser	Ser	Arg	Ile	Val	Ser	Ile
			980					985					990		
Lys	Pro	Cys	Ser	Ile	Ser	Asp	Leu	Gly	Asp	Lys	Ser	Phe	Met	Glu	Thr
		995					1000					1005			

ES 2 623 610 T3

Tyr Asn Val Ser Ala Pro Leu	Tyr Thr Gly Ala Met	Ala Lys Gly
1010	1015	1020
Ile Ala Ser Ala Asp Leu Val	Ile Ala Ala Gly Lys	Arg Lys Ile
1025	1030	1035
Leu Gly Ser Phe Gly Ala Gly	Gly Leu Pro Ile Ser	Ile Val Arg
1040	1045	1050
Glu Ala Leu Glu Lys Ile Gln	Gln His Leu Pro His	Gly Pro Tyr
1055	1060	1065
Ala Val Asn Leu Ile His Ser	Pro Phe Asp Ser Asn	Leu Glu Lys
1070	1075	1080
Gly Asn Val Asp Leu Phe Leu	Glu Met Gly Val Thr	Val Val Glu
1085	1090	1095
Cys Ser Ala Phe Met Glu Leu	Thr Ala Gln Val Val	Arg Tyr Arg
1100	1105	1110
Ala Ser Gly Leu Ser Lys Ser	Ala Asp Gly Ser Ile	Arg Ile Ala
1115	1120	1125
His Arg Ile Ile Gly Lys Val	Ser Arg Thr Glu Leu	Ala Glu Met
1130	1135	1140
Phe Ile Arg Pro Ala Pro Gln	His Leu Leu Gln Lys	Leu Val Ala
1145	1150	1155
Ser Gly Glu Leu Thr Ala Glu	Gln Ala Glu Leu Ala	Thr Gln Val
1160	1165	1170
Pro Val Ala Asp Asp Ile Ala	Val Glu Ala Asp Ser	Gly Gly His
1175	1180	1185
Thr Asp Asn Arg Pro Ile His	Val Ile Leu Pro Leu	Ile Ile Asn
1190	1195	1200
Leu Arg Asn Arg Leu His Lys	Glu Leu Asp Tyr Pro	Ser His Leu
1205	1210	1215
Arg Val Arg Val Gly Ala Gly	Gly Gly Ile Gly Cys	Pro Gln Ala
1220	1225	1230
Ala Leu Ala Ala Phe Gln Met	Gly Ala Ala Phe Leu	Ile Thr Gly
1235	1240	1245

Thr Val	Asn Gln Leu	Ala Arg	Glu Ser Gly Thr	Cys	Asp Asn Val
1250		1255		1260	
Arg Leu	Gln Leu Ser Lys	Ala	Thr Tyr Ser Asp	Val	Cys Met Ala
1265		1270		1275	
Pro Ala	Ala Asp Met Phe	Asp	Gln Gly Val Glu	Leu	Gln Val Leu
1280		1285		1290	
Lys Lys	Gly Thr Leu Phe	Pro	Ser Arg Ala Lys	Lys	Leu Tyr Glu
1295		1300		1305	
Leu Phe	Cys Lys Tyr Asp	Ser	Phe Glu Ala Met	Pro	Ala Glu Glu
1310		1315		1320	
Leu Gln	Arg Val Glu Lys	Arg	Ile Phe Gln Lys	Ser	Leu Ala Glu
1325		1330		1335	
Val Trp	Gln Glu Thr Ser	Asp	Phe Tyr Ile His	Arg	Ile Lys Asn
1340		1345		1350	
Pro Glu	Lys Ile Asn Arg	Ala	Ala Ser Asp Gly	Lys	Leu Lys Met
1355		1360		1365	
Ser Leu	Cys Phe Arg Trp	Tyr	Leu Gly Leu Ser	Ser	Phe Trp Ala
1370		1375		1380	
Asn Ser	Gly Ala Gln Asp	Arg	Val Met Asp Tyr	Gln	Ile Trp Cys
1385		1390		1395	
Gly Pro	Ala Ile Gly Ala	Phe	Asn Asp Phe Thr	Lys	Gly Thr Tyr
1400		1405		1410	
Leu Asp	Val Thr Val Ala	Lys	Ser Tyr Pro Cys	Val	Ala Gln Ile
1415		1420		1425	
Asn Leu	Gln Ile Leu Gln	Gly	Ala Ala Tyr Leu	Lys	Arg Leu Gly
1430		1435		1440	
Val Ile	Arg Phe Asp Arg	Met	Leu Leu Gln Ala	Val	Asp Ile Asp
1445		1450		1455	
Asp Pro	Val Phe Thr Tyr	Val	Pro Thr Gln Pro	Leu	
1460		1465		1470	

<210> 61
<211> 1503
<212> PRT
<213> *Schizochytrium* sp.

5

<400> 61

ES 2 623 610 T3

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

Ala	Ser	Val	Phe	Gly	Ser	Lys	Asn	Gly	Met	Pro	Glu	Ile	Asn	Tyr	Lys		
			260					265					270				
Leu	Cys	Ala	Arg	Lys	Met	Leu	Met	Ile	Asp	Arg	Val	Thr	Ser	Ile	Asp		
		275					280					285					
His	Lys	Gly	Gly	Val	Tyr	Gly	Leu	Gly	Gln	Leu	Val	Gly	Glu	Lys	Ile		
	290					295					300						
Leu	Glu	Arg	Asp	His	Trp	Tyr	Phe	Pro	Cys	His	Phe	Val	Lys	Asp	Gln		
305					310					315					320		
Val	Met	Ala	Gly	Ser	Leu	Val	Ser	Asp	Gly	Cys	Ser	Gln	Met	Leu	Lys		
				325					330					335			
Met	Tyr	Met	Ile	Trp	Leu	Gly	Leu	His	Leu	Thr	Thr	Gly	Pro	Phe	Asp		
			340					345					350				
Phe	Arg	Pro	Val	Asn	Gly	His	Pro	Asn	Lys	Val	Arg	Cys	Arg	Gly	Gln		
		355					360					365					
Ile	Ser	Pro	His	Lys	Gly	Lys	Leu	Val	Tyr	Val	Met	Glu	Ile	Lys	Glu		
	370					375					380						
Met	Gly	Phe	Asp	Glu	Asp	Asn	Asp	Pro	Tyr	Ala	Ile	Ala	Asp	Val	Asn		
385					390					395					400		
Ile	Ile	Asp	Val	Asp	Phe	Glu	Lys	Gly	Gln	Asp	Phe	Ser	Leu	Asp	Arg		
				405					410					415			
Ile	Ser	Asp	Tyr	Gly	Lys	Gly	Asp	Leu	Asn	Lys	Lys	Ile	Val	Val	Asp		
		420						425					430				
Phe	Lys	Gly	Ile	Ala	Leu	Lys	Met	Gln	Lys	Arg	Ser	Thr	Asn	Lys	Asn		
		435					440					445					
Pro	Ser	Lys	Val	Gln	Pro	Val	Phe	Ala	Asn	Gly	Ala	Ala	Thr	Val	Gly		
	450					455					460						
Pro	Glu	Ala	Ser	Lys	Ala	Ser	Ser	Gly	Ala	Ser	Ala	Ser	Ala	Ser	Ala		
465					470					475					480		
Ala	Pro	Ala	Lys	Pro	Ala	Phe	Ser	Ala	Asp	Val	Leu	Ala	Pro	Lys	Pro		
				485					490					495			
Val	Ala	Leu	Pro	Glu	His	Ile	Leu	Lys	Gly	Asp	Ala	Leu	Ala	Pro	Lys		
			500					505					510				

ES 2 623 610 T3

Glu Met Ser Trp His Pro Met Ala Arg Ile Pro Gly Asn Pro Thr Pro
 515 520 525
 Ser Phe Ala Pro Ser Ala Tyr Lys Pro Arg Asn Ile Ala Phe Thr Pro
 530 535 540
 Phe Pro Gly Asn Pro Asn Asp Asn Asp His Thr Pro Gly Lys Met Pro
 545 550 555 560
 Leu Thr Trp Phe Asn Met Ala Glu Phe Met Ala Gly Lys Val Ser Met
 565 570 575
 Cys Leu Gly Pro Glu Phe Ala Lys Phe Asp Asp Ser Asn Thr Ser Arg
 580 585 590
 Ser Pro Ala Trp Asp Leu Ala Leu Val Thr Arg Ala Val Ser Val Ser
 595 600 605
 Asp Leu Lys His Val Asn Tyr Arg Asn Ile Asp Leu Asp Pro Ser Lys
 610 615 620
 Gly Thr Met Val Gly Glu Phe Asp Cys Pro Ala Asp Ala Trp Phe Tyr
 625 630 635 640
 Lys Gly Ala Cys Asn Asp Ala His Met Pro Tyr Ser Ile Leu Met Glu
 645 650 655
 Ile Ala Leu Gln Thr Ser Gly Val Leu Thr Ser Val Leu Lys Ala Pro
 660 665 670
 Leu Thr Met Glu Lys Asp Asp Ile Leu Phe Arg Asn Leu Asp Ala Asn
 675 680 685
 Ala Glu Phe Val Arg Ala Asp Leu Asp Tyr Arg Gly Lys Thr Ile Arg
 690 695 700
 Asn Val Thr Lys Cys Thr Gly Tyr Ser Met Leu Gly Glu Met Gly Val
 705 710 715 720
 His Arg Phe Thr Phe Glu Leu Tyr Val Asp Asp Val Leu Phe Tyr Lys
 725 730 735
 Gly Ser Thr Ser Phe Gly Trp Phe Val Pro Glu Val Phe Ala Ala Gln
 740 745 750
 Ala Gly Leu Asp Asn Gly Arg Lys Ser Glu Pro Trp Phe Ile Glu Asn
 755 760 765
 Lys Val Pro Ala Ser Gln Val Ser Ser Phe Asp Val Arg Pro Asn Gly

ES 2 623 610 T3

770	775	780
Ser Gly Arg Thr Ala Ile Phe Ala Asn Ala Pro Ser Gly Ala Gln Leu 785 790 795 800		
Asn Arg Arg Thr Asp Gln Gly Gln Tyr Leu Asp Ala Val Asp Ile Val 805 810 815		
Ser Gly Ser Gly Lys Lys Ser Leu Gly Tyr Ala His Gly Ser Lys Thr 820 825 830		
Val Asn Pro Asn Asp Trp Phe Phe Ser Cys His Phe Trp Phe Asp Ser 835 840 845		
Val Met Pro Gly Ser Leu Gly Val Glu Ser Met Phe Gln Leu Val Glu 850 855 860		
Ala Ile Ala Ala His Glu Asp Leu Ala Gly Lys Ala Arg His Cys Gln 865 870 875 880		
Pro His Leu Cys Ala Arg Pro Arg Ala Arg Ser Ser Trp Lys Tyr Arg 885 890 895		
Gly Gln Leu Thr Pro Lys Ser Lys Lys Met Asp Ser Glu Val His Ile 900 905 910		
Val Ser Val Asp Ala His Asp Gly Val Val Asp Leu Val Ala Asp Gly 915 920 925		
Phe Leu Trp Ala Asp Ser Leu Arg Val Tyr Ser Val Ser Asn Ile Arg 930 935 940		
Val Arg Ile Ala Ser Gly Glu Ala Pro Ala Ala Ala Ser Ser Ala Ala 945 950 955 960		
Ser Val Gly Ser Ser Ala Ser Ser Val Glu Arg Thr Arg Ser Ser Pro 965 970 975		
Ala Val Ala Ser Gly Pro Ala Gln Thr Ile Asp Leu Lys Gln Leu Lys 980 985 990		
Thr Glu Leu Leu Glu Leu Asp Ala Pro Leu Tyr Leu Ser Gln Asp Pro 995 1000 1005		
Thr Ser Gly Gln Leu Lys Lys His Thr Asp Val Ala Ser Gly Gln 1010 1015 1020		
Ala Thr Ile Val Gln Pro Cys Thr Leu Gly Asp Leu Gly Asp Arg 1025 1030 1035		

Ser Phe 1040	Met Glu Thr Tyr Gly 1045	Val Val Ala Pro Leu 1050	Tyr Thr Gly
Ala Met 1055	Ala Lys Gly Ile Ala 1060	Ser Ala Asp Leu Val 1065	Ile Ala Ala
Gly Lys 1070	Arg Lys Ile Leu Gly 1075	Ser Phe Gly Ala Gly 1080	Gly Leu Pro
Met His 1085	His Val Arg Ala Ala 1090	Leu Glu Lys Ile Gln 1095	Ala Ala Leu
Pro Gln 1100	Gly Pro Tyr Ala Val 1105	Asn Leu Ile His Ser 1110	Pro Phe Asp
Ser Asn 1115	Leu Glu Lys Gly Asn 1120	Val Asp Leu Phe Leu 1125	Glu Lys Gly
Val Thr 1130	Val Val Glu Ala Ser 1135	Ala Phe Met Thr Leu 1140	Thr Pro Gln
Val Val 1145	Arg Tyr Arg Ala Ala 1150	Gly Leu Ser Arg Asn 1155	Ala Asp Gly
Ser Val 1160	Asn Ile Arg Asn Arg 1165	Ile Ile Gly Lys Val 1170	Ser Arg Thr
Glu Leu 1175	Ala Glu Met Phe Ile 1180	Arg Pro Ala Pro Glu 1185	His Leu Leu
Glu Lys 1190	Leu Ile Ala Ser Gly 1195	Glu Ile Thr Gln Glu 1200	Gln Ala Glu
Leu Ala 1205	Arg Arg Val Pro Val 1210	Ala Asp Asp Ile Ala 1215	Val Glu Ala
Asp Ser 1220	Gly Gly His Thr Asp 1225	Asn Arg Pro Ile His 1230	Val Ile Leu
Pro Leu 1235	Ile Ile Asn Leu Arg 1240	Asn Arg Leu His Arg 1245	Glu Cys Gly
Tyr Pro 1250	Ala His Leu Arg Val 1255	Arg Val Gly Ala Gly 1260	Gly Gly Val
Gly Cys 1265	Pro Gln Ala Ala Ala 1270	Ala Ala Leu Thr Met 1275	Gly Ala Ala

Phe	Ile	Val	Thr	Gly	Thr	Val	Asn	Gln	Val	Ala	Lys	Gln	Ser	Gly
1280						1285					1290			
Thr	Cys	Asp	Asn	Val	Arg	Lys	Gln	Leu	Ser	Gln	Ala	Thr	Tyr	Ser
1295						1300					1305			
Asp	Ile	Cys	Met	Ala	Pro	Ala	Ala	Asp	Met	Phe	Glu	Glu	Gly	Val
1310						1315					1320			
Lys	Leu	Gln	Val	Leu	Lys	Lys	Gly	Thr	Met	Phe	Pro	Ser	Arg	Ala
1325						1330					1335			
Asn	Lys	Leu	Tyr	Glu	Leu	Phe	Cys	Lys	Tyr	Asp	Ser	Phe	Asp	Ser
1340						1345					1350			
Met	Pro	Pro	Ala	Glu	Leu	Glu	Arg	Ile	Glu	Lys	Arg	Ile	Phe	Lys
1355						1360					1365			
Arg	Ala	Leu	Gln	Glu	Val	Trp	Glu	Glu	Thr	Lys	Asp	Phe	Tyr	Ile
1370						1375					1380			
Asn	Gly	Leu	Lys	Asn	Pro	Glu	Lys	Ile	Gln	Arg	Ala	Glu	His	Asp
1385						1390					1395			
Pro	Lys	Leu	Lys	Met	Ser	Leu	Cys	Phe	Arg	Trp	Tyr	Leu	Gly	Leu
1400						1405					1410			
Ala	Ser	Arg	Trp	Ala	Asn	Met	Gly	Ala	Pro	Asp	Arg	Val	Met	Asp
1415						1420					1425			
Tyr	Gln	Val	Trp	Cys	Gly	Pro	Ala	Ile	Gly	Ala	Phe	Asn	Asp	Phe
1430						1435					1440			
Ile	Lys	Gly	Thr	Tyr	Leu	Asp	Pro	Ala	Val	Ser	Asn	Glu	Tyr	Pro
1445						1450					1455			
Cys	Val	Val	Gln	Ile	Asn	Leu	Gln	Ile	Leu	Arg	Gly	Ala	Cys	Tyr
1460						1465					1470			
Leu	Arg	Arg	Leu	Asn	Ala	Leu	Arg	Asn	Asp	Pro	Arg	Ile	Asp	Leu
1475						1480					1485			
Glu	Thr	Glu	Asp	Ala	Ala	Phe	Val	Tyr	Glu	Pro	Thr	Asn	Ala	Leu
1490						1495					1500			

<210> 62
 <211> 27

 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> Cebador sintético prDS173
 <220>
 10 <221> misc_feature
 <222> (19)..(19)
 <223> n es a, c, g o t

 <220>
 15 <221> misc_feature
 <222> (22)..(22)
 <223> n es a, c, g o t

 <400> 62
 20 gatctactgc aagcgcgng gnttyat 27

 <210> 63
 <211> 21
 <212> ADN
 25 <213> Secuencia artificial

 <220>
 <223> Cebador sintético prDS174

 <220>
 30 <221> misc_feature
 <222> (16)..(16)
 <223> n es a, c, g o t

 <220>
 35 <221> misc_feature
 <222> (19)..(19)
 <223> n es a, c, g o t

 <400> 63
 40 ggcgcaggcg gortcnacna c 21

 <210> 64
 <211> 23
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador sintético JGM190
 50
 <220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n es a, c, g o t
 55
 <400> 64
 caytggtayt tyccntgyca ytt 23

 <210> 65
 60 <211> 17
 <212> ADN
 <213> Secuencia artificial

 <220>
 65 <223> Cebador sintético BLR242

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> n es a, c, g o t
 5
 <220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n es a, c, g o t
 10
 <220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n es a, c, g o t
 15
 <400> 65
 ccnggcatna cnggrtc 17
 <210> 66
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 20
 <220>
 <223> Cebador sintético CAX055
 25
 <400> 66
 gtcagtgatg aacaagatgg attgcac 27
 30
 <210> 67
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> Cebador sintético CAX056
 <400> 67
 ccacgtgtca gaagaactcg tcaagaa 27
 40
 <210> 68
 <211> 8976
 <212> ADN
 <213> *Thraustochytrium* sp.
 45
 <400> 68

 atggaagatc aaagaattgc tattgttgga ttatctgcga ttttaccaag tggtgaaaat 60
 gttagagaat cttgggaagc aatacgtgat ggtttgaatt gtttaagtga tttacctgcg 120
 gatcgtgttg atgttactgc gtattataat ccaacaaaag gtgtaaagga taaaatttat 180
 tgtaaacgtg gtgggtttat tcctgaatat gaatttgatt ctagagaatt tggacttaat 240
 atgttacaaa tggaagattc tgatgctaata caaacgttaa ctttattaaa gggttaaagaa 300
 gcattagatg atgctaatat acctgcattt actaatgaga aaaaaaatat tggttgtggtt 360

cttggtattg gtgggtggca aaaagcatct catgaatddd attcaagact taattatggt 420
 gttgtggata aagttttaag aaaaatggga ttacctgatg aggatggtga aactgctggt 480
 gaaaagttda aagctaattt tccatgaatg agattagatt cctttccctg ttttcttggt 540
 aatgttactg ctggccgttg tactaataca ttcaatatgg aaggatgaa ttgtgttgta 600
 gatgctgctt gtgctagttc ttttaattgct attaaagttg ctattgatga attattacat 660
 ggtgattgtg atgcaatgat tgctgggtgca acttggtactg ataacgctct tggatgtgat 720
 atggcatddd caaaaacacc tgttttttca actgatcaaa gttgtcttgc atatgatgaa 780
 aaaacaaaag gtatgcttat tgggtgaaggc tcagctatgt ttgtttttaa acgttatgct 840
 gacgcagtga gagatggtga tactgtacat gctgttatac gttcatgttc atcatcatct 900
 gacggtaaag catctggtat ttataacca actattdctg gtcaagaaga agctattctt 960
 agagcatatc gtagagctgg tgtatcacca aatactatta ctttagttga aggacatggt 1020
 actggtacac cagtgggtga taaaattgaa ttaacagctt tacgcaatgt atttgataaa 1080
 gcatatggtc ctgggtcataa ggaagaagtt gctgttgga gtattaaaag tcaaattggt 1140
 catttgaaag ctgttgctgg ttgtgctggt cttgtgaaat tggttatggc attgaaacat 1200
 aaaacactac ctcaaagtat taatgttgaa aatccacctt atttagtgga tgggtactgc 1260
 attagtata ctactttata tattaataca atgaatcgtc catggattac taagcctggt 1320
 gttccaagaa gagctggtat atctagtttc ggatttggtg gtgcaaatta tcatgctggt 1380
 ttagaagaat ttgagccgga acaaactaaa ccatatagat tgaatgtatc tgcacaacca 1440
 atgcttcttc atgcggtaaa tgcaaatca ttacaaaagc tatgtgaaga tcaattaaaa 1500
 cttttgaaag aatcaagaga aaaatgtgtc aacacaaaaa acactgatta tgttgctgtt 1560
 tcaaaatttc aagattcttt taaattgaaa ggttctgttc catcacaaca tgctagagtt 1620
 ggttttgcat caaaatctat tgaagatact atttctattd tatctgctat cgttaataga 1680
 tttcaaaaag atattacaac aactagttgg gctttacca aagaagggtc tattttttaga 1740
 tctactgcat tgattaatga caataaaagt gtagctgctt tattttctgg acaaggcgca 1800
 caatataccc atatgtttta tgatgttgca atgcaatggc cacaatttcg tttatgtgta 1860
 aatgatattg agaaagcaca ggaagaagtt atcaatgata aaagtgtgaa acgtatcagt 1920
 caagttatgt ttccctcgta accatattga agagaatcac ctttagacaa taaagaaatc 1980
 tctaagactg aatattctca aacaacaact gtcgctagtt cagtaggttt atttgaaatt 2040
 ttccgtgatg ctgggttcgc tctgctttt gttgctgggc attcttttag tgaatttagt 2100
 gcattgtatg cagctggatt gattgatcgc gaagatttat tcaagttggt atgtaatcgt 2160
 gcaatggcta tgagagatgc accaaaaaaa tctgctgatg gagcaatggc tgctgttatt 2220
 ggtccaaatg cttcttcaat taagctttca gctcctgaag tatgggttgc taacaataac 2280

tctccatctc	aaactgttat	taccggtgca	aattctggtg	tacaagctga	aacaagtaaa	2340
ttgaaaactc	aaggtttccg	tgtggttcat	ttggcatgtg	atggggcatt	tcattcgcct	2400
catatggaaa	atgctgaaaa	gcaatttcaa	aaagctcttt	cagcagttaa	gtttaataaaa	2460
ccaactgggt	cttctccaaa	aattttcagc	aatgtaactg	gtggtgtatt	tacggatcca	2520
aaaactgctt	tgtcaagaca	tatgactagt	tctgtacaat	ttcttactca	aattaagaat	2580
atgtacgcgg	ctggagctcg	tgtctttatt	gaatttggac	caaaacaagt	actttccaaa	2640
ttggtcaatg	aaatttttcc	tggtgataca	agcgttttaa	ctgtttcggg	gaatccagct	2700
agtgtctaaag	atagtgcacat	tcaattgcgt	caagctgcag	ttcaaatggc	cgttgcctgg	2760
gtagctctta	ccgattttga	taaatgggaa	ctcaaagatc	ctacccgtat	gaaggaattc	2820
ccacgtaaga	agactacttt	gactttgtct	gcagcaactt	atgtctccaa	gaaaactcta	2880
caggagcgtg	aacgaatcat	gaatgatggg	cgaactgttt	catgtgttca	acgtattgaa	2940
aacactaata	ctggtgagtt	ggagaaattg	aagaagcaat	tgcaagataa	agaaaatgag	3000
gttgtaagag	ttcaagctct	tgcaactcaa	gcttcagctg	atttgcaaaa	taccaaagca	3060
gaattacaaa	aagctcaagc	aacaaaatct	agtaatgcag	catctgatgc	ggtggtggca	3120
aaacataagg	caattttatt	ggcaatgtta	gaagaacttg	aaaccggcaa	ggctgtagat	3180
tattcttcat	tttcgaaagg	tcaagttgca	agtcacagct	cgttctgtgt	cgtttcagct	3240
cctgttcaag	cggctgctcc	tgtgcaggta	tctgcttctg	ttgattctgg	tttgttggca	3300
aaagcggaac	aagttgtatt	ggaagtattg	gcatcgaaga	ctggttatga	gactgagttg	3360
attgaattgg	atatggaatt	ggaaactgaa	cttggtattg	attctatcaa	gagagtagaa	3420
attctttctg	aagttcaagc	tcaattgaat	gttgaagcta	aagatgtaga	tgctcttagt	3480
agaactcgta	ctgttggtga	agtgattgat	gcaatgaaag	ccgaaattgc	tggtggtcaa	3540
ccagctgctc	ctgttcaagt	tgagctcct	actcaagtag	ttgctcctgt	tcaagcatct	3600
gctcctgttg	attctggttt	gttagcaaaa	gcggaacaag	ttgtattgga	agtattggca	3660
togaagactg	gttatgagac	tgagttgatt	gaattggata	tggaattgga	aaccgaactt	3720
ggtattgatt	ctatcaagag	agtagaaatt	ctttctgaag	ttcaagctca	attgagtgtt	3780
gaagctaaag	atgtagatgc	tcttagtaga	actcgtactg	ttggtgaagt	gattgatgca	3840
atgaaagccg	aaattgctgg	tggtcaacca	gctgctcctg	ttcaagttgc	agctcctact	3900
caagtagttg	ctcctgttca	agcatctgct	cctgttgatt	ctggtttgtt	agcaaaagcg	3960
gaacaagtgg	tattggaagt	attggcatcg	aagactgggt	atgagactga	gttgattgaa	4020
ttggatatgg	aattggaaac	cgaacttggt	attgattcta	tcaagagagt	agaaattctt	4080
tctgaagttc	aagctcaatt	gagtgttgaa	gctaaagatg	tagatgctct	tagtagaact	4140
cgtactgttg	gtgaagtgat	tgatgcaatg	aaagctgaaa	tttctggtgg	tcagccagct	4200
gctcctgttc	aagttgcagc	tcctactcaa	atagttgctc	ctgttcaagt	atccgctcct	4260

gttgattctg gtttgtttagc aaaggcggaa caagtagtat tggaaagtatt ggcatccaag 4320
 actggttatg agactgagtt gattgaattg gatatggaat tggaaactga acttggtatt 4380
 gattctatca agagagtaga aattccttct gaagttcaag ctcaattgag tgttgaagct 4440
 aaagatgtag atgctccttag tagaactcgt actgttggtg aagtgattga tgcaatgaaa 4500
 gctgaaattt ctggtggtca accaactgct cctgttcaag ttgcagctcc tactcaaata 4560
 gttgctcctg ttcaagtatc tgctcctggt gattctgggt ttgtagcaaa ggcggaacaa 4620
 gttgtattgg aagtattggc atcgaagact ggttatgaga ctgagttgat tgaattggat 4680
 atggaattgg aaaccgaact tgggtattgat tctatcaaga gagtagaaat tctttctgaa 4740
 gttcaagctc aattgagtg tgaagctaaa gatgtagatg ctcttagtag aactcgtact 4800
 gttggtgaag tgattgatgc aatgaaagcc gaaatttctg gtggtcagcc agctgctcct 4860
 gttcaagttg cagctcctac tcaaatagtt gctcctgttc aagcatctgc tcctgttgat 4920
 tctggtttgt tggcaaaagc ggaacaagtt gtattggaag tggtagcatc caagactggt 4980
 tatgaaactg agttgattga attagatatg gaattggaaa ccgaacttgg tattgattct 5040
 atcaagagag tagaaattct ttctgaagtt caagctcaat tgagtgttga agctaaagat 5100
 gtagatgctc ttagtagaac tcgtactgtt ggtgaagtga ttgatgcaat gaaagctgaa 5160
 atttctggtg gtcaaccagc tgctcctggt caagttgcag ctctactca aatagttgct 5220
 cctgttcaag tatctgctcc tgttgattct ggtttgtag caaaggcggg acaagttgta 5280
 ttggaagtat tggcatctaa gactggttat gagactgagt tgattgaatt ggatatggaa 5340
 ttgaaactg aacttggtat tgattctatc aagagagtag aaattccttc tgaagttcaa 5400
 gctcaattga atgttgaagc taaagatgta gatgctctta gtagaactcg tactgttggt 5460
 gaagtgattg atgcaatgaa agccgaaatt gctggtggtc aaccagctgc tcctgttcaa 5520
 gttgcagctc ctgctccagt agttgctcct gttcaagtat ctactcctgt tgattctggt 5580
 ttgttggcaa aagcgggaaca agttgtattg gaagtgttag catgcaagac tggttatgaa 5640
 actgagttga ttgaattgga tatggaattg gaaactgaac ttggtattga ttctatcaag 5700
 agagtagaaa ttctttctga agttcaagct caattgagtg ttgaagctaa agatgtagat 5760
 gctcttagta gaactcgtac tgttggtgaa gtgattgatg caatgaaagc cgaaatttct 5820
 ggtggtcaac caactgctcc tgttcaagtt gcagctccta ctcaagtagt tgctcctgtt 5880
 aaagtatcta ctctgttgga ttctggtttg ttagcaaagg cggaacaagt agtattggaa 5940
 gtattggcat ctaagactgg ttatgaaact gagttgattg aattagatat ggaattggaa 6000
 actgaacttg gtattgattc tatcaagaga gtagaaattc tttctgaagt tcaagctcaa 6060
 ttgaatgtgg aagctaaaga tgtggatgct cttagtagaa ctcgtactgt tgggtgaagtg 6120
 attgatgcaa tgaaagccga aattgctggt gatcaacctg ctccagctgt agttccagtt 6180

caagctaaga	gtggtgtagc	caaccctgca	cttttgccaa	aggcggaa	agtagtattg	6240
gaagtattgg	catccaagac	cggttatgaa	actgagctga	ttgaattgga	tatggaattg	6300
gaaactgaac	ttggtattga	ttcaatcaag	agagtagaaa	ttctgtccga	agttcaagca	6360
gaattgagtg	ttgaagcaaa	agatgtagac	gctctaagta	gaaccogtac	tgttggggaa	6420
gtgatcgatg	caatgaaagc	tgaaattgct	ggcagtgctg	tcacggttgc	aactttggat	6480
gattcaacaa	ttatggagga	gacagatgat	gaagatgaag	actttatttt	atacgatcat	6540
gtatacggaa	gcgaatgtga	agatcttagt	ctgagttttt	catccgtaaa	gagcatcccg	6600
cgcgctgata	aacttttggt	ggataacatt	gctgaaaggc	caattgttat	tgtggattgt	6660
ggaacaaaagc	ttacaactga	acttgcaaaa	gctattggag	aacgtgccgt	ggttgctaca	6720
ttcagtgcac	agagcttggg	atcccgtaga	ttcgttggtg	aatcattttac	tctaggaaat	6780
acagaagaaa	gtgagatcga	aaagatgggt	tcaagcattg	aatcttcgta	tggaaaaatt	6840
ggtggctttg	tttatcaaca	ttttcatgat	agcgactatg	gtatgcaact	tggatgggag	6900
ttaatggcag	cgaacatttt	gaaagagtcc	ctcaacgacc	cgattaagaa	tggagaagcc	6960
ttcttttttg	ctgttgcgag	tatgaatggg	aaacttggtg	tggacaatgc	ttcagttcat	7020
gatcaaggaa	tagtggaaac	atgcggtatc	gccgaacgtg	gtgctatctt	tggtttgtgc	7080
aaaacttttg	atttggaaatg	gcctaattgt	tttgctcgtg	gtgttgatat	tgtggaaggt	7140
atgagttata	gtttggctgc	ggaattgatt	gttgatgaga	tttcttgtgc	aaatctttcc	7200
attcgggaat	ctggttacac	gattagcggg	gaaagattca	caactgaagc	tcacaaattg	7260
gttactggaa	agcctcatgc	tccgattaag	aagaaggatg	ctttcctagt	atctggtggg	7320
gctcgtggta	ttactccact	ttgtattcgt	gaaattgcta	aagcagtga	aggtggcact	7380
tacattttga	tgggtcgatc	agctttgggt	gatgaaccct	tgtgggctaa	tggtaaatcc	7440
ggaaaagatt	tagataaagc	tggtttggca	tttttgaagg	aagagtttgc	agctgggcgt	7500
ggtagtaaac	caactccaaa	agttcacaaa	tctttgattg	ataaagtgtc	cggatttagg	7560
gaggttagag	catctattgc	aaatatagaa	gcccatggag	caaaagctat	atatttgtct	7620
tgcgatgtat	cttccgctga	gaaagtaaag	gctgcagtgc	aaaaagttga	aaaggagcat	7680
ctagttcgta	ttactgggtat	tgtgcatgca	tcaggcggtt	tgggggataa	attggttgag	7740
aacaaaactt	tggatgattt	caacgcagta	tatggaacca	aagtaactgg	actagtaaac	7800
ttgctgtcag	cagtgaacat	gaattttgtt	cgtcatttgg	ttatgtttag	ttctttggct	7860
ggatatcatg	gaaatgttgg	tcaatctgat	tatgcaatgg	ctaacgaatc	acttaacaag	7920
attggtttta	gattgggtgc	agcttattct	caattgtgtg	ttaaactctat	ttgttttgga	7980
ccttgggatg	gtggaatggg	aactccagct	ttgaaaaaac	aatttcaatc	aatgggtgtc	8040
cagattatcc	ctcgtgaagg	tggcgcgagg	actgttgcaa	gaatagtctt	atcttcaaat	8100
ccttctcaag	ttttagttgg	caactggggg	gttcctccag	tttcaacttt	gtcaaaatcg	8160

```

gcaactattg ttcaaacttt taccctgag ttaaattccat ttctaaagtc tcatcaaatt 8220
catggtaaaa atgttttgcc tatgactgta gcaattggat atcttgctca cttgggtaag 8280
aatttttatg ctggcatca tttgtgggga gttgaagatg ctcaattgtt cagtgggtgtt 8340
gtaattgacc atgcggtgca agctcaagtg aaattaacgg aacagagttt ggatgatgat 8400
ggcaaggtaa aagttcaagc tgttctgact gcttcaaacg ataatggaaa aatgggtacct 8460
gcatacaaag cagtattgt tttgggaaaa acaagtagac ctgcgtttat tttgaaagat 8520
ttttcattgc aagaatctaa ttctcgagc gctgatgagt tgtatgatgg taaaactttg 8580
tttcatgggc cattatttcg tggaattacc aagttgttga atgtatctga tacttcacta 8640
acaactcaat gtaccaatat tgatttgact gctactgaac gtgggtcaatt tgcggatatc 8700
gaacctgtga atccttttat ggcggtatgct gcatttcaag ctatgcttgt atggggttaga 8760
aatttaagga atagtgcac tttaccaaac aattgtgaaa gagtagatat ctataaacca 8820
atagcacctg gtgaaaagta ttacactact ttgcaagctt tgggtaatac ctccggttct 8880
gttctcaagt ctgtatttta tatgcacgat gaacaaggag aagtatttct atctggaaga 8940
gctagtgttg ttgtgaatga caagatggag ttttag 8976

```

<210> 69

<211> 2991

<212> PRT

<213> *Thraustochytrium* sp.

<400> 69

5

ES 2 623 610 T3

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

ES 2 623 610 T3

115		120		125
Ala Ser His Glu Phe Tyr	Ser Arg Leu Asn Tyr	Val Val Val Asp Lys		
130	135	140		
Val Leu Arg Lys Met Gly	Leu Pro Asp Glu Asp	Val Glu Thr Ala Val		
145	150	155	160	
Glu Lys Phe Lys Ala Asn	Phe Pro Glu Trp Arg	Leu Asp Ser Phe Pro		
	165	170	175	
Gly Phe Leu Gly Asn Val	Thr Ala Gly Arg Cys	Thr Asn Thr Phe Asn		
	180	185	190	
Met Glu Gly Met Asn Cys	Val Val Asp Ala Ala	Cys Ala Ser Ser Leu		
	195	200	205	
Ile Ala Ile Lys Val Ala	Ile Asp Glu Leu Leu	His Gly Asp Cys Asp		
210	215	220		
Ala Met Ile Ala Gly Ala	Thr Cys Thr Asp Asn	Ala Leu Gly Met Tyr		
225	230	235	240	
Met Ala Phe Ser Lys Thr	Pro Val Phe Ser Thr	Asp Gln Ser Cys Leu		
	245	250	255	
Ala Tyr Asp Glu Lys Thr	Lys Gly Met Leu Ile	Gly Glu Gly Ser Ala		
	260	265	270	
Met Phe Val Leu Lys Arg	Tyr Ala Asp Ala Val	Arg Asp Gly Asp Thr		
	275	280	285	
Val His Ala Val Ile Arg	Ser Cys Ser Ser Ser	Ser Asp Gly Lys Ala		
290	295	300		
Ser Gly Ile Tyr Thr Pro	Thr Ile Ser Gly Gln	Glu Glu Ala Ile Leu		
305	310	315	320	
Arg Ala Tyr Arg Arg Ala	Gly Val Ser Pro Asn	Thr Ile Thr Leu Val		
	325	330	335	
Glu Gly His Gly Thr Gly	Thr Pro Val Gly Asp	Lys Ile Glu Leu Thr		
	340	345	350	
Ala Leu Arg Asn Val Phe	Asp Lys Ala Tyr Gly	Pro Gly His Lys Glu		
	355	360	365	
Glu Val Ala Val Gly Ser	Ile Lys Ser Gln Ile	Gly His Leu Lys Ala		
370	375	380		

Val Ala Gly Cys Ala Gly Leu Val Lys Leu Val Met Ala Leu Lys His
 385 390 395 400
 Lys Thr Leu Pro Gln Ser Ile Asn Val Glu Asn Pro Pro Asn Leu Val
 405 410 415
 Asp Gly Thr Val Ile Ser Asp Thr Thr Leu Tyr Ile Asn Thr Met Asn
 420 425 430
 Arg Pro Trp Ile Thr Lys Pro Gly Val Pro Arg Arg Ala Gly Ile Ser
 435 440 445
 Ser Phe Gly Phe Gly Gly Ala Asn Tyr His Ala Val Leu Glu Glu Phe
 450 455 460
 Glu Pro Glu Gln Thr Lys Pro Tyr Arg Leu Asn Val Ser Ala Gln Pro
 465 470 475 480
 Met Leu Leu His Ala Val Asn Ala Asn Ser Leu Gln Lys Leu Cys Glu
 485 490 495
 Asp Gln Leu Lys Leu Leu Lys Glu Ser Arg Glu Lys Cys Val Asn Thr
 500 505 510
 Lys Asn Thr Asp Tyr Val Ala Phe Ser Lys Phe Gln Asp Ser Phe Lys
 515 520 525
 Leu Lys Gly Ser Val Pro Ser Gln His Ala Arg Val Gly Phe Ala Ser
 530 535 540
 Lys Ser Ile Glu Asp Thr Ile Ser Ile Leu Ser Ala Ile Val Asn Arg
 545 550 555 560
 Phe Gln Lys Asp Ile Thr Thr Thr Ser Trp Ala Leu Pro Lys Glu Gly
 565 570 575
 Ala Ile Phe Arg Ser Thr Ala Leu Ile Asn Asp Asn Lys Ser Val Ala
 580 585 590
 Ala Leu Phe Ser Gly Gln Gly Ala Gln Tyr Thr His Met Phe Asn Asp
 595 600 605
 Val Ala Met Gln Trp Pro Gln Phe Arg Leu Cys Val Asn Asp Met Glu
 610 615 620
 Lys Ala Gln Glu Glu Val Ile Asn Asp Lys Ser Val Lys Arg Ile Ser
 625 630 635 640

Gln	Val	Met	Phe	Pro	Arg	Lys	Pro	Tyr	Ala	Arg	Glu	Ser	Pro	Leu	Asp		
				645					650						655		
Asn	Lys	Glu	Ile	Ser	Lys	Thr	Glu	Tyr	Ser	Gln	Thr	Thr	Thr	Val	Ala		
			660					665						670			
Ser	Ser	Val	Gly	Leu	Phe	Glu	Ile	Phe	Arg	Asp	Ala	Gly	Phe	Ala	Pro		
		675					680					685					
Ala	Phe	Val	Ala	Gly	His	Ser	Leu	Gly	Glu	Phe	Ser	Ala	Leu	Tyr	Ala		
	690					695						700					
Ala	Gly	Leu	Ile	Asp	Arg	Glu	Asp	Leu	Phe	Lys	Leu	Val	Cys	Asn	Arg		
705					710					715					720		
Ala	Met	Ala	Met	Arg	Asp	Ala	Pro	Lys	Lys	Ser	Ala	Asp	Gly	Ala	Met		
				725					730					735			
Ala	Ala	Val	Ile	Gly	Pro	Asn	Ala	Ser	Ser	Ile	Lys	Leu	Ser	Ala	Pro		
			740					745					750				
Glu	Val	Trp	Val	Ala	Asn	Asn	Asn	Ser	Pro	Ser	Gln	Thr	Val	Ile	Thr		
	755						760					765					
Gly	Ala	Asn	Ser	Gly	Val	Gln	Ala	Glu	Thr	Ser	Lys	Leu	Lys	Thr	Gln		
	770					775					780						
Gly	Phe	Arg	Val	Val	His	Leu	Ala	Cys	Asp	Gly	Ala	Phe	His	Ser	Pro		
785					790					795					800		
His	Met	Glu	Asn	Ala	Glu	Lys	Gln	Phe	Gln	Lys	Ala	Leu	Ser	Ala	Val		
			805						810					815			
Lys	Phe	Asn	Lys	Pro	Thr	Gly	Ser	Ser	Pro	Lys	Ile	Phe	Ser	Asn	Val		
			820					825					830				
Thr	Gly	Gly	Val	Phe	Thr	Asp	Pro	Lys	Thr	Ala	Leu	Ser	Arg	His	Met		
	835						840					845					
Thr	Ser	Ser	Val	Gln	Phe	Leu	Thr	Gln	Ile	Lys	Asn	Met	Tyr	Ala	Ala		
	850					855					860						
Gly	Ala	Arg	Val	Phe	Ile	Glu	Phe	Gly	Pro	Lys	Gln	Val	Leu	Ser	Lys		
865					870					875					880		
Leu	Val	Asn	Glu	Ile	Phe	Pro	Gly	Asp	Thr	Ser	Val	Leu	Thr	Val	Ser		
				885					890						895		

Val	Asn	Pro	Ala	Ser	Ala	Lys	Asp	Ser	Asp	Ile	Gln	Leu	Arg	Gln	Ala	
			900						905					910		
Ala	Val	Gln	Met	Ala	Val	Ala	Gly	Val	Ala	Leu	Thr	Asp	Phe	Asp	Lys	
		915					920					925				
Trp	Glu	Leu	Lys	Asp	Pro	Thr	Arg	Met	Lys	Glu	Phe	Pro	Arg	Lys	Lys	
	930					935					940					
Thr	Thr	Leu	Thr	Leu	Ser	Ala	Ala	Thr	Tyr	Val	Ser	Lys	Lys	Thr	Leu	
945					950					955					960	
Gln	Glu	Arg	Glu	Arg	Ile	Met	Asn	Asp	Gly	Arg	Thr	Val	Ser	Cys	Val	
				965					970					975		
Gln	Arg	Ile	Glu	Asn	Thr	Asn	Thr	Gly	Glu	Leu	Glu	Lys	Leu	Lys	Lys	
			980					985					990			
Gln	Leu	Gln	Asp	Lys	Glu	Asn	Glu	Val	Val	Arg	Val	Gln	Ala	Leu	Ala	
	995						1000					1005				
Thr	Gln	Ala	Ser	Ala	Asp	Leu	Gln	Asn	Thr	Lys	Ala	Glu	Leu	Gln		
	1010					1015					1020					
Lys	Ala	Gln	Ala	Thr	Lys	Ser	Ser	Asn	Ala	Ala	Ser	Asp	Ala	Val		
	1025					1030					1035					
Val	Ala	Lys	His	Lys	Ala	Ile	Leu	Leu	Ala	Met	Leu	Glu	Glu	Leu		
	1040					1045					1050					
Glu	Thr	Gly	Lys	Ala	Val	Asp	Tyr	Ser	Ser	Phe	Ser	Lys	Gly	Gln		
	1055					1060					1065					
Val	Ala	Ser	Pro	Ala	Thr	Val	Arg	Val	Val	Ser	Ala	Pro	Val	Gln		
	1070					1075					1080					
Ala	Ala	Ala	Pro	Val	Gln	Val	Ser	Ala	Ser	Val	Asp	Ser	Gly	Leu		
	1085					1090					1095					
Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val	Leu	Ala	Ser	Lys		
	1100					1105					1110					
Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met	Glu	Leu	Glu		
	1115					1120					1125					
Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser		
	1130					1135					1140					
Glu	Val	Gln	Ala	Gln	Leu	Asn	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala		

ES 2 623 610 T3

1145	1150	1155
Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys 1160 1165 1170		
Ala Glu Ile Ala Gly Gly Gln Pro Ala Ala Pro Val Gln Val Ala 1175 1180 1185		
Ala Pro Thr Gln Val Val Ala Pro Val Gln Ala Ser Ala Pro Val 1190 1195 1200		
Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val 1205 1210 1215		
Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp 1220 1225 1230		
Met Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val 1235 1240 1245		
Glu Ile Leu Ser Glu Val Gln Ala Gln Leu Ser Val Glu Ala Lys 1250 1255 1260		
Asp Val Asp Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile 1265 1270 1275		
Asp Ala Met Lys Ala Glu Ile Ala Gly Gly Gln Pro Ala Ala Pro 1280 1285 1290		
Val Gln Val Ala Ala Pro Thr Gln Val Val Ala Pro Val Gln Ala 1295 1300 1305		
Ser Ala Pro Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val 1310 1315 1320		
Val Leu Glu Val Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu 1325 1330 1335		
Ile Glu Leu Asp Met Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser 1340 1345 1350		
Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln Ala Gln Leu Ser 1355 1360 1365		
Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr Arg Thr Val 1370 1375 1380		
Gly Glu Val Ile Asp Ala Met Lys Ala Glu Ile Ser Gly Gly Gln 1385 1390 1395		

Pro	Ala	Ala	Pro	Val	Gln	Val	Ala	Ala	Pro	Thr	Gln	Ile	Val	Ala
1400						1405					1410			
Pro	Val	Gln	Val	Ser	Ala	Pro	Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys
1415						1420					1425			
Ala	Glu	Gln	Val	Val	Leu	Glu	Val	Leu	Ala	Ser	Lys	Thr	Gly	Tyr
1430						1435					1440			
Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu
1445						1450					1455			
Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln
1460						1465					1470			
Ala	Gln	Leu	Ser	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg
1475						1480					1485			
Thr	Arg	Thr	Val	Gly	Glu	Val	Ile	Asp	Ala	Met	Lys	Ala	Glu	Ile
1490						1495					1500			
Ser	Gly	Gly	Gln	Pro	Thr	Ala	Pro	Val	Gln	Val	Ala	Ala	Pro	Thr
1505						1510					1515			
Gln	Ile	Val	Ala	Pro	Val	Gln	Val	Ser	Ala	Pro	Val	Asp	Ser	Gly
1520						1525					1530			
Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val	Leu	Ala	Ser
1535						1540					1545			
Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met	Glu	Leu
1550						1555					1560			
Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu
1565						1570					1575			
Ser	Glu	Val	Gln	Ala	Gln	Leu	Ser	Val	Glu	Ala	Lys	Asp	Val	Asp
1580						1585					1590			
Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Ile	Asp	Ala	Met
1595						1600					1605			
Lys	Ala	Glu	Ile	Ser	Gly	Gly	Gln	Pro	Ala	Ala	Pro	Val	Gln	Val
1610						1615					1620			
Ala	Ala	Pro	Thr	Gln	Ile	Val	Ala	Pro	Val	Gln	Ala	Ser	Ala	Pro
1625						1630					1635			

ES 2 623 610 T3

Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu
1640						1645					1650			
Val	Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu
1655						1660					1665			
Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg
1670						1675					1680			
Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu	Ser	Val	Glu	Ala
1685						1690					1695			
Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val
1700						1705					1710			
Ile	Asp	Ala	Met	Lys	Ala	Glu	Ile	Ser	Gly	Gly	Gln	Pro	Ala	Ala
1715						1720					1725			
Pro	Val	Gln	Val	Ala	Ala	Pro	Thr	Gln	Ile	Val	Ala	Pro	Val	Gln
1730						1735					1740			
Val	Ser	Ala	Pro	Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln
1745						1750					1755			
Val	Val	Leu	Glu	Val	Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu
1760						1765					1770			
Leu	Ile	Glu	Leu	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp
1775						1780					1785			
Ser	Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu
1790						1795					1800			
Asn	Val	Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr
1805						1810					1815			
Val	Gly	Glu	Val	Ile	Asp	Ala	Met	Lys	Ala	Glu	Ile	Ala	Gly	Gly
1820						1825					1830			
Gln	Pro	Ala	Ala	Pro	Val	Gln	Val	Ala	Ala	Pro	Ala	Pro	Val	Val
1835						1840					1845			
Ala	Pro	Val	Gln	Val	Ser	Thr	Pro	Val	Asp	Ser	Gly	Leu	Leu	Ala
1850						1855					1860			
Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val	Leu	Ala	Cys	Lys	Thr	Gly
1865						1870					1875			

Tyr Glu 1880	Thr Glu Leu Ile Glu 1885	Leu Asp Met Glu Leu 1890	Glu Thr Glu
Leu Gly 1895	Ile Asp Ser Ile Lys 1900	Arg Val Glu Ile Leu 1905	Ser Glu Val
Gln Ala 1910	Gln Leu Ser Val Glu 1915	Ala Lys Asp Val Asp 1920	Ala Leu Ser
Arg Thr 1925	Arg Thr Val Gly Glu 1930	Val Ile Asp Ala Met 1935	Lys Ala Glu
Ile Ser 1940	Gly Gly Gln Pro Thr 1945	Ala Pro Val Gln Val 1950	Ala Ala Pro
Thr Gln 1955	Val Val Ala Pro Val 1960	Lys Val Ser Thr Pro 1965	Val Asp Ser
Gly Leu 1970	Leu Ala Lys Ala Glu 1975	Gln Val Val Leu Glu 1980	Val Leu Ala
Ser Lys 1985	Thr Gly Tyr Glu Thr 1990	Glu Leu Ile Glu Leu 1995	Asp Met Glu
Leu Glu 2000	Thr Glu Leu Gly Ile 2005	Asp Ser Ile Lys Arg 2010	Val Glu Ile
Leu Ser 2015	Glu Val Gln Ala Gln 2020	Leu Asn Val Glu Ala 2025	Lys Asp Val
Asp Ala 2030	Leu Ser Arg Thr Arg 2035	Thr Val Gly Glu Val 2040	Ile Asp Ala
Met Lys 2045	Ala Glu Ile Ala Gly 2050	Asp Gln Pro Ala Pro 2055	Ala Val Val
Pro Val 2060	Gln Ala Lys Ser Gly 2065	Val Ala Asn Pro Ala 2070	Leu Leu Ala
Lys Ala 2075	Glu Gln Val Val Leu 2080	Glu Val Leu Ala Ser 2085	Lys Thr Gly
Tyr Glu 2090	Thr Glu Leu Ile Glu 2095	Leu Asp Met Glu Leu 2100	Glu Thr Glu
Leu Gly 2105	Ile Asp Ser Ile Lys 2110	Arg Val Glu Ile Leu 2115	Ser Glu Val
Gln Ala	Glu Leu Ser Val Glu	Ala Lys Asp Val Asp	Ala Leu Ser

ES 2 623 610 T3

2120	2125	2130
Arg Thr 2135	Arg Thr Val Gly Glu 2140	Val Ile Asp Ala Met 2145
Ile Ala 2150	Gly Ser Ala Val Thr 2155	Val Ala Thr Leu Asp 2160
Ile Met 2165	Glu Glu Thr Asp Asp 2170	Glu Asp Glu Asp Phe 2175
Asp His 2180	Val Tyr Gly Ser Glu 2185	Cys Glu Asp Leu Ser 2190
Ser Ser 2195	Val Lys Ser Ile Pro 2200	Arg Ala Asp Lys Leu 2205
Asn Ile 2210	Ala Glu Arg Pro Ile 2215	Val Ile Val Asp Cys 2220
Leu Thr 2225	Thr Glu Leu Ala Lys 2230	Ala Ile Gly Glu Arg 2235
Ala Thr 2240	Phe Ser Ala Gln Ser 2245	Leu Val Ser Arg Gly 2250
Lys Ser 2255	Phe Thr Leu Gly Asn 2260	Thr Glu Glu Ser Glu 2265
Met Val 2270	Ser Ser Ile Glu Ser 2275	Ser Tyr Gly Lys Ile 2280
Val Tyr 2285	Gln His Phe His Asp 2290	Ser Asp Tyr Gly Met 2295
Trp Ala 2300	Leu Met Ala Ala Lys 2305	His Leu Lys Glu Ser 2310
Pro Ile 2315	Lys Asn Gly Arg Thr 2320	Phe Phe Leu Ala Val 2325
Asn Gly 2330	Lys Leu Gly Met Asp 2335	Asn Ala Ser Val His 2340
Ile Val 2345	Glu Ser Cys Gly Ile 2350	Ala Glu Arg Gly Ala 2355
Leu Cys 2360	Lys Thr Leu Asp Leu 2365	Glu Trp Pro Asn Val 2370
		Phe Ala Arg

Gly Val 2375	Asp Ile Ala Glu 2380	Gly Met Ser Tyr Ser 2385	Leu Ala Ala Glu 2385
Leu Ile 2390	Val Asp Glu Ile 2395	Ser Cys Ala Asn Leu 2400	Ser Ile Arg Glu 2400
Ser Gly 2405	Tyr Thr Ile Ser 2410	Gly Glu Arg Phe Thr 2415	Thr Glu Ala His 2415
Lys Leu 2420	Val Thr Gly Lys 2425	Pro His Ala Pro Ile 2430	Lys Lys Lys Asp 2430
Ala Phe 2435	Leu Val Ser Gly 2440	Gly Ala Arg Gly Ile 2445	Thr Pro Leu Cys 2445
Ile Arg 2450	Glu Ile Ala Lys 2455	Ala Val Lys Gly Gly 2460	Thr Tyr Ile Leu 2460
Met Gly 2465	Arg Ser Ala Leu 2470	Ala Asp Glu Pro Leu 2475	Trp Ala Asn Gly 2475
Lys Ser 2480	Gly Lys Asp Leu 2485	Asp Lys Ala Gly Leu 2490	Ala Phe Leu Lys 2490
Glu Glu 2495	Phe Ala Ala Gly 2500	Arg Gly Ser Lys Pro 2505	Thr Pro Lys Val 2505
His Lys 2510	Ser Leu Ile Asp 2515	Lys Val Leu Gly Ile 2520	Arg Glu Val Arg 2520
Ala Ser 2525	Ile Ala Asn Ile 2530	Glu Ala His Gly Ala 2535	Lys Ala Ile Tyr 2535
Leu Ser 2540	Cys Asp Val Ser 2545	Ser Ala Glu Lys Val 2550	Lys Ala Ala Val 2550
Gln Lys 2555	Val Glu Lys Glu 2560	His Leu Val Arg Ile 2565	Thr Gly Ile Val 2565
His Ala 2570	Ser Gly Val Leu 2575	Arg Asp Lys Leu Val 2580	Glu Asn Lys Thr 2580
Leu Asp 2585	Asp Phe Asn Ala 2590	Val Tyr Gly Thr Lys 2595	Val Thr Gly Leu 2595
Val Asn 2600	Leu Leu Ser Ala 2605	Val Asn Met Asn Phe 2610	Val Arg His Leu 2610

Val Met 2615	Phe Ser Ser Leu 2620	Ala Gly Tyr His Gly Asn 2625	Val Gly Gln
Ser Asp 2630	Tyr Ala Met Ala Asn 2635	Glu Ser Leu Asn Lys 2640	Ile Gly Phe
Arg Leu 2645	Gly Ala Ala Tyr Ser 2650	Gln Leu Cys Val Lys 2655	Ser Ile Cys
Phe Gly 2660	Pro Trp Asp Gly Gly 2665	Met Val Thr Pro Ala 2670	Leu Lys Lys
Gln Phe 2675	Gln Ser Met Gly Val 2680	Gln Ile Ile Pro Arg 2685	Glu Gly Gly
Ala Glu 2690	Thr Val Ala Arg Ile 2695	Val Leu Ser Ser Asn 2700	Pro Ser Gln
Val Leu 2705	Val Gly Asn Trp Gly 2710	Val Pro Pro Val Ser 2715	Pro Leu Ser
Lys Ser 2720	Ala Thr Ile Val Gln 2725	Thr Phe Thr Pro Glu 2730	Leu Asn Pro
Phe Leu 2735	Lys Ser His Gln Ile 2740	His Gly Lys Asn Val 2745	Leu Pro Met
Thr Val 2750	Ala Ile Gly Tyr Leu 2755	Ala His Leu Val Lys 2760	Asn Phe Tyr
Ala Gly 2765	His His Leu Trp Gly 2770	Val Glu Asp Ala Gln 2775	Leu Phe Ser
Gly Val 2780	Val Ile Asp His Ala 2785	Val Gln Ala Gln Val 2790	Lys Leu Thr
Glu Gln 2795	Ser Leu Asp Asp Asp 2800	Gly Lys Val Lys Val 2805	Gln Ala Val
Leu Thr 2810	Ala Ser Asn Asp Asn 2815	Gly Lys Met Val Pro 2820	Ala Tyr Lys
Ala Val 2825	Ile Val Leu Gly Lys 2830	Thr Ser Arg Pro Ala 2835	Phe Ile Leu
Lys Asp 2840	Phe Ser Leu Gln Glu 2845	Ser Asn Ser Arg Ser 2850	Ala Asp Glu

Leu	Tyr	Asp	Gly	Lys	Thr	Leu	Phe	His	Gly	Pro	Leu	Phe	Arg	Gly
2855						2860					2865			
Ile	Thr	Lys	Leu	Leu	Asn	Val	Ser	Asp	Thr	Ser	Leu	Thr	Thr	Gln
2870						2875					2880			
Cys	Thr	Asn	Ile	Asp	Leu	Thr	Ala	Thr	Glu	Arg	Gly	Gln	Phe	Ala
2885						2890					2895			
Asp	Ile	Glu	Pro	Val	Asn	Pro	Phe	Met	Ala	Asp	Ala	Ala	Phe	Gln
2900						2905					2910			
Ala	Met	Leu	Val	Trp	Val	Arg	Asn	Leu	Arg	Asn	Ser	Ala	Ser	Leu
2915						2920					2925			
Pro	Asn	Asn	Cys	Glu	Arg	Val	Asp	Ile	Tyr	Lys	Pro	Ile	Ala	Pro
2930						2935					2940			
Gly	Glu	Lys	Tyr	Tyr	Thr	Thr	Leu	Gln	Ala	Leu	Gly	Asn	Thr	Ser
2945						2950					2955			
Gly	Ser	Val	Leu	Lys	Ser	Val	Phe	Tyr	Met	His	Asp	Glu	Gln	Gly
2960						2965					2970			
Glu	Val	Phe	Leu	Ser	Gly	Arg	Ala	Ser	Val	Val	Val	Asn	Asp	Lys
2975						2980					2985			
Met	Glu	Phe												
2990														

<210> 70

<211> 5880

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 70

atggtgaaat taagtgttgg tgataatatt tgtcatgac aacgtgttgc tgttgttgg	60
atggctgtta tgtatgctgg ttgtcaaaat caacatgaat tttggcaatc tttacaaggt	120
aaaaatatga attcaaaatc gatttcacaa aatcgtttag gttctgagta tagagaagaa	180
cattttaaac ctgaaagaag taaatattcc gatacctttt gtaatgaaag atatggttgt	240
attgatgaga atgttcaaag tgaacatgaa cttttattaa aacttgcaaa agatgctatt	300
gcggatacaa aaggttctat tgatttgaat aaaaccggaa tcgttagtgg ttgcttatct	360
tttccaatgg ataatttaca aggtgattta ttaaatttgt atcaatgtca cattgaaaag	420
aaaattgggc caaatgcatt aaaagatgtg aatttatggt ctaaaagaac caccaacgga	480
aaagatgata aaaaagctta ttttgatcct gcctctttcg tagctgaaca attagatatg	540
ggaccattac attatagttt agatgctgct tgtgctctg cactttatgt attaagactt	600

gctcaagatc atttattaag tgggtgctgct gataacaatgt tatgtgggtgc atcttggttta 660
 cctgaacctt tttttatttt atctgggtttt tctacttttc atgcaatgcc attatctgggt 720
 gatgtttctg ctcccttgca taaaacttca caaggtctta cacctggtga aggtgggtgct 780
 attatgggtac ttaaacgatt aaatgatgca atccgtgatg gtgatagaat ttatgggtact 840
 ttacttgggtg ctgaattaag taatgctgggt tgtgggtttac cattgagtcc acatatgccca 900
 agtgaatttg attgtatgga aaaagcttta caaagagtac acagattacc atcatctatt 960
 caatatgttg agtgtcatgc aactgggtaca ccacaagggtg ataaagtga aattgatgct 1020
 atgacaaaat gttttgggtga acattttacca aggtttgggtt caacgaaagg gaattttgggt 1080
 catacaacttg ttgctgctgg ttttgctgggt atgtgtaaag ttttattatc aatgcaatat 1140
 ggtgaaatac caccaactcc aggtcttgaa aatccagaca atattatgca tgatttagtt 1200
 gttactgaaa caattccatg gcctaataca aatgggtgatt tgaaacgtgc atgtttatct 1260
 gcttttggat tcggtggtac taatgcacat gctgtatttg aagagtatcg ttcagattta 1320
 caagcaaata aaactcttga aaatgaaagt aaaagtcatg aaatcttttc ttcatttaaa 1380
 attgctattg ttgggtatgga atctgaattt ggtactttga aaggattaca agaatttgaa 1440
 cgtgctattt acaatgggtg tcatgggtgca tgtgatttac ctgaaaatag atggagattt 1500
 cttggagaag ataaagaatt ttacaagct tgtgggtttac aaaaattacc aagagggttgt 1560
 tatattaaag aagtggaaac tgatttttaa aggttacgtt taccaatgat acaggaggat 1620
 attctaagac ctttacagtt gttagctggt togattatcg acagagcaact taacgcatct 1680
 ggtgtttaa acaaatggcaa agttgcagtt ttagttggat taggtactga tcttgaatta 1740
 tatcgtcatc gtgctcgtgt tgcattaaag gaacgcctcc aaactgcggt caaagaagat 1800
 attcctttac ttgaaaagtt aatgaactat gtcaatgata gaggtacaag tacatcatat 1860
 acatcttata ttggaaattt ggttgcaact cgagtttcat cattatgggg ttttactgggt 1920
 ccatcattca cgattactga aggtgaaaat tccgtatata gttgtcttga tttgggaaga 1980
 tgggtcttag ctaatgggtga agtagatgct gttgttgttg ccgggggttga tttatgtgggt 2040
 agtgcctgaaa atctttttgt aaaatctcgt agaagtaaag tttccacaca aaatgaacca 2100
 tttgcaaatt ttgaatcaaa tgctgatgga tattttgctg gagatggttg tggagctttg 2160
 gttttgaaac gattgagtga ttgtacggat tcaactgaaa aaatttatgc aacgggtggat 2220
 tcaattgctg ttgggtgatga agttggccca actattaaac aagctttgaa gaatgcatcc 2280
 atagcagcga aagatattga actggcagag ctatcagcaa gttcaggcaa acatcattct 2340
 ggtagaatca cttgtgaaga tgaactaaat gaactgggtg aaattttcaa tgaagggtata 2400
 caaagagttg caattggtag tgtgaaagct aatggtggag atgttggata tgcactctgggt 2460
 gcagcaagtt taatcaaaac ggctttgtgc ctgtacaacc gatatttacc aaagttacca 2520

aattggaata	agccaacgaa	agatgttgaa	tgggtccaaat	cattttttgt	atgtgaacat	2580
tctagagcat	ggttgaaaaa	tgttgatgaa	aatagacatg	ctgtcgtttc	tggagtttgc	2640
gaaaatgggt	cgtgttatgg	aatcgtaatg	tctgatgtac	aaggacatca	tgaagaatcg	2700
aatcttggtt	gtttagacaa	aatgaacca	aaagtactgg	gtatttacgg	agattcagtt	2760
gatgatatcc	tagttcagct	caacaaatat	cttgaaaaat	tccttcaaga	aactggaacg	2820
gctgcggtg	cacaaaaagt	taaatcacct	acaatagata	ttgactccaa	tgtgtttgct	2880
gagatgctta	atctaccgca	ggataaaaa	aaaaaatttg	cggtcgcatt	ggttaccaca	2940
ccaaataaac	tccagcgtga	aatagaactt	gctgtgaagg	gtattccacg	ttgcgtaaaa	3000
gcaaaaagag	attggtgttc	tccatctgga	agtatttttg	cttgtaatcc	actcaaaagt	3060
gataatattg	catttatgta	tggatgaagg	cgaagcccat	atgctggact	gggatatgat	3120
ttgcatcgaa	tttggcctat	gctacacgag	ttggttaaca	atagaactac	agaactttgg	3180
gatcaagggt	atagttggta	tttacctcga	tctagctctg	ttgctgaaaa	agaaaaagtc	3240
ttcggagatt	ttgataagaa	tcaaattgaa	atgttttagat	tgggtatttt	tgtatcaatg	3300
tgtttcactg	atatggccac	tgaacttttg	ggtttaaaac	ccaaagccgc	gtttggttta	3360
agtttgggtg	aaatatctat	gctttttgca	ttttctaaaa	agaataccaa	gttgtccaaa	3420
gaattgaccc	gtcgtctaaa	agaagcaaaa	gtttgggcat	cacaattagc	tgttgaaattt	3480
gcagctattc	gagattttgt	gaatattcca	gctgataaat	ctattgatga	attttggcaa	3540
gggtattttg	tttacgcaaa	tcgaaccctg	gtcgagaaca	caattgggga	gaataaattt	3600
gttcgtttgt	tgattgtaaa	tgattcgcaa	agttgtctaa	ttgccgggaa	accagatgaa	3660
tgtcaaaaag	ttattgagaa	gcttcatttg	aagctaccgg	cggttccagt	aactcagggg	3720
atgatcgggc	attgcccaga	agcaattcct	tatctagatc	aatcagtc	tattcatgaa	3780
atgcttgaaa	ttccaaaacc	cgaaaatgtg	aaattgttta	caactagtga	aaacagagaa	3840
ttagtgtcga	tgaagattc	cgtgtcaaaa	ttggttgctg	agatttatca	gcattgtgct	3900
gattttccaa	acatcgtgaa	caagggttaa	gaaacttgca	aaactgatat	atttattgaa	3960
ttgggatcga	acaattatcg	atctggagct	gtcaaaaaca	ttttaggtcc	agaaatcggt	4020
tctgttgcaa	ttgataggca	aatgaaact	gcattgggtc	aactaatgaa	gatgggtgca	4080
tcgttgataa	gtcatcgagt	tccgggtgtt	gaattgaaaa	aactctatca	tcctgaattg	4140
ctgaaatttg	atccacaggc	aaaaccgaat	cgtttcatca	gaaatataga	actgaatgga	4200
ttttttgatc	gtacgaatat	tattgttgat	aagcaactat	cccctgcgga	tccgaaactc	4260
gctgaaattg	tgaacaatcg	aaatatgcct	aaagataatg	tttatgtacc	aattgaacgg	4320
gtgaaaacga	tgataaaggc	ggaaccagct	aatttacaag	tcagcgtggg	aagtaaacca	4380
gttggttactg	aaagaattag	ttcggacgat	aatctatttg	aaaagttgtc	agaaattaca	4440
aaatcttttg	atgggtgtaa	tgcgtgtact	gaagcaatgt	tgggagactc	tggattttctc	4500


```

aaaacatatg aggttgacta tcctttgtac acaggtgcc a tggctaaagg aattgcgtct 4560
gctgatttgg ttattgctgc tggtaaataa aagatcttgg catcatttgg agctgggtggg 4620
ttggccttac aagtggtaga agatgccatt aaacaaatta aagctgaatt ggggaacggt 4680
ccgtttgctg taaatttgat tcattcacca ttcgatccta gcttggagaa gggtaacggt 4740
gatctttttc taaaatataa cgttcgattt gttgaagtat cgcatttat gtcattaacc 4800
cctcaggttg tacgatacag agccgctggg ttggccaaag caagagatgg atctgtgaaa 4860
attcaaaatc gtattattgc caaaatttca agaacagagt tagcggaact gttcttgaaa 4920
ccagcaccca aaaatatttt agatgcattg gttgcggatg gatctattag toaagaacaa 4980
gcccaacttg cattacttgt gccaatggct gatgatatta ctgtggaagc tgattctggt 5040
gggcatactg acaatcgacc aattcatgtt ttgttacctt tgataattca gcaaagaaat 5100
agaatttgta aacaataccc aaaacattta aaagttcgaa tcggagcagc tgggtggtatt 5160
ggatgcccga aggcagcatt tgctgcgttt gagatgggtg ctgcatacat tgcaactgga 5220
acggtaaatc aactttcaaa ggaagcaggt acttgtgact atgtacgtaa agtattgaat 5280
aaagctacat attcggtatg taccatggct ccagccgcag atatgttcga toatggtgtt 5340
gaattacaag ttttgaagaa aggtactatg tttccttcac gtgctaaaaa actatacgat 5400
ttgttcaaaa aatacaaatac gattgaggaa ttaccagcag atgaggtgaa aaaacttgag 5460
caaaaagttt tcaaaaagtc gtttgatgaa gtatgggatg agaccaagaa ttactatatt 5520
aatcgtttac attctcccga aaaaattgaa cgtgctgaaa gagatgcaaa acttaaaatg 5580
tcgttatgtt ttcgttggtt tttgtcgaag tcttccagat gggctaatac cgggtgaatct 5640
ggaagagtgc aggattatca aatttggtgt ggtccagcaa ttgggtcata taatgatttt 5700
gcgaaaggat caccatgttt ggatcctgag attttgggtg gttttccaag tgttgttcag 5760
attaataaac atatttttacg tgggtgcttg tttctatcaa gactctctca gttgaaatat 5820
ctgaatttta actatgagga attagatacg ttaacatact ctgcacgaa ttttatttaa 5880

```

<210> 71

<211> 1959

<212> PRT

<213> *Thraustochytrium* sp

<400> 71

5

ES 2 623 610 T3

Met Val Lys Leu Ser Val Gly Asp Asn Ile Cys His Asp Gln Arg Val
1 5 10 15

Ala Val Val Gly Met Ala Val Met Tyr Ala Gly Cys Gln Asn Gln His
20 25 30

Glu Phe Trp Gln Ser Leu Gln Gly Lys Asn Met Asn Ser Lys Ser Ile
35 40 45

ES 2 623 610 T3

Ser	Gln	Asn	Arg	Leu	Gly	Ser	Glu	Tyr	Arg	Glu	Glu	His	Phe	Lys	Pro	50	55	60
Glu	Arg	Ser	Lys	Tyr	Ser	Asp	Thr	Phe	Cys	Asn	Glu	Arg	Tyr	Gly	Cys	65	70	75
Ile	Asp	Glu	Asn	Val	Gln	Ser	Glu	His	Glu	Leu	Leu	Leu	Lys	Leu	Ala	85	90	95
Lys	Asp	Ala	Ile	Ala	Asp	Thr	Lys	Gly	Ser	Ile	Asp	Leu	Asn	Lys	Thr	100	105	110
Gly	Ile	Val	Ser	Gly	Cys	Leu	Ser	Phe	Pro	Met	Asp	Asn	Leu	Gln	Gly	115	120	125
Asp	Leu	Leu	Asn	Leu	Tyr	Gln	Cys	His	Ile	Glu	Lys	Lys	Ile	Gly	Pro	130	135	140
Asn	Ala	Leu	Lys	Asp	Val	Asn	Leu	Trp	Ser	Lys	Arg	Thr	Thr	Asn	Gly	145	150	155
Lys	Asp	Asp	Lys	Lys	Ala	Tyr	Phe	Asp	Pro	Ala	Ser	Phe	Val	Ala	Glu	165	170	175
Gln	Leu	Asp	Met	Gly	Pro	Leu	His	Tyr	Ser	Leu	Asp	Ala	Ala	Cys	Ala	180	185	190
Ser	Ala	Leu	Tyr	Val	Leu	Arg	Leu	Ala	Gln	Asp	His	Leu	Leu	Ser	Gly	195	200	205
Ala	Ala	Asp	Thr	Met	Leu	Cys	Gly	Ala	Ser	Cys	Leu	Pro	Glu	Pro	Phe	210	215	220
Phe	Ile	Leu	Ser	Gly	Phe	Ser	Thr	Phe	His	Ala	Met	Pro	Leu	Ser	Gly	225	230	235
Asp	Val	Ser	Ala	Pro	Leu	His	Lys	Thr	Ser	Gln	Gly	Leu	Thr	Pro	Gly	245	250	255
Glu	Gly	Gly	Ala	Ile	Met	Val	Leu	Lys	Arg	Leu	Asn	Asp	Ala	Ile	Arg	260	265	270
Asp	Gly	Asp	Arg	Ile	Tyr	Gly	Thr	Leu	Leu	Gly	Ala	Glu	Leu	Ser	Asn	275	280	285
Ala	Gly	Cys	Gly	Leu	Pro	Leu	Ser	Pro	His	Met	Pro	Ser	Glu	Phe	Asp	290	295	300

Cys Met Glu Lys Ala Leu Gln Arg Val His Arg Leu Pro Ser Ser Ile
 305 310 315 320
 Gln Tyr Val Glu Cys His Ala Thr Gly Thr Pro Gln Gly Asp Lys Val
 325 330 335
 Glu Ile Asp Ala Met Thr Lys Cys Phe Gly Glu His Leu Pro Arg Phe
 340 345 350
 Gly Ser Thr Lys Gly Asn Phe Gly His Thr Leu Val Ala Ala Gly Phe
 355 360 365
 Ala Gly Met Cys Lys Val Leu Leu Ser Met Gln Tyr Gly Glu Ile Pro
 370 375 380
 Pro Thr Pro Gly Leu Glu Asn Pro Asp Asn Ile Met His Asp Leu Val
 385 390 395 400
 Val Thr Glu Thr Ile Pro Trp Pro Asn Thr Asn Gly Asp Leu Lys Arg
 405 410 415
 Ala Cys Leu Ser Ala Phe Gly Phe Gly Gly Thr Asn Ala His Ala Val
 420 425 430
 Phe Glu Glu Tyr Arg Ser Asp Leu Gln Ala Asn Lys Thr Leu Glu Asn
 435 440 445
 Glu Ser Lys Ser His Glu Ile Phe Ser Ser Phe Lys Ile Ala Ile Val
 450 455 460
 Gly Met Glu Ser Glu Phe Gly Thr Leu Lys Gly Leu Gln Glu Phe Glu
 465 470 475 480
 Arg Ala Ile Tyr Asn Gly Gly His Gly Ala Cys Asp Leu Pro Glu Asn
 485 490 495
 Arg Trp Arg Phe Leu Gly Glu Asp Lys Glu Phe Leu Gln Ala Cys Gly
 500 505 510
 Leu Gln Lys Leu Pro Arg Gly Cys Tyr Ile Lys Glu Val Glu Thr Asp
 515 520 525
 Phe Lys Arg Leu Arg Leu Pro Met Ile Gln Glu Asp Ile Leu Arg Pro
 530 535 540
 Leu Gln Leu Leu Ala Val Ser Ile Ile Asp Arg Ala Leu Asn Ala Ser
 545 550 555 560

ES 2 623 610 T3

Gly Val Lys Pro Asn Gly Lys Val Ala Val Leu Val Gly Leu Gly Thr
 565 570 575
 Asp Leu Glu Leu Tyr Arg His Arg Ala Arg Val Ala Leu Lys Glu Arg
 580 585 590
 Leu Gln Thr Ala Val Lys Glu Asp Ile Pro Leu Leu Glu Lys Leu Met
 595 600 605
 Asn Tyr Val Asn Asp Arg Gly Thr Ser Thr Ser Tyr Thr Ser Tyr Ile
 610 615 620
 Gly Asn Leu Val Ala Thr Arg Val Ser Ser Leu Trp Gly Phe Thr Gly
 625 630 635 640
 Pro Ser Phe Thr Ile Thr Glu Gly Glu Asn Ser Val Tyr Arg Cys Leu
 645 650 655
 Asp Leu Gly Arg Trp Phe Leu Ala Asn Gly Glu Val Asp Ala Val Val
 660 665 670
 Val Ala Gly Val Asp Leu Cys Gly Ser Ala Glu Asn Leu Phe Val Lys
 675 680 685
 Ser Arg Arg Ser Lys Val Ser Thr Gln Asn Glu Pro Phe Ala Asn Phe
 690 695 700
 Glu Ser Asn Ala Asp Gly Tyr Phe Ala Gly Asp Gly Cys Gly Ala Leu
 705 710 715 720
 Val Leu Lys Arg Leu Ser Asp Cys Thr Asp Ser Thr Glu Lys Ile Tyr
 725 730 735
 Ala Thr Val Asp Ser Ile Ala Val Gly Asp Glu Val Gly Pro Thr Ile
 740 745 750
 Lys Gln Ala Leu Lys Asn Ala Ser Ile Ala Ala Lys Asp Ile Glu Leu
 755 760 765
 Ala Glu Leu Ser Ala Ser Ser Gly Lys His His Ser Gly Arg Ile Thr
 770 775 780
 Cys Glu Asp Glu Leu Asn Glu Leu Gly Glu Ile Phe Asn Glu Gly Ile
 785 790 795 800
 Gln Arg Val Ala Ile Gly Ser Val Lys Ala Asn Val Gly Asp Val Gly
 805 810 815
 Tyr Ala Ser Gly Ala Ala Ser Leu Ile Lys Thr Ala Leu Cys Leu Tyr

ES 2 623 610 T3

820	825	830
Asn Arg Tyr Leu Pro Lys Leu Pro Asn Trp Asn Lys Pro Thr Lys Asp		
835	840	845
Val Glu Trp Ser Lys Ser Phe Phe Val Cys Glu His Ser Arg Ala Trp		
850	855	860
Leu Lys Asn Val Asp Glu Asn Arg His Ala Val Val Ser Gly Val Cys		
865	870	875
Glu Asn Gly Ser Cys Tyr Gly Ile Val Met Ser Asp Val Gln Gly His		
885	890	895
His Glu Glu Ser Asn Leu Val Ser Leu Asp Lys Asn Glu Pro Lys Val		
900	905	910
Leu Gly Ile Tyr Gly Asp Ser Val Asp Asp Ile Leu Val Gln Leu Asn		
915	920	925
Lys Tyr Leu Glu Lys Phe Leu Gln Glu Thr Gly Thr Ala Ala Ala Ala		
930	935	940
Gln Lys Val Lys Ser Pro Thr Ile Asp Ile Asp Ser Asn Val Phe Ala		
945	950	955
Glu Met Leu Asn Leu Pro Gln Asp Lys Asn Lys Lys Phe Ala Val Ala		
965	970	975
Leu Val Thr Thr Pro Asn Lys Leu Gln Arg Glu Ile Glu Leu Ala Val		
980	985	990
Lys Gly Ile Pro Arg Cys Val Lys Ala Lys Arg Asp Trp Cys Ser Pro		
995	1000	1005
Ser Gly Ser Ile Phe Ala Cys Asn Pro Leu Lys Ser Asp Asn Ile		
1010	1015	1020
Ala Phe Met Tyr Gly Glu Gly Arg Ser Pro Tyr Ala Gly Leu Gly		
1025	1030	1035
Tyr Asp Leu His Arg Ile Trp Pro Met Leu His Glu Leu Val Asn		
1040	1045	1050
Asn Arg Thr Thr Glu Leu Trp Asp Gln Gly Asp Ser Trp Tyr Leu		
1055	1060	1065
Pro Arg Ser Ser Ser Val Ala Glu Lys Glu Lys Val Phe Gly Asp		
1070	1075	1080

Phe	Asp	Lys	Asn	Gln	Ile	Glu	Met	Phe	Arg	Leu	Gly	Ile	Phe	Val
1085						1090					1095			
Ser	Met	Cys	Phe	Thr	Asp	Met	Ala	Thr	Glu	Leu	Leu	Gly	Leu	Lys
1100						1105					1110			
Pro	Lys	Ala	Ala	Phe	Gly	Leu	Ser	Leu	Gly	Glu	Ile	Ser	Met	Leu
1115						1120					1125			
Phe	Ala	Phe	Ser	Lys	Lys	Asn	Thr	Lys	Leu	Ser	Lys	Glu	Leu	Thr
1130						1135					1140			
Arg	Arg	Leu	Lys	Glu	Ala	Lys	Val	Trp	Ala	Ser	Gln	Leu	Ala	Val
1145						1150					1155			
Glu	Phe	Ala	Ala	Ile	Arg	Asp	Leu	Trp	Asn	Ile	Pro	Ala	Asp	Lys
1160						1165					1170			
Ser	Ile	Asp	Glu	Phe	Trp	Gln	Gly	Tyr	Phe	Val	Tyr	Ala	Asn	Arg
1175						1180					1185			
Thr	Leu	Val	Glu	Asn	Thr	Ile	Gly	Glu	Asn	Lys	Phe	Val	Arg	Leu
1190						1195					1200			
Leu	Ile	Val	Asn	Asp	Ser	Gln	Ser	Cys	Leu	Ile	Ala	Gly	Lys	Pro
1205						1210					1215			
Asp	Glu	Cys	Gln	Lys	Val	Ile	Glu	Lys	Leu	His	Leu	Lys	Leu	Pro
1220						1225					1230			
Ala	Val	Pro	Val	Thr	Gln	Gly	Met	Ile	Gly	His	Cys	Pro	Glu	Ala
1235						1240					1245			
Ile	Pro	Tyr	Leu	Asp	Gln	Ile	Ser	His	Ile	His	Glu	Met	Leu	Glu
1250						1255					1260			
Ile	Pro	Lys	Pro	Glu	Asn	Val	Lys	Leu	Phe	Thr	Thr	Ser	Glu	Asn
1265						1270					1275			
Arg	Glu	Leu	Val	Ser	Met	Lys	Asp	Ser	Val	Ser	Lys	Leu	Val	Ala
1280						1285					1290			
Glu	Ile	Tyr	Gln	His	Val	Ala	Asp	Phe	Pro	Asn	Ile	Val	Asn	Lys
1295						1300					1305			
Val	Lys	Glu	Thr	Cys	Lys	Thr	Asp	Ile	Phe	Ile	Glu	Leu	Gly	Ser
1310						1315					1320			

Asn	Asn	Tyr	Arg	Ser	Gly	Ala	Val	Lys	Thr	Ile	Leu	Gly	Pro	Glu
1325						1330					1335			
Ile	Val	Ser	Val	Ala	Ile	Asp	Arg	Gln	Asn	Glu	Thr	Ala	Trp	Gly
1340						1345					1350			
Gln	Leu	Met	Lys	Met	Val	Ala	Ser	Leu	Ile	Ser	His	Arg	Val	Pro
1355						1360					1365			
Gly	Val	Glu	Leu	Lys	Lys	Leu	Tyr	His	Pro	Glu	Leu	Leu	Lys	Phe
1370						1375					1380			
Asp	Pro	Gln	Ala	Lys	Pro	Asn	Arg	Phe	Ile	Arg	Asn	Ile	Glu	Leu
1385						1390					1395			
Asn	Gly	Phe	Phe	Asp	Arg	Thr	Asn	Ile	Ile	Val	Asp	Lys	Gln	Leu
1400						1405					1410			
Ser	Pro	Ala	Asp	Pro	Lys	Leu	Ala	Glu	Ile	Val	Asn	Asn	Arg	Asn
1415						1420					1425			
Met	Pro	Lys	Asp	Asn	Val	Tyr	Val	Pro	Ile	Glu	Arg	Val	Lys	Thr
1430						1435					1440			
Met	Ile	Lys	Ala	Glu	Pro	Ala	Asn	Leu	Gln	Val	Ser	Val	Gly	Ser
1445						1450					1455			
Lys	Pro	Val	Val	Thr	Glu	Arg	Ile	Ser	Ser	Asp	Asp	Asn	Leu	Phe
1460						1465					1470			
Glu	Lys	Leu	Ser	Glu	Ile	Thr	Lys	Ser	Phe	Asp	Gly	Val	Asn	Ala
1475						1480					1485			
Cys	Thr	Glu	Ala	Met	Leu	Gly	Asp	Ser	Gly	Phe	Leu	Lys	Thr	Tyr
1490						1495					1500			
Glu	Val	Asp	Tyr	Pro	Leu	Tyr	Thr	Gly	Ala	Met	Ala	Lys	Gly	Ile
1505						1510					1515			
Ala	Ser	Ala	Asp	Leu	Val	Ile	Ala	Ala	Gly	Lys	Ser	Lys	Ile	Leu
1520						1525					1530			
Ala	Ser	Phe	Gly	Ala	Gly	Gly	Leu	Ala	Leu	Gln	Val	Val	Glu	Asp
1535						1540					1545			
Ala	Ile	Lys	Gln	Ile	Lys	Ala	Glu	Leu	Gly	Asn	Gly	Pro	Phe	Ala
1550						1555					1560			

Val	Asn	Leu	Ile	His	Ser	Pro	Phe	Asp	Pro	Ser	Leu	Glu	Lys	Gly
1565						1570					1575			
Asn	Val	Asp	Leu	Phe	Leu	Lys	Tyr	Asn	Val	Arg	Phe	Val	Glu	Val
1580						1585					1590			
Ser	Ala	Phe	Met	Ser	Leu	Thr	Pro	Gln	Val	Val	Arg	Tyr	Arg	Ala
1595						1600					1605			
Ala	Gly	Leu	Ala	Lys	Ala	Arg	Asp	Gly	Ser	Val	Lys	Ile	Gln	Asn
1610						1615					1620			
Arg	Ile	Ile	Ala	Lys	Ile	Ser	Arg	Thr	Glu	Leu	Ala	Glu	Leu	Phe
1625						1630					1635			
Leu	Lys	Pro	Ala	Pro	Lys	Asn	Ile	Leu	Asp	Ala	Leu	Val	Ala	Asp
1640						1645					1650			
Gly	Ser	Ile	Ser	Gln	Glu	Gln	Ala	Gln	Leu	Ala	Leu	Leu	Val	Pro
1655						1660					1665			
Met	Ala	Asp	Asp	Ile	Thr	Val	Glu	Ala	Asp	Ser	Gly	Gly	His	Thr
1670						1675					1680			
Asp	Asn	Arg	Pro	Ile	His	Val	Leu	Leu	Pro	Leu	Ile	Ile	Gln	Gln
1685						1690					1695			
Arg	Asn	Arg	Ile	Cys	Lys	Gln	Tyr	Pro	Lys	His	Leu	Lys	Val	Arg
1700						1705					1710			
Ile	Gly	Ala	Ala	Gly	Gly	Ile	Gly	Cys	Pro	Lys	Ala	Ala	Phe	Ala
1715						1720					1725			
Ala	Phe	Glu	Met	Gly	Ala	Ala	Tyr	Ile	Ala	Thr	Gly	Thr	Val	Asn
1730						1735					1740			
Gln	Leu	Ser	Lys	Glu	Ala	Gly	Thr	Cys	Asp	Tyr	Val	Arg	Lys	Val
1745						1750					1755			
Leu	Asn	Lys	Ala	Thr	Tyr	Ser	Asp	Val	Thr	Met	Ala	Pro	Ala	Ala
1760						1765					1770			
Asp	Met	Phe	Asp	His	Gly	Val	Glu	Leu	Gln	Val	Leu	Lys	Lys	Gly
1775						1780					1785			
Thr	Met	Phe	Pro	Ser	Arg	Ala	Lys	Lys	Leu	Tyr	Asp	Leu	Phe	Lys
1790						1795					1800			
Lys	Tyr	Lys	Ser	Ile	Glu	Glu	Leu	Pro	Ala	Asp	Glu	Val	Lys	Lys

1805		1810		1815
Leu Glu Gln Lys Val Phe Lys Lys Ser Phe Asp Glu Val Trp Asp				
1820		1825		1830
Glu Thr Lys Asn Tyr Tyr Ile Asn Arg Leu His Ser Pro Glu Lys				
1835		1840		1845
Ile Glu Arg Ala Glu Arg Asp Ala Lys Leu Lys Met Ser Leu Cys				
1850		1855		1860
Phe Arg Trp Tyr Leu Ser Lys Ser Ser Arg Trp Ala Asn Thr Gly				
1865		1870		1875
Glu Ser Gly Arg Val Gln Asp Tyr Gln Ile Trp Cys Gly Pro Ala				
1880		1885		1890
Ile Gly Ser Tyr Asn Asp Phe Ala Lys Gly Ser Pro Cys Leu Asp				
1895		1900		1905
Pro Glu Ile Leu Gly Ser Phe Pro Ser Val Val Gln Ile Asn Lys				
1910		1915		1920
His Ile Leu Arg Gly Ala Cys Phe Tyr Gln Arg Leu Ser Gln Leu				
1925		1930		1935
Lys Tyr Leu Asn Phe Asn Tyr Glu Glu Leu Asp Thr Leu Thr Tyr				
1940		1945		1950
Ser Ala Ser Asn Phe Ile				
1955				

<210> 72

<211> 4368

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 72

atggttggtt	tacaaatgaa	aaagaaacca	gtatgggaga	tgagtaagga	agaacaaagt	60
tctggaaaga	atgttggtatt	tgactatgat	gaattggttg	aatttgctga	aggtgatatt	120
ggtaaagtct	ttggacctaa	gtttgatatt	atcgataagt	atagtcgacg	tgtacgttta	180
cctgcgagag	aatatcttct	agttaccaga	gttactttga	tggatgctga	agttgggaat	240
ttcagagttg	gatctagaat	ggttactgaa	tatgatgttc	cagtaaatgg	tgaactttca	300
caaggtggtg	atgttccatg	ggctgttctt	gttgaatctg	gacaatgtga	tcttatgtta	360
atatcttata	tgggtattga	ttttcaatgt	aaaggtgatc	gtgtctatcg	attattaaat	420
actacgttga	cgttttacgg	tgttgctcat	gagggtgaaa	cactagtata	cgatattcgt	480

ES 2 623 610 T3

gtaactggat ttgcaaaagg tatgcacggg gaaatctcca tgtttttttt tgaatatgat	540
tgttatgtga atggacgatt attaatcgaa atgagagatg gttgtgcggg attttttact	600
gatgaagaac ttgcagcagg taaaggagtt attaaaactg ttgctgaact tcataaaaga	660
aaatctattg ttccaaaatc cattaaacct tttgctctaa atccagcagt acacaaaaca	720
atgttttctg aaaatgatat ggaaaaattg tgtgagcgtc aatgggaaaa tgtattgggt	780
agtggacttc aaggatttga ctacaagtta tgtgcacgga aaatgcttat gattgatcgt	840
attactaaaa tacaacataa tgggtggtgca tatggtcttg gattattggt tggcgaaaaa	900
attcttgaac gtgatcattg gtattttcca tgccattttg taaaggatca agttatggct	960
ggctcacttg ttagtgatgg ttgcagtcag ctactaaaac tttatatgtt atggttgggt	1020
ttacatgatg tggttccaga ttttcaattt cgtccagttc ctggacaacc aaataaagtt	1080
cgttgccgtg gacaaattag tccacatcgt ggtaaaacttg tttatgttat ggaaataaga	1140
gaaatgggat tcaatgaatc aactggacaa ccatatgcta ttgctgatgt tgatattatt	1200
gatgtaaaact atgaacttgg tcaatcattt gatatggctg atattgatag ttatggacgt	1260
ggtaatttgt caaagaaaat tgtggttgat tttaaaggaa ttgctttgca aatggaaggt	1320
accgtgaaat catcaaatat cattgattct tcacaaaaat caactattat acaaccacct	1380
ccaaattgtc ttcgtggtga tccactggca ccatcacaag ttacatggca tccaatggca	1440
ggagttaatg gggcaccagc tccttcattt agtccatctg attatccacc acgtgctgtt	1500
tgcttcaaac catttctctg taatccttta gataacgac atacacctgg taaaatgcct	1560
ttaacatggt ttaatatgtc cgagtttatg tgtggttaaag tatcaaattg tcttggacca	1620
gaatttaaga gatttgataa ctctaaaaca tccagaagtc ctgcctttga tcttgcactt	1680
gttacacgtg ttgtgagtgt atcagatatg gaatttaaac ctcatTTaaa tattgatgtt	1740
aatccaagta agggatcaat gataggtgaa tttgattgcc ctgcagatgc gtggtttttt	1800
caaggatcat gtaacgatgg tcatatgccg tattctattg ttatggaaat tgctcttcaa	1860
acttctggtg tattaacttc agttttgaaa gcaccttga ctatggataa agatgatatt	1920
cttttccgca atttggatgc cactgctgaa atggttcgaa gtgatgttga ttgtagaggt	1980
aaaactatca aaaactttac tcaatgtacc ggttacagta tgctcggaaa aatgggaatt	2040
catagattca catttgaatt atctgttgat gatgtagttt tctacaaagg atcaacatct	2100
tttggttggt tcaccctga agtattcgag tcacaagttg gtcttgataa tggtaaaaaa	2160
gtacaaccat ggtatttggg acaaaaatca tctaagttag taacttatga cgttgcgctc	2220
actgctggca aggataagtt attttcaaag attggatcta aggatgcaca agttcaaaga	2280
agaaatacac aatgtgagtt tctagatact atgcatatta ttccaaatac tggaaagtac	2340
aacaaaggtt atgctcatgg agaaaagaaa gttaatccaa acgactggtt cttttcctgt	2400
catttctggt ttgatcctgt gatgcctggt tcattaggta ttgaaagtat gtttcaactc	2460

attgaagcat	tttcaattga	tcaaggaatc	gcttcaaaac	atgggtattgt	gaatccaact	2520
tttgctcatt	ccaatggaaa	aacttcttgg	aaatacagag	gtcaattgaa	taacaaaggt	2580
aaacgaatgg	atagtgaaat	tcatatcaaa	gatattgtca	aaaatgctga	tggtactgtt	2640
gatttgattg	ctgatggatt	tttattgggt	gattcactaa	gagtatactc	tgcagatgat	2700
cttcgcgtaa	aaattgtacc	gggaacccaa	gctgcaccta	aatcagtagc	tgctgctcca	2760
agacatgttg	caacaccaat	tccaggagtg	ccttcgaata	caagcagtgt	tgaaatcagt	2820
ttggaatctt	tgaagaaaga	attgttaaata	cttgagaaac	cattgtatct	tgaaacttcc	2880
aatcatattg	taaaacaatt	cggtgacgtt	aacaatggcc	aagcatccgt	tattccacca	2940
tgcaccatca	atgatttggg	tgagcgtagt	tttatggaaa	catacaatgt	tggtgcacca	3000
ctttacactg	gagccatggc	taaagggtatt	gcactctgtg	atttggtaat	tgcagctggg	3060
aaaagaaaaa	ttttgggttc	ttttggcgct	ggaggcttac	caatgcactt	ggttcgtgct	3120
tctgttgaaa	aaatccaagc	cgcacttcca	gaaggtccat	acgctgtcaa	cttgattcat	3180
agtccattcg	actcaaactc	tgaaaaggga	aatgtagatc	tatttttgga	aaaagggtgtt	3240
catgttggtg	aagcatctgc	attcactgct	ctgaccactc	aagtagttcg	ttaccgtgca	3300
tgtggtttat	ctcgggctaa	agacggatct	gtattgatca	aaaatagaat	catcggtaaa	3360
gtttcaagaa	cogaattggc	tgaaatgttt	ttcagacctg	caccacaaaa	cttgcttgac	3420
aagcttattg	ctagtggaga	aatcactaaa	gaacaagctt	cattggcttt	ggaagtacca	3480
atggctgatg	atgtagctgt	tgaagctgat	agcggtgga	atactgataa	tagaccaatt	3540
catgtaatcc	tacctttgat	tatcaatcta	cgaaatagaa	ttcataaaga	atgtgggtttt	3600
cctgctgctt	tgagagttcg	cgttggtgct	ggtggtggaa	ttgggtgtcc	aagtgtgca	3660
gttgctgcat	tcaatatggg	agctgcattc	ttgattactg	gcagcgtcaa	ccaagttagc	3720
aaacaatctg	gtacgtgtga	tatcgttaga	aagcaattat	ctgaagcttc	gtattcagat	3780
attaccatgg	caccagcggc	tgatatgttt	gatcaaggag	tcgagcttca	agtattaaaa	3840
aaaggaacta	tgtttccatc	tcgtgcaaag	aaattgtatg	aattattctg	tatgtacaac	3900
tcatttgatg	acatgccaaa	aagcgaactt	caaagactag	agaagcgaat	ttttcaaaaa	3960
tcgcttgccg	aagtttggga	agaaactaaa	gatttttata	tcaatcgttt	gaataatcct	4020
gagaagattg	aacatgctga	gaagaaagat	ccaaagtga	agatgtcatt	atgctttaga	4080
tggtatttgg	gtttaagttc	attttgggca	aacaatggaa	ttaaagaaag	atcaatggac	4140
tatcaaattt	ggtgtggtcc	agcgattggg	tcatacaatg	attttgtaaa	aggaacttat	4200
ttggatcctg	cagtagcagg	ttcatatcca	tgtgtgtgtc	aaattaacat	gcaaatctta	4260
cgcgggtgctt	gttttcttca	acgagttcgt	gcaatcaagc	acgatccacg	attggatatt	4320
gatgtcgatg	aagatgtatt	tacctatcgt	ccagaatcaa	ccctatag		4368

<210> 73
<211> 1455
<212> PRT
<213> *Thraustochytrium* sp.

5

<400> 73

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

225						230										235						240
Met	Phe	Ser	Glu	Asn	Asp	Met	Glu	Lys	Leu	Cys	Glu	Arg	Gln	Trp	Glu							
				245					250					255								
Asn	Val	Leu	Gly	Ser	Gly	Leu	Gln	Gly	Ile	Asp	Tyr	Lys	Leu	Cys	Ala							
			260					265					270									
Arg	Lys	Met	Leu	Met	Ile	Asp	Arg	Ile	Thr	Lys	Ile	Gln	His	Asn	Gly							
		275					280					285										
Gly	Ala	Tyr	Gly	Leu	Gly	Leu	Leu	Val	Gly	Glu	Lys	Ile	Leu	Glu	Arg							
	290					295				300												
Asp	His	Trp	Tyr	Phe	Pro	Cys	His	Phe	Val	Lys	Asp	Gln	Val	Met	Ala							
305					310					315					320							
Gly	Ser	Leu	Val	Ser	Asp	Gly	Cys	Ser	Gln	Leu	Leu	Lys	Leu	Tyr	Met							
				325					330					335								
Leu	Trp	Leu	Gly	Leu	His	Asp	Val	Val	Pro	Asp	Phe	Gln	Phe	Arg	Pro							
			340					345					350									
Val	Pro	Gly	Gln	Pro	Asn	Lys	Val	Arg	Cys	Arg	Gly	Gln	Ile	Ser	Pro							
		355					360					365										
His	Arg	Gly	Lys	Leu	Val	Tyr	Val	Met	Glu	Ile	Arg	Glu	Met	Gly	Phe							
	370					375					380											
Asn	Glu	Ser	Thr	Gly	Gln	Pro	Tyr	Ala	Ile	Ala	Asp	Val	Asp	Ile	Ile							
385					390					395				400								
Asp	Val	Asn	Tyr	Glu	Leu	Gly	Gln	Ser	Phe	Asp	Met	Ala	Asp	Ile	Asp							
				405					410					415								
Ser	Tyr	Gly	Arg	Gly	Asn	Leu	Ser	Lys	Lys	Ile	Val	Val	Asp	Phe	Lys							
			420					425					430									
Gly	Ile	Ala	Leu	Gln	Met	Glu	Gly	Thr	Val	Lys	Ser	Ser	Asn	Ile	Ile							
		435					440					445										
Asp	Ser	Ser	Pro	Lys	Ser	Thr	Ile	Ile	Gln	Pro	Pro	Pro	Asn	Cys	Leu							
	450					455				460												
Arg	Gly	Asp	Pro	Leu	Ala	Pro	Ser	Gln	Val	Thr	Trp	His	Pro	Met	Ala							
465					470					475					480							
Gly	Val	Asn	Gly	Ala	Pro	Ala	Pro	Ser	Phe	Ser	Pro	Ser	Asp	Tyr	Pro							
				485					490					495								

Pro Arg Ala Val Cys Phe Lys Pro Phe Pro Gly Asn Pro Leu Asp Asn
500 505 510

Asp His Thr Pro Gly Lys Met Pro Leu Thr Trp Phe Asn Met Ser Glu
515 520 525

Phe Met Cys Gly Lys Val Ser Asn Cys Leu Gly Pro Glu Phe Lys Arg
530 535 540

Phe Asp Asn Ser Lys Thr Ser Arg Ser Pro Ala Phe Asp Leu Ala Leu
545 550 555 560

Val Thr Arg Val Val Ser Val Ser Asp Met Glu Phe Lys Pro His Leu
565 570 575

Asn Ile Asp Val Asn Pro Ser Lys Gly Thr Met Ile Gly Glu Phe Asp
580 585 590

Cys Pro Ala Asp Ala Trp Phe Phe Gln Gly Ser Cys Asn Asp Gly His
595 600 605

Met Pro Tyr Ser Ile Val Met Glu Ile Ala Leu Gln Thr Ser Gly Val
610 615 620

Leu Thr Ser Val Leu Lys Ala Pro Leu Thr Met Asp Lys Asp Asp Ile
625 630 635 640

Leu Phe Arg Asn Leu Asp Ala Thr Ala Glu Met Val Arg Ser Asp Val
645 650 655

Asp Cys Arg Gly Lys Thr Ile Lys Asn Phe Thr Gln Cys Thr Gly Tyr
660 665 670

Ser Met Leu Gly Lys Met Gly Ile His Arg Phe Thr Phe Glu Leu Ser
675 680 685

Val Asp Asp Val Val Phe Tyr Lys Gly Ser Thr Ser Phe Gly Trp Phe
690 695 700

Thr Pro Glu Val Phe Glu Ser Gln Val Gly Leu Asp Asn Gly Lys Lys
705 710 715 720

Val Gln Pro Trp Tyr Leu Glu Gln Lys Ser Ser Asn Val Val Thr Tyr
725 730 735

Asp Val Ala Ser Thr Ala Gly Lys Asp Lys Leu Phe Ser Lys Ile Gly
740 745 750

ES 2 623 610 T3

Ser Lys Asp Ala Gln Val Gln Arg Arg Asn Thr Gln Cys Glu Phe Leu
 755 760 765
 Asp Thr Met His Ile Ile Pro Asn Thr Gly Lys Tyr Asn Lys Gly Tyr
 770 775 780
 Ala His Gly Glu Lys Lys Val Asn Pro Asn Asp Trp Phe Phe Ser Cys
 785 790 795 800
 His Phe Trp Phe Asp Pro Val Met Pro Gly Ser Leu Gly Ile Glu Ser
 805 810 815
 Met Phe Gln Leu Ile Glu Ala Phe Ser Ile Asp Gln Gly Ile Ala Ser
 820 825 830
 Lys His Gly Ile Val Asn Pro Thr Phe Ala His Ser Asn Gly Lys Thr
 835 840 845
 Ser Trp Lys Tyr Arg Gly Gln Leu Asn Asn Lys Gly Lys Arg Met Asp
 850 855 860
 Ser Glu Ile His Ile Lys Asp Ile Val Lys Asn Ala Asp Gly Thr Val
 865 870 875 880
 Asp Leu Ile Ala Asp Gly Phe Leu Leu Val Asp Ser Leu Arg Val Tyr
 885 890 895
 Ser Ala Asp Asp Leu Arg Val Lys Ile Val Pro Gly Thr Lys Ala Ala
 900 905 910
 Pro Lys Ser Val Ala Ala Ala Pro Arg His Val Ala Thr Pro Ile Pro
 915 920 925
 Gly Val Pro Ser Asn Thr Ser Ser Val Glu Ile Ser Leu Glu Ser Leu
 930 935 940
 Lys Lys Glu Leu Leu Asn Leu Glu Lys Pro Leu Tyr Leu Glu Thr Ser
 945 950 955 960
 Asn His Ile Val Lys Gln Phe Gly Asp Val Asn Asn Gly Gln Ala Ser
 965 970 975
 Val Ile Pro Pro Cys Thr Ile Asn Asp Leu Gly Glu Arg Ser Phe Met
 980 985 990
 Glu Thr Tyr Asn Val Val Ala Pro Leu Tyr Thr Gly Ala Met Ala Lys
 995 1000 1005

Gly Ile 1010	Ala Ser Ala Asp	Leu Val Ile Ala Ala Gly 1015 1020	Lys Arg Lys
Ile Leu 1025	Gly Ser Phe Gly Ala 1030	Gly Gly Leu Pro Met 1035	His Leu Val
Arg Ala 1040	Ser Val Glu Lys Ile 1045	Gln Ala Ala Leu Pro 1050	Glu Gly Pro
Tyr Ala 1055	Val Asn Leu Ile His 1060	Ser Pro Phe Asp Ser 1065	Asn Leu Glu
Lys Gly 1070	Asn Val Asp Leu Phe 1075	Leu Glu Lys Gly Val 1080	His Val Val
Glu Ala 1085	Ser Ala Phe Thr Ala 1090	Leu Thr Thr Gln Val 1095	Val Arg Tyr
Arg Ala 1100	Cys Gly Leu Ser Arg 1105	Ala Lys Asp Gly Ser 1110	Val Leu Ile
Lys Asn 1115	Arg Ile Ile Gly Lys 1120	Val Ser Arg Thr Glu 1125	Leu Ala Glu
Met Phe 1130	Phe Arg Pro Ala Pro 1135	Gln Asn Leu Leu Asp 1140	Lys Leu Ile
Ala Ser 1145	Gly Glu Ile Thr Lys 1150	Glu Gln Ala Ser Leu 1155	Ala Leu Glu
Val Pro 1160	Met Ala Asp Asp Val 1165	Ala Val Glu Ala Asp 1170	Ser Gly Gly
His Thr 1175	Asp Asn Arg Pro Ile 1180	His Val Ile Leu Pro 1185	Leu Ile Ile
Asn Leu 1190	Arg Asn Arg Ile His 1195	Lys Glu Cys Gly Phe 1200	Pro Ala Ala
Leu Arg 1205	Val Arg Val Gly Ala 1210	Gly Gly Gly Ile Gly 1215	Cys Pro Ser
Ala Ala 1220	Val Ala Ala Phe Asn 1225	Met Gly Ala Ala Phe 1230	Leu Ile Thr
Gly Ser 1235	Val Asn Gln Val Ser 1240	Lys Gln Ser Gly Thr 1245	Cys Asp Ile
Val Arg	Lys Gln Leu Ser Glu	Ala Ser Tyr Ser Asp	Ile Thr Met

1250		1255		1260
Ala Pro	Ala Ala Asp Met	Phe	Asp Gln Gly Val	Glu Leu Gln Val
1265		1270		1275
Leu Lys	Lys Gly Thr Met	Phe	Pro Ser Arg Ala	Lys Lys Leu Tyr
1280		1285		1290
Glu Leu	Phe Cys Met Tyr	Asn	Ser Phe Asp Asp	Met Pro Lys Ser
1295		1300		1305
Glu Leu	Gln Arg Leu Glu	Lys	Arg Ile Phe Gln	Lys Ser Leu Ala
1310		1315		1320
Glu Val	Trp Glu Glu Thr	Lys	Asp Phe Tyr Ile	Asn Arg Leu Asn
1325		1330		1335
Asn Pro	Glu Lys Ile Glu	His	Ala Glu Lys Lys	Asp Pro Lys Leu
1340		1345		1350
Lys Met	Ser Leu Cys Phe	Arg	Trp Tyr Leu Gly	Leu Ser Ser Phe
1355		1360		1365
Trp Ala	Asn Asn Gly Ile	Lys	Glu Arg Ser Met	Asp Tyr Gln Ile
1370		1375		1380
Trp Cys	Gly Pro Ala Ile	Gly	Ser Tyr Asn Asp	Phe Val Lys Gly
1385		1390		1395
Thr Tyr	Leu Asp Pro Ala	Val	Ala Gly Ser Tyr	Pro Cys Val Val
1400		1405		1410
Gln Ile	Asn Met Gln Ile	Leu	Arg Gly Ala Cys	Phe Leu Gln Arg
1415		1420		1425
Val Arg	Ala Ile Lys His	Asp	Pro Arg Leu Asp	Ile Asp Val Asp
1430		1435		1440
Glu Asp	Val Phe Thr Tyr	Arg	Pro Glu Ser Thr	Leu
1445		1450		1455

<210> 74

<211> 1350

<212> ADN

<213> *Thraustochytrium* sp.

<400> 74

```

agaattgcta ttgttggatt atctgogatt ttaccaagtg gtgaaaatgt tagagaatct      60
tggaagcaa tacgtgatgg tttgaattgt ttaagtgatt tacctgogga tcgtgttgat      120

gttactgCGT attataatcc aacaaaaggt gtaaaggata aaatttattg taaacgtggt      180
gggtttattc ctgaatatga atttgattct agagaatttg gacttaatat gttacaaatg      240
gaagattctg atgctaatac aacgttaact ttattaaagg ttaaagaagc attagatgat      300
gctaataatac ctgcatttac taatgagaaa aaaaatattg gttgtgttct tgggtattggt      360
ggtggtcaaa aagcatctca tgaattttat tcaagactta attatgttgt tgtggataaa      420
gttttaagaa aaatgggatt acctgatgag gatgttgaaa ctgctgttga aaagtttaaa      480
gctaattttc ctgaatggag attagattcc tttcctgggt ttcttggtaa tgttactgct      540
ggcgttgta ctaatacatt caatatggaa ggtatgaatt gtgttgtaga tgctgcttgt      600
gctagttctt taattgctat taaagttgct attgatgaat tattacatgg tgattgtgat      660
gcaatgattg ctggtgcaac ttgtactgat aacgctcttg gtatgtatat ggcattttca      720
aaaacacctg ttttttcaac tgatcaaagt tgtcttgcat atgatgaaaa aacaaaaggt      780
atgcttattg gtgaagggtc agctatgttt gttttaaaac gttatgctga cgcagtgaga      840
gatggtgata ctgtacatgc tgttatacgt tcatgttcat catcatctga cggtaaagca      900
tctggtatth atacaccaac tatttctggt caagaagaag ctattcttag agcatatcgt      960
agagctggtg tatcaccaa tactattact ttagttgaag gacatggtac tggtagacca      1020
gtgggtgata aaattgaatt aacagcttta cgcaatgtat ttgataaagc atatggctct      1080
ggtcataagg aagaagttgc tgttggaagt attaaaagtc aaattggtca tttgaaagct      1140
gttgctggtt gtgctggtct tgtgaaattg gttatggcat tgaaacataa aacactacct      1200
caaagtatta atgttgaaaa tccacctaat ttagtggtat gtactgtcat tagtgatact      1260
actttatata ttaatacaat gaatcgtcca tggattacta agcctggtgt tccaagaaga      1320
gctggtatat ctagtttcgg atttgggtgt      1350

```

<210> 75
 <211> 450
 <212> PRT
 <213> *Thraustochytrium* sp.

 <400> 75

5

ES 2 623 610 T3

Arg Ile Ala Ile Val Gly Leu Ser Ala Ile Leu Pro Ser Gly Glu Asn
1 5 10 15

Val Arg Glu Ser Trp Glu Ala Ile Arg Asp Gly Leu Asn Cys Leu Ser
20 25 30

Asp Leu Pro Ala Asp Arg Val Asp Val Thr Ala Tyr Tyr Asn Pro Thr
35 40 45

Lys Gly Val Lys Asp Lys Ile Tyr Cys Lys Arg Gly Gly Phe Ile Pro
50 55 60

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

Arg Ala Gly Val Ser Pro Asn Thr Ile Thr Leu Val Glu Gly His Gly
325 330 335

Thr Gly Thr Pro Val Gly Asp Lys Ile Glu Leu Thr Ala Leu Arg Asn
340 345 350

Val Phe Asp Lys Ala Tyr Gly Pro Gly His Lys Glu Glu Val Ala Val
355 360 365

Gly Ser Ile Lys Ser Gln Ile Gly His Leu Lys Ala Val Ala Gly Cys
370 375 380

Ala Gly Leu Val Lys Leu Val Met Ala Leu Lys His Lys Thr Leu Pro
385 390 395 400

Gln Ser Ile Asn Val Glu Asn Pro Pro Asn Leu Val Asp Gly Thr Val
405 410 415

Ile Ser Asp Thr Thr Leu Tyr Ile Asn Thr Met Asn Arg Pro Trp Ile
420 425 430

Thr Lys Pro Gly Val Pro Arg Arg Ala Gly Ile Ser Ser Phe Gly Phe
435 440 445

Gly Gly
450

<210> 76
<211> 921
5 <212> ADN
<213> *Thraustochytrium* sp.
<400> 76


```

ttttctggac aaggcgcaca atatacccat atgtttaatg atgttgcaat gcaatggcca      60
caatttcggt tatgtgtaaa tgatatggag aaagcacagg aagaagttat caatgataaa    120
agtgtgaaac gtatcagtc aagttatgttt cctcgtaaac catatgcaag agaatcacct    180
ttagacaata aagaaatctc taagactgaa tattctcaaa caacaactgt cgctagttca    240
gtaggtttat ttgaaatddd ccgtgatgct ggtttcgctc ctgcttttgt tgctgggtcat    300
tcttttaggtg aatttagtgc attgtatgca gctggattga ttgatcgcca agattttatc    360
aagttggtat gtaatcgtgc aatggctatg agagatgcac caaaaaaatc tgctgatgga    420
gcaatggctg ctgttattgg tccaaatgct tcttcaatta agctttcagc tctgaagta    480
tgggttgcta acaataactc tccatctcaa actgttatta ccggtgcaaa ttctgggtga    540
caagctgaaa caagtaaatt gaaaactcaa ggtttccgtg tggttcattt ggcattgtgat    600
ggggcatttc attcgctca tatggaaaat gctgaaaagc aatttcaaaa agctctttca    660
gcagttaagt ttaataaacc aactggttct tctccaaaaa ttttcagcaa tgtaactggt    720

gggtgtattta cggatccaaa aactgctttg tcaagacata tgactagttc tgtacaattt    780
cttactcaaa ttaagaatat gtacgcggct ggagctcgtg tctttattga atttggacca    840
aaacaagtac tttccaaatt ggtcaatgaa atttttcctg gtgatacaag cgttttaact    900
gtttcgggtga atccagctag t                                           921

```

<210> 77

<211> 307

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 77

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

Arg Val Val His Leu Ala Cys Asp Gly Ala Phe His Ser Pro His Met
195 200 205

Glu Asn Ala Glu Lys Gln Phe Gln Lys Ala Leu Ser Ala Val Lys Phe
210 215 220

Asn Lys Pro Thr Gly Ser Ser Pro Lys Ile Phe Ser Asn Val Thr Gly
225 230 235 240

Gly Val Phe Thr Asp Pro Lys Thr Ala Leu Ser Arg His Met Thr Ser
245 250 255

Ser Val Gln Phe Leu Thr Gln Ile Lys Asn Met Tyr Ala Ala Gly Ala
260 265 270

Arg Val Phe Ile Glu Phe Gly Pro Lys Gln Val Leu Ser Lys Leu Val
275 280 285

Asn Glu Ile Phe Pro Gly Asp Thr Ser Val Leu Thr Val Ser Val Asn
290 295 300

Pro Ala Ser
305

<210> 78

<211> 3303

<212> ADN

<213> *Thraustochytrium* sp.

<400> 78

gcaagtcacg ctacogttcg tgtcgtttca gtcctgttc aagcggtgc tcctgtgcag	60
gtatctgctt ctgttgattc tggtttggtg gcaaaagcgg aacaagttgt attggaagta	120
ttggcatcga agactgggta tgagactgag ttgattgaat tggatatgga attggaaact	180
gaacttggtg ttgattctat caagagagta gaaattcttt ctgaagttca agctcaattg	240
aatggtgaag ctaaagatgt agatgctctt agtagaactc gtactgttg tgaagtgatt	300
gatgcaatga aagccgaaat tgctgggtgg caaccagctg ctctgttca agttgcagct	360
cctactcaag tagttgctcc tgttcaagca tctgctcctg ttgattctgg tttgttagca	420
aaagcggaac aagttgtatt ggaagtattg gcatcgaaga ctggttatga gactgagttg	480
attgaattgg atatggaatt ggaaaccgaa cttggtattg attctatcaa gagagtagaa	540
attctttctg aagttcaagc tcaattgagt gttgaagcta aagatgtaga tgctcttagt	600
agaactcgta ctgttggtga agtgattgat gcaatgaaag ccgaaattgc tgggtggtcaa	660
ccagctgctc ctgttcaagt tgcagctcct actcaagtag ttgctcctgt tcaagcatct	720
gctcctgttg attctgggtt gttagcaaaa gcggaacaag ttgtattgga agtattggca	780
tcgaagactg gttatgagac tgagttgatt gaattggata tgggaattgga aaccgaactt	840

ggtattgatt ctatcaagag agtagaaatt ctttctgaag ttcaagctca attgagtgtt 900
 gaagctaaaag atgtagatgc tcttagtaga actcgtactg ttggtgaagt gattgatgca 960
 atgaaagctg aaatttctgg tggtcagcca gctgctcctg ttcaagttgc agctcctact 1020
 caaatagttg ctctgttca agtatccgct cctgttgatt ctgggtttgtt agcaaaggcg 1080
 gaacaagtag tattggaagt attggcatcc aagactgggt atgagactga gttgattgaa 1140
 ttggatatgg aattggaaac tgaacttgggt attgattcta tcaagagagt agaaattctt 1200
 tctgaagttc aagctcaatt gagtgttgaa gctaaagatg tagatgctct tagtagaact 1260
 cgtactgttg gtgaagtgat tgatgcaatg aaagctgaaa tttctggtgg tcaaccaact 1320
 gctcctgttc aagttgcagc tcctactcaa atagttgctc ctgttcaagt atctgctcct 1380
 gttgattctg gtttgttagc aaaggcggaa caagttgtat tggaagtatt ggcacgaa 1440
 actggttatg agactgagtt gattgaattg gatatggaat tggaaaccga acttgggtatt 1500
 gattctatca agagagtaga aattctttct gaagttcaag ctcaattgag tgttgaagct 1560
 aaagatgtag atgctcttag tagaactcgt actgttggtg aagtgattga tgcaatgaaa 1620
 gccgaaattt ctggtggtca gccagctgct cctgttcaag ttgcagctcc tactcaaata 1680
 gttgctcctg ttcaagcatc tgctcctgtt gattctggtt tgttggcaaa agcggaaaca 1740
 gttgtattgg aagtgttagc atccaagact ggttatgaaa ctgagttgat tgaattagat 1800
 atggaattgg aaaccgaact tgggtattgat tctatcaaga gtagtagaat tctttctgaa 1860
 gttcaagctc aattgagtgt tgaagctaaa gatgtagatg ctcttagtag aactcgtact 1920
 gttggtgaag tgattgatgc aatgaaagct gaaatttctg gtggtcaacc agctgctcct 1980
 gttcaagttg cagctcctac tcaaatagtt gctcctgttc aagtatctgc tcctgttgat 2040
 tctggtttgt tagcaaaggc ggaacaagtt gtattggaag tattggcatc taagactggt 2100
 tatgagactg agttgattga attggatatg gaattggaaa ctgaacttggt tattgattct 2160
 atcaagagag tagaaattct ttctgaagtt caagctcaat tgaatgttga agctaaagat 2220
 gtagatgctc ttagtagaac tcgtactgtt ggtgaagtga ttgatgcaat gaaagccgaa 2280
 attgctggtg gtcaaccagc tgctcctgtt caagttgcag ctctgctcc agtagttgct 2340
 cctgttcaag tatctactcc tgttgattct ggtttgttgg caaaagcgga acaagttgta 2400
 ttggaagtgt tagcatgcaa gactggttat gaaactgagt tgattgaatt ggatatggaa 2460
 ttggaaactg aacttggtat tgattctatc aagagagtag aaattcttct tgaagttcaa 2520
 gctcaattga gtgttgaaagc taaagatgta gatgctctta gtagaactcg tactgttggt 2580
 gaagtgattg atgcaatgaa agccgaaatt tctggtggtc aaccaactgc tcctgttcaa 2640
 gttgcagctc ctactcaagt agttgctcct gttaaagtat ctactcctgt tgattctggt 2700
 ttgttagcaa aggcggaaca agtagtattg gaagtattgg catctaagac tgggttatgaa 2760

actgagttga ttgaattaga tatggaattg gaaactgaac ttggtattga ttctatcaag	2820
agagtagaaa ttctttctga agttcaagct caattgaatg tggagctaa agatgtggat	2880
gctcttagta gaactcgtac tgttggtgaa gtgattgatg caatgaaagc cgaaattgct	2940
ggtgatcaac ctgctccagc tgtagttcca gttcaagcta agagtgggtg agccaaccct	3000
gcacttttgg caaaggcgga acaagtagta ttggaagtat tggcatccaa gaccgggtat	3060
gaaactgagc tgattgaatt ggatatggaa ttggaaaactg aacttgggtat tgattcaatc	3120
aagagagtag aaattctgtc cgaagttcaa gcagaattga gtgttgaagc aaaagatgta	3180
gacgctctaa gtagaaccgc tactgttggg gaagtgatcg atgcaatgaa agctgaaatt	3240
gctggcagtg ctgtcacggt tgcaactttg gatgattcaa caattatgga ggagacagat	3300
gat	3303

<210> 79

<211> 1101

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 79

ES 2 623 610 T3

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

ES 2 623 610 T3

145		150		155		160
Ile Glu Leu Asp	Met Glu Leu Glu Thr	Glu Leu Gly Ile Asp Ser Ile				
	165		170		175	
Lys Arg Val Glu	Ile Leu Ser Glu Val	Gln Ala Gln Leu Ser Val Glu				
	180		185		190	
Ala Lys Asp Val Asp	Ala Leu Ser Arg Thr Arg Thr	Val Gly Glu Val				
	195		200		205	
Ile Asp Ala Met Lys	Ala Glu Ile Ala Gly Gly	Gln Pro Ala Ala Pro				
	210		215		220	
Val Gln Val Ala Ala	Pro Thr Gln Val Val	Ala Pro Val Gln Ala Ser				
	225		230		235	
Ala Pro Val Asp Ser	Gly Leu Leu Ala Lys	Ala Glu Gln Val Val Leu				
	245		250		255	
Glu Val Leu Ala Ser	Lys Thr Gly Tyr Glu Thr	Glu Leu Ile Glu Leu				
	260		265		270	
Asp Met Glu Leu Glu	Thr Glu Leu Gly Ile Asp Ser	Ile Lys Arg Val				
	275		280		285	
Glu Ile Leu Ser Glu	Val Gln Ala Gln Leu Ser	Val Glu Ala Lys Asp				
	290		295		300	
Val Asp Ala Leu Ser	Arg Thr Arg Thr Val	Gly Glu Val Ile Asp Ala				
	305		310		315	
Met Lys Ala Glu Ile	Ser Gly Gly Gln Pro	Ala Ala Pro Val Gln Val				
	325		330		335	
Ala Ala Pro Thr	Gln Ile Val Ala Pro	Val Gln Val Ser Ala Pro Val				
	340		345		350	
Asp Ser Gly Leu Leu	Ala Lys Ala Glu Gln Val	Val Leu Glu Val Leu				
	355		360		365	
Ala Ser Lys Thr Gly	Tyr Glu Thr Glu Leu Ile	Glu Leu Asp Met Glu				
	370		375		380	
Leu Glu Thr Glu Leu	Gly Ile Asp Ser Ile	Lys Arg Val Glu Ile Leu				
	385		390		395	
Ser Glu Val Gln Ala	Gln Leu Ser Val Glu	Ala Lys Asp Val Asp Ala				
	405		410		415	

Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys Ala
 420 425 430

Glu Ile Ser Gly Gly Gln Pro Thr Ala Pro Val Gln Val Ala Ala Pro
 435 440 445

Thr Gln Ile Val Ala Pro Val Gln Val Ser Ala Pro Val Asp Ser Gly
 450 455 460

Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val Leu Ala Ser Lys
 465 470 475 480

Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met Glu Leu Glu Thr
 485 490 495

Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val
 500 505 510

Gln Ala Gln Leu Ser Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg
 515 520 525

Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys Ala Glu Ile Ser
 530 535 540

Gly Gly Gln Pro Ala Ala Pro Val Gln Val Ala Ala Pro Thr Gln Ile
 545 550 555 560

Val Ala Pro Val Gln Ala Ser Ala Pro Val Asp Ser Gly Leu Leu Ala
 565 570 575

Lys Ala Glu Gln Val Val Leu Glu Val Leu Ala Ser Lys Thr Gly Tyr
 580 585 590

Glu Thr Glu Leu Ile Glu Leu Asp Met Glu Leu Glu Thr Glu Leu Gly
 595 600 605

Ile Asp Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln Ala Gln
 610 615 620

Leu Ser Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr Arg Thr
 625 630 635 640

Val Gly Glu Val Ile Asp Ala Met Lys Ala Glu Ile Ser Gly Gly Gln
 645 650 655

Pro Ala Ala Pro Val Gln Val Ala Ala Pro Thr Gln Ile Val Ala Pro
 660 665 670

ES 2 623 610 T3

Val	Gln	Val	Ser	Ala	Pro	Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	675	680	685	
Gln	Val	Val	Leu	Glu	Val	Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	690	695	700	
Leu	Ile	Glu	Leu	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	705	710	715	720
Ile	Lys	Arg	Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu	Asn	Val	725	730	735	
Glu	Ala	Lys	Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	740	745	750	
Val	Ile	Asp	Ala	Met	Lys	Ala	Glu	Ile	Ala	Gly	Gly	Gln	Pro	Ala	Ala	755	760	765	
Pro	Val	Gln	Val	Ala	Ala	Pro	Ala	Pro	Val	Val	Ala	Pro	Val	Gln	Val	770	775	780	
Ser	Thr	Pro	Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	785	790	795	800
Leu	Glu	Val	Leu	Ala	Cys	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	805	810	815	
Leu	Asp	Met	Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	820	825	830	
Val	Glu	Ile	Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu	Ser	Val	Glu	Ala	Lys	835	840	845	
Asp	Val	Asp	Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Ile	Asp	850	855	860	
Ala	Met	Lys	Ala	Glu	Ile	Ser	Gly	Gly	Gln	Pro	Thr	Ala	Pro	Val	Gln	865	870	875	880
Val	Ala	Ala	Pro	Thr	Gln	Val	Val	Ala	Pro	Val	Lys	Val	Ser	Thr	Pro	885	890	895	
Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val	900	905	910	
Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met	915	920	925	

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
930 935 940

Leu Ser Glu Val Gln Ala Gln Leu Asn Val Glu Ala Lys Asp Val Asp
945 950 955 960

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
965 970 975

Ala Glu Ile Ala Gly Asp Gln Pro Ala Pro Ala Val Val Pro Val Gln
980 985 990

Ala Lys Ser Gly Val Ala Asn Pro Ala Leu Leu Ala Lys Ala Glu Gln
995 1000 1005

Val Val Leu Glu Val Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu
1010 1015 1020

Leu Ile Glu Leu Asp Met Glu Leu Glu Thr Glu Leu Gly Ile Asp
1025 1030 1035

Ser Ile Lys Arg Val Glu Ile Leu Ser Glu Val Gln Ala Glu Leu
1040 1045 1050

Ser Val Glu Ala Lys Asp Val Asp Ala Leu Ser Arg Thr Arg Thr
1055 1060 1065

Val Gly Glu Val Ile Asp Ala Met Lys Ala Glu Ile Ala Gly Ser
1070 1075 1080

Ala Val Thr Val Ala Thr Leu Asp Asp Ser Thr Ile Met Glu Glu
1085 1090 1095

Thr Asp Asp
1100

<210> 80

<211> 255

<212> ADN

<213> *Thraustochytrium* sp.

<400> 80

```

gttgattctg gtttgttggc aaaagcggaa caagttgtat tggaaagtatt ggcacggaag      60
actggttatg agactgagtt gattgaattg gatatggaat tggaaactga acttggtatt      120
gattctatca agagagtaga aattctttct gaagttcaag ctcaattgaa tgttgaagct      180
aaagatgtag atgctcttag tagaactcgt actgttggtg aagtgattga tgcaatgaaa      240
gccgaaattg ctggt                                     255

```

<210> 81

<211> 85

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 81

```

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1          5          10          15

```

```

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
          20          25          30

```

```

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35          40          45

```

```

Leu Ser Glu Val Gln Ala Gln Leu Asn Val Glu Ala Lys Asp Val Asp
50          55          60

```

```

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65          70          75          80

```

```

Ala Glu Ile Ala Gly
          85

```

10

<210> 82

<211> 255

<212> ADN

15 <213> *Thraustochytrium* sp.

<400> 82

```

gttgattctg gtttgttagc aaaagcggaa caagttgtat tggaaagtatt ggcacggaag      60
actggttatg agactgagtt gattgaattg gatatggaat tggaaaccga acttggtatt      120
gattctatca agagagtaga aattctttct gaagttcaag ctcaattgag tgttgaagct      180
aaagatgtag atgctcttag tagaactcgt actgttggtg aagtgattga tgcaatgaaa      240
gccgaaattg ctggt                                     255

```

20

<210> 83

<211> 85

<212> PRT

<213> *Thraustochytrium* sp.

<400> 83

Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val
1				5					10					15	
Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met
			20					25						30	
Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile
		35					40					45			
Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu	Ser	Val	Glu	Ala	Lys	Asp	Val	Asp
		50				55						60			
Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Ile	Asp	Ala	Met	Lys
65					70					75					80
Ala	Glu	Ile	Ala	Gly											
				85											

5

<210> 84

<211> 252

<212> ADN

10 <213> *Thraustochytrium* sp.

<400> 84

gttgattctg	gtttgtttagc	aaaagcggaa	caagttgtat	tggaagtatt	ggcatogaag	60
actgggttatg	agactgagtt	gattgaattg	gatatggaat	tggaaacoga	acttggtatt	120
gattctatca	agagagtaga	aattctttct	gaagttcaag	ctcaattgag	tgttgaagct	180
aaagatgtag	atgctcttag	tagaactcgt	actgttggtg	aagtgattga	tgcaatgaaa	240
gctgaaattt	ct					252

15

<210> 85

<211> 85

<212> PRT

20 <213> *Thraustochytrium* sp.

<400> 85

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Ser Val Glu Ala Lys Asp Val Asp
50 55 60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ser Gly
85

<210> 86

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 86

gttgattctg gtttgtttagc aaaggcggaa caagtagtat tggaagtatt ggcattccaag 60

actggttatg agactgagtt gattgaattg gatattggaat tggaaactga acttggtatt 120

gattctatca agagagtaga aattctttct gaagttcaag ctcaattgag tgttgaagct 180

aaagatgtag atgctcttag tagaactcgt actggttggtg aagtgattga tgcaatgaaa 240

gctgaaattt ctggt 255

10

<210> 87

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 87

Val	Asp	Ser	Gly	Leu	Leu	Ala	Lys	Ala	Glu	Gln	Val	Val	Leu	Glu	Val
1				5					10					15	
Leu	Ala	Ser	Lys	Thr	Gly	Tyr	Glu	Thr	Glu	Leu	Ile	Glu	Leu	Asp	Met
			20					25					30		
Glu	Leu	Glu	Thr	Glu	Leu	Gly	Ile	Asp	Ser	Ile	Lys	Arg	Val	Glu	Ile
		35					40					45			
Leu	Ser	Glu	Val	Gln	Ala	Gln	Leu	Ser	Val	Glu	Ala	Lys	Asp	Val	Asp
	50					55					60				
Ala	Leu	Ser	Arg	Thr	Arg	Thr	Val	Gly	Glu	Val	Ile	Asp	Ala	Met	Lys
65					70					75				80	
Ala	Glu	Ile	Ser	Gly											
				85											

<210> 88
 <211> 254
 5 <212> ADN
 <213> *Thraustochytrium* sp.

<400> 88

ttgattctgg	tttgtttagca	aaggcggaac	aagttgtatt	ggaagtattg	gcatcgaaga	60
ctggttatga	gactgagttg	attgaattgg	atatggaatt	ggaaaccgaa	cttggtattg	120
attctatcaa	gagagtagaa	attctttctg	aagttcaagc	tcaattgagt	gttgaagcta	180
aagatgtaga	tgctcttagt	agaactcgta	ctggttggtga	agtgattgat	gcaatgaaag	240
ccgaaatttc	tggt					254

<210> 89
 <211> 85
 <212> PRT
 15 <213> *Thraustochytrium* sp.

<400> 89

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Ser Val Glu Ala Lys Asp Val Asp
50 55 60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ser Gly
85

<210> 90

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 90

gttgattctg gtttgttggc aaaagcggaa caagttgtat tggaagtgtt agcatccaag 60

actggttatg aaactgagtt gattgaatta gatattggaat tggaaaccga acttgggtatt 120

gattctatca agagagtaga aattctttct gaagttcaag ctcaattgag tggtgaagct 180

aaagatgtag atgctcttag tagaactcgt actgttgggtg aagtgattga tgcaatgaaa 240

gctgaaattt ctggt 255

10

<210> 91

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 91

ES 2 623 610 T3

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Ser Val Glu Ala Lys Asp Val Asp

50

55

60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ser Gly
85

<210> 92

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 92

gttgattctg gtttgtagc aaaggcggaa caagttgtat tggaagtatt ggcattctaag 60

actggttatg agactgagtt gattgaattg gatatggaat tggaaactga acttggtatt 120

gattctatca agagagtaga aattctttct gaagttcaag ctcaattgaa tgttgaagct 180

aaagatgtag atgctcttag tagaactcgt actggtggtg aagtgattga tgcaatgaaa 240

gccgaaattg ctggt 255

10

<210> 93

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 93

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Asn Val Glu Ala Lys Asp Val Asp
50 55 60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ala Gly
85

<210> 94

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 94

gttgattctg gtttgttggc aaaagcggaa caagttgtat tggaaagtgtt agcatgcaag 60

actggttatg aaactgagtt gattgaattg gatattggaat tggaaactga acttggtatt 120

gattctatca agagagtaga aattctttct gaagttcaag ctcaattgag tgttgaagct 180

aaagatgtag atgctcttag tagaactcgt actgttggtg aagtgattga tgcaatgaaa 240

gccgaaattt ctggt 255

10

<210> 95

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 95

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Cys Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Ser Val Glu Ala Lys Asp Val Asp
50 55 60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ser Gly
85

<210> 96

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 96

gttgattctg gtttgtttagc aaaggcggaa caagtagtat tggaaagtatt ggcattctaag 60

actggttatg aaactgagtt gattgaatta gatattggaat tggaaactga acttggtatt 120

gattctatca agagagtaga aattctttct gaagttcaag ctcaattgaa tgtggaagct 180

aaagatgtgg atgctcttag tagaactcgt actggttggtg aagtgattga tgcaatgaaa 240

10 gccgaaattg ctggt 255

<210> 97

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 97

ES 2 623 610 T3

Val Asp Ser Gly Leu Leu Ala Lys Ala Glu Gln Val Val Leu Glu Val
1 5 10 15

Leu Ala Ser Lys Thr Gly Tyr Glu Thr Glu Leu Ile Glu Leu Asp Met
20 25 30

Glu Leu Glu Thr Glu Leu Gly Ile Asp Ser Ile Lys Arg Val Glu Ile
35 40 45

Leu Ser Glu Val Gln Ala Gln Leu Asn Val Glu Ala Lys Asp Val Asp
50 55 60

Ala Leu Ser Arg Thr Arg Thr Val Gly Glu Val Ile Asp Ala Met Lys
65 70 75 80

Ala Glu Ile Ala Gly
85

<210> 98

<211> 255

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 98

gccaacctg cacttttggc aaaggcggaa caagtagtat tggaagtatt ggcacccaag 60

accggttatg aaactgagct gattgaattg gatattggaat tggaaactga acttggtatt 120

gattcaatca agagagtaga aattctgtcc gaagttcaag cagaattgag tggtgaagca 180

aaagatgtag acgctctaag tagaaccogt actgttgggg aagtgatcga tgcaatgaaa 240

10 gctgaaattg ctggc 255

<210> 99

<211> 85

<212> PRT

15 <213> *Thraustochytrium* sp.

<400> 99

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

<210> 100
 <211> 2151
 <212> ADN
 <213> *Thraustochytrium* sp.
 <400> 100

5

gtttcaagca ttgaatcttc gtatggaaaa attggtggct ttgtttatca acattttcat	60
gatagcgact atggtatgca acttggatgg gcgttaatgg cagcgaaaca tttgaaagag	120
tccctcaacg acccgattaa gaatggaaga accttctttt tggctgttgc gcgtatgaat	180
ggtaaaacttg gtatggacaa tgcttcagtt catgatcaag gaatagtgga atcatgcggt	240
atcgccgaac gtggtgctat ctttggtttg tgcaaaactt tggatttgga atggccta	300
gtttttgctc gtggtgttga tattgctgaa ggtatgagtt atagtttggc tgcggaattg	360
attgttgatg agatttcttg tgcaaactct tccattcggg aatctggtta cacgattagc	420
ggagaaagat tcacaactga agctcacaaa ttggttactg gaaagcctca tgctccgatt	480
aagaagaagg atgctttcct agtatctggt ggtgctcgtg gtattactcc actttgtatt	540
cgtgaaattg ctaaagcagt gaaagggtggc acttacattt tgatgggtcg atcagctttg	600
gctgatgaac ccttgtgggc taatggtaaa tccggaaaag atttagataa agctggtttg	660
gcatttttga aggaagagtt tgcagctggg cgtggtagta aaccaactcc aaaagtccac	720
aaatctttga ttgataaagt gctcgggtatt agggagggtta gagcatctat tgcaaata	780
gaagcccatg gagcaaaagc tatatatattg tcttgcatg tatcttccgc tgagaaagta	840
aaggctgcag tgcaaaaagt tgaaaaggag catctagttc gtattactgg tattgtgcat	900
gcatcaggcg ttttgaggga taaattggtt gagaacaaaa ctttggatga tttcaacgca	960
gtatatggaa ccaaagtaac tggactagta aacttgctgt cagcagtga catgaatttt	1020
gttcgtcatt tggttatggt tagttctttg gctggatatc atggaaatgt tggccaatct	1080
gattatgcaa tggctaacga atcacttaac aagattggtt ttagattggg tgcagcttat	1140
tctcaattgt gtgttaaata tatttgtttt ggaccttggg atggtggaat ggtaactcca	1200
gctttgaaaa aacaatttca atcaatgggt gtccagatta ttctcgtga aggtggcgcg	1260
gagactgttg caagaatagt cttatcttca aatccttctc aagttttagt tggcaactgg	1320
ggtgttcctc cagtttcacc tttgtcaaaa tcggcaacta ttgttcaaac ttttaccct	1380
gagttaaatc catcttctaa gtctcatcaa attcatggta aaaatgtttt gcctatgact	1440
gtagcaattg gatattctgc tcaattgggt aagaattttt atgctggtca tcatgttg	1500
ggagttgaag atgctcaatt gttcagtggt gttgtaattg accatgcggt gcaagctcaa	1560

gtgaaattaa cggaacagag tttggatgat gatggcaagg taaaagttca agctgttctg 1620
 actgottcaa acgataatgg aaaaatggta cctgcataca aagcagtgat tgttttggga 1680
 aaaacaagta gacctgcgtt tattttgaaa gatttttcat tgcaagaatc taattctcgc 1740
 agtgcgtgat agttgtatga tggtaaaaact ttgtttcatg gtccattatt tegtgggaatt 1800
 accaagttgt tgaatgtatc tgataacttca ctaacaactc aatgtaccaa tattgatttg 1860
 actgctactg aacgtgggtca atttgcggat atcgaacctg tgaatccttt tatggcggat 1920
 gctgcatttc aagctatgct tgtatgggtt agaaatttaa ggaatagtgc atctttacca 1980
 aacaattgtg aaagagtaga tatctataaa ccaatagcac ctggtgaaaa gtattacact 2040
 actttgcaag ctttgggtaa tacctccggt tctgttctca agtctgtatt ttatatgcac 2100
 gatgaacaag gagaagtatt tctatctgga agagctagtg ttgttgtgaa t 2151

<210> 101

<211> 717

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 101

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

Thr Thr Glu Ala His Lys Leu Val Thr Gly Lys Pro His Ala Pro Ile
 145 150 155 160

 Lys Lys Lys Asp Ala Phe Leu Val Ser Gly Gly Ala Arg Gly Ile Thr
 165 170 175

 Pro Leu Cys Ile Arg Glu Ile Ala Lys Ala Val Lys Gly Gly Thr Tyr
 180 185 190

 Ile Leu Met Gly Arg Ser Ala Leu Ala Asp Glu Pro Leu Trp Ala Asn
 195 200 205

 Gly Lys Ser Gly Lys Asp Leu Asp Lys Ala Gly Leu Ala Phe Leu Lys
 210 215 220

 Glu Glu Phe Ala Ala Gly Arg Gly Ser Lys Pro Thr Pro Lys Val His
 225 230 235 240

 Lys Ser Leu Ile Asp Lys Val Leu Gly Ile Arg Glu Val Arg Ala Ser
 245 250 255

 Ile Ala Asn Ile Glu Ala His Gly Ala Lys Ala Ile Tyr Leu Ser Cys
 260 265 270

 Asp Val Ser Ser Ala Glu Lys Val Lys Ala Ala Val Gln Lys Val Glu
 275 280 285

 Lys Glu His Leu Val Arg Ile Thr Gly Ile Val His Ala Ser Gly Val
 290 295 300

 Leu Arg Asp Lys Leu Val Glu Asn Lys Thr Leu Asp Asp Phe Asn Ala
 305 310 315 320

 Val Tyr Gly Thr Lys Val Thr Gly Leu Val Asn Leu Leu Ser Ala Val
 325 330 335

 Asn Met Asn Phe Val Arg His Leu Val Met Phe Ser Ser Leu Ala Gly
 340 345 350

 Tyr His Gly Asn Val Gly Gln Ser Asp Tyr Ala Met Ala Asn Glu Ser
 355 360 365

 Leu Asn Lys Ile Gly Phe Arg Leu Gly Ala Ala Tyr Ser Gln Leu Cys
 370 375 380

 Val Lys Ser Ile Cys Phe Gly Pro Trp Asp Gly Gly Met Val Thr Pro
 385 390 395 400

 Ala Leu Lys Lys Gln Phe Gln Ser Met Gly Val Gln Ile Ile Pro Arg

405										410					415						
Glu	Gly	Gly	Ala	Glu	Thr	Val	Ala	Arg	Ile	Val	Leu	Ser	Ser	Asn	Pro						
			420					425					430								
Ser	Gln	Val	Leu	Val	Gly	Asn	Trp	Gly	Val	Pro	Pro	Val	Ser	Pro	Leu						
		435					440					445									
Ser	Lys	Ser	Ala	Thr	Ile	Val	Gln	Thr	Phe	Thr	Pro	Glu	Leu	Asn	Pro						
	450					455					460										
Phe	Leu	Lys	Ser	His	Gln	Ile	His	Gly	Lys	Asn	Val	Leu	Pro	Met	Thr						
465					470					475					480						
Val	Ala	Ile	Gly	Tyr	Leu	Ala	His	Leu	Val	Lys	Asn	Phe	Tyr	Ala	Gly						
			485						490					495							
His	His	Leu	Trp	Gly	Val	Glu	Asp	Ala	Gln	Leu	Phe	Ser	Gly	Val	Val						
		500						505					510								
Ile	Asp	His	Ala	Val	Gln	Ala	Gln	Val	Lys	Leu	Thr	Glu	Gln	Ser	Leu						
	515					520						525									
Asp	Asp	Asp	Gly	Lys	Val	Lys	Val	Gln	Ala	Val	Leu	Thr	Ala	Ser	Asn						
	530					535					540										
Asp	Asn	Gly	Lys	Met	Val	Pro	Ala	Tyr	Lys	Ala	Val	Ile	Val	Leu	Gly						
545					550					555					560						
Lys	Thr	Ser	Arg	Pro	Ala	Phe	Ile	Leu	Lys	Asp	Phe	Ser	Leu	Gln	Glu						
			565						570					575							
Ser	Asn	Ser	Arg	Ser	Ala	Asp	Glu	Leu	Tyr	Asp	Gly	Lys	Thr	Leu	Phe						
			580					585					590								
His	Gly	Pro	Leu	Phe	Arg	Gly	Ile	Thr	Lys	Leu	Leu	Asn	Val	Ser	Asp						
	595					600						605									
Thr	Ser	Leu	Thr	Thr	Gln	Cys	Thr	Asn	Ile	Asp	Leu	Thr	Ala	Thr	Glu						
	610				615						620										
Arg	Gly	Gln	Phe	Ala	Asp	Ile	Glu	Pro	Val	Asn	Pro	Phe	Met	Ala	Asp						
625					630					635					640						
Ala	Ala	Phe	Gln	Ala	Met	Leu	Val	Trp	Val	Arg	Asn	Leu	Arg	Asn	Ser						
			645						650					655							
Ala	Ser	Leu	Pro	Asn	Asn	Cys	Glu	Arg	Val	Asp	Ile	Tyr	Lys	Pro	Ile						
		660						665					670								

ES 2 623 610 T3

Ala Pro Gly Glu Lys Tyr Tyr Thr Thr Leu Gln Ala Leu Gly Asn Thr
675 680 685

Ser Gly Ser Val Leu Lys Ser Val Phe Tyr Met His Asp Glu Gln Gly
690 695 700

Glu Val Phe Leu Ser Gly Arg Ala Ser Val Val Val Asn
705 710 715

<210> 102

<211> 1311

<212> ADN

<213> *Thraustochytrium* sp.

<400> 102

```

ttaagtgttg gtgataatat ttgtcatgat caacgtgttg ctggtgttgg tatggctggt      60
atgtatgctg gttgtcaaaa tcaacatgaa ttttggcaat ctttacaagg taaaaatatg    120
aattcaaaat cgatttcaca aaatcgttta gggtctgagt atagagaaga acattttaaa    180
cctgaaagaa gtaaataattc cgataccttt tgtaatgaaa gatatggttg tattgatgag    240
aatgttcaaa gtgaacatga acttttatta aaacttgcaa aagatgctat tgcggataca    300
aaaggttcta ttgatttgaa taaaaccgga atcgttagtg gttgcttata ttttccaatg    360
gataatttac aaggtgattt attaaatttg tatcaatgtc acattgaaaa gaaaattggg    420
ccaaatgcat taaaagatgt gaatttatgg tctaaaagaa ccaccaacgg aaaagatgat    480
aaaaaagctt attttgatcc tgcctctttc gtagctgaac aattagatat gggaccatta    540
cattatagtt tagatgctgc ttgtgcgtct gcactttatg tattaagact tgctcaagat    600
catttattaa gtggtgctgc tgatacaatg ttatgtggtg catcttgttt acctgaacct    660
ttttttattt tatctggttt ttctactttt catgcaatgc cattatctgg tgatgtttct    720
gtctctttgc ataaaacttc acaaggtctt acacctggtg aaggtggtgc tattatggta    780
cttaaacgat taaatgatgc aatccgtgat ggtgatagaa tttatggtac tttacttggg    840
gctgaattaa gtaatgctgg ttgtggttta ccattgagtc cacatatgcc aagtgaattt    900
gattgtatgg aaaaagcttt acaaagagta cacagattac catcatctat tcaatatgtt    960
gagtgatcat caactggtac accacaaggt gataaagttg aaattgatgc tatgacaaaa   1020
tgttttggtg aacatttacc aaggtttggt tcaacgaaag ggaattttgg tcatacactt   1080
gttgctgctg gttttgctgg tatgtgtaaa gttttattat caatgcaata tggtgaaata   1140
ccaccaactc caggtcttga aaatccagac aatattatgc atgatttagt tgttactgaa   1200
acaattccat ggcctaatac aaatggtgat ttgaaacgtg catgtttata tgcttttgga   1260
ttcggtggtg ctaatgcaca tgctgtattt gaagagtatc gttcagattt a             1311

```

<210> 103

<211> 437

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 103

ES 2 623 610 T3

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

Ala Pro Leu His Lys Thr Ser Gln Gly Leu Thr Pro Gly Glu Gly Gly
245 250 255

Ala Ile Met Val Leu Lys Arg Leu Asn Asp Ala Ile Arg Asp Gly Asp
260 265 270

Arg Ile Tyr Gly Thr Leu Leu Gly Ala Glu Leu Ser Asn Ala Gly Cys
275 280 285

Gly Leu Pro Leu Ser Pro His Met Pro Ser Glu Phe Asp Cys Met Glu
290 295 300

Lys Ala Leu Gln Arg Val His Arg Leu Pro Ser Ser Ile Gln Tyr Val
305 310 315 320

Glu Cys His Ala Thr Gly Thr Pro Gln Gly Asp Lys Val Glu Ile Asp
325 330 335

Ala Met Thr Lys Cys Phe Gly Glu His Leu Pro Arg Phe Gly Ser Thr
340 345 350

Lys Gly Asn Phe Gly His Thr Leu Val Ala Ala Gly Phe Ala Gly Met
355 360 365

Cys Lys Val Leu Leu Ser Met Gln Tyr Gly Glu Ile Pro Pro Thr Pro
370 375 380

Gly Leu Glu Asn Pro Asp Asn Ile Met His Asp Leu Val Val Thr Glu
385 390 395 400

Thr Ile Pro Trp Pro Asn Thr Asn Gly Asp Leu Lys Arg Ala Cys Leu
405 410 415

Ser Ala Phe Gly Phe Gly Gly Thr Asn Ala His Ala Val Phe Glu Glu
420 425 430

Tyr Arg Ser Asp Leu
435

<210> 104

<211> 1323

<212> ADN

<213> *Thraustochytrium* sp.

<400> 104

aaaattgcta ttgttggtat ggaatctgaa tttggtactt tgaaaggatt acaagaattt	60
gaacgtgcta tttacaatgg tggatcatgg gcatgtgatt tacctgaaaa tagatggaga	120
tttcttggag aagataaaga attttttacaa gcttgtgggt taaaaaatt accaagaggt	180
tggtatatta aagaagtga aactgatttt aaaaggttac gtttaccaat gatacaggag	240
gatattctaa gacctttaca gttgttagct gtttcgatta tcgacagagc acttaacgca	300
tctggtgcta aaccaaattg caaagttgca gttttagttg gattaggtac tgatcttgaa	360
ttatatcgtc atcgtgctcg tgttgcatga aaggaacgcc tccaaactgc ggtcaaagaa	420
gatattcctt tacttgaaaa gttaatgaac tatgtcaatg atagaggtac aagtacatca	480
tatacatctt atattggaaa tttggttgca actcgagttt catcattatg gggttttact	540
ggtccatcat tcacgattac tgaaggtgaa aattccgtat atcgttgtct tgatttggga	600
agatggttct tagctaattg tgaagtagat gctgttggtg ttgccggggg tgatttatgt	660
ggtagtgctg aaaatctttt tgtaaaatct cgtagaagta aagtttccac acaaatgaa	720
ccatttgcaa attttgaatc aaatgctgat ggatattttg ctggagatgg ttgtggagct	780
ttggttttga aacgattgag tgattgtacg gattcaactg aaaaaattta tgcaacgggtg	840
gattcaattg ctgttggtga tgaagttggc ccaactatta aacaagcttt gaagaatgca	900
tccatagcag cgaaagatat tgaactggca gagctatcag caagttcagg caaacatcat	960
tctggtagaa tcacttgtga agatgaacta aatgaactgg gtgaaatttt caatgaaggt	1020
atacaaagag ttgcaattgg tagtgtgaaa gctaattgtg gagatgttgg atatgcatct	1080
ggtgcagcaa gtttaaatcaa aacggctttg tgctgtgaca accgatattt accaaagtta	1140
ccaaattgga ataagccaac gaaagatgtt gaatgggtcca aatcattttt tgtatgtgaa	1200
cattctagag catggttgaa aaatgttgat gaaaatagac atgctgtcgt ttctggagtt	1260
tgcgaaaatg gttcgtgtta tggaatcgta atgtctgatg tacaaggaca tcatgaagaa	1320
tcg	1323

<210> 105

<211> 441

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 105

ES 2 623 610 T3

Lys Ile Ala Ile Val Gly Met Glu Ser Glu Phe Gly Thr Leu Lys Gly
1 5 10 15

Leu Gln Glu Phe Glu Arg Ala Ile Tyr Asn Gly Gly His Gly Ala Cys
20 25 30

Asp Leu Pro Glu Asn Arg Trp Arg Phe Leu Gly Glu Asp Lys Glu Phe
35 40 45

Leu Gln Ala Cys Gly Leu Gln Lys Leu Pro Arg Gly Cys Tyr Ile Lys
50 55 60

ES 2 623 610 T3

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

325								330				335					
Phe	Asn	Glu	Gly	Ile	Gln	Arg	Val	Ala	Ile	Gly	Ser	Val	Lys	Ala	Asn		
340								345				350					
Val	Gly	Asp	Val	Gly	Tyr	Ala	Ser	Gly	Ala	Ala	Ser	Leu	Ile	Lys	Thr		
355								360				365					
Ala	Leu	Cys	Leu	Tyr	Asn	Arg	Tyr	Leu	Pro	Lys	Leu	Pro	Asn	Trp	Asn		
370								375				380					
Lys	Pro	Thr	Lys	Asp	Val	Glu	Trp	Ser	Lys	Ser	Phe	Phe	Val	Cys	Glu		
385								390				395				400	
His	Ser	Arg	Ala	Trp	Leu	Lys	Asn	Val	Asp	Glu	Asn	Arg	His	Ala	Val		
405								410				415					
Val	Ser	Gly	Val	Cys	Glu	Asn	Gly	Ser	Cys	Tyr	Gly	Ile	Val	Met	Ser		
420								425				430					
Asp	Val	Gln	Gly	His	His	Glu	Glu	Ser									
435								440									

<210> 106
<211> 1353
<212> ADN
<213> *Thraustochytrium* sp.

<400> 106

```

cctacaatag atattgactc caatgtgttt gctgagatgc ttaatctacc gcaggataaa      60
aacaaaaaat ttgcggctgc attggttacc acaccaaata aactccagcg tgaaatagaa      120
cttgctgtga aggggtattcc acgttgcgta aaagcaaaaa gagattgggtg ttctccatct      180
ggaagtatth ttgcttgtaa tccactcaaa agtgataata ttgcatttat gtatgggtgaa      240
ggccgaagcc catatgctgg actgggatat gatttgcacg gaatttggcc tatgctacac      300
gagttgggta acaatagaac tacagaaactt tgggatcaag gtgatagttg gtatttacct      360
cgatctagct ctgttgctga aaaagaaaaa gtcttcggag attttgataa gaatcaaatt      420
gaaatgttta gattgggtat ttttgtatca atgtgtttca ctgatatggc cactgaactt      480
ttgggtttta aacccaaagc cgcgtttggg ttaagtttgg gtgaaatac tatgcttttt      540
gcattttcta aaaagaatac caagttgtcc aaagaattga cccgtcgtct aaaagaagca      600
aaagtttggg catcacaatt agctgttgaa tttgcagcta ttcgagattt gtggaatatt      660
ccagctgata aatctattga tgaatttttg caagggattt ttgtttacgc aaatcgaacc      720
ctggctcgaga acacaattgg ggagaataaa tttgttcggt tgttgattgt aaatgattcg      780
caaagttgtc taattgccgg gaaaccagat gaatgtcaaa aagttattga gaagcttcat      840

ttgaagctac cggcgggttc agtaactcag ggtatgatcg gtcattgccc agaagcaatt      900
ccttatctag atcaaactcag tcatattcat gaaatgcttg aaattccaaa acccgaaaat      960
gtgaaattgt ttacaactag tgaaaacaga gaattagtgt cgatgaaaga ttccgtgtca     1020
aaattgggtg ctgagattta tcagcatggt gctgattttc caaacatcgt gaacaagggt     1080
aaagaaactt gcaaaactga tatatttatt gaattgggat cgaacaatta tcgatctgga     1140
gctgtcaaaa caattttagg tccagaaatc gtttctgttg caattgatag gcaaaatgaa     1200
actgcatggg gtcaactaat gaagatgggt gcatcggtga taagtcacg agttccgggt     1260
gttgaattga aaaaacteta tcatcctgaa ttgctgaaat ttgatccaca ggcaaaaccg     1320
aatcgtttca tcagaaatat agaactgaat gga                                     1353

```

<210> 107

<211> 451

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 107

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

Leu Gly Leu Lys Pro Lys Ala Ala Phe Gly Leu Ser Leu Gly Glu Ile
 165 170 175

Ser Met Leu Phe Ala Phe Ser Lys Lys Asn Thr Lys Leu Ser Lys Glu
 180 185 190

Leu Thr Arg Arg Leu Lys Glu Ala Lys Val Trp Ala Ser Gln Leu Ala
 195 200 205

Val Glu Phe Ala Ala Ile Arg Asp Leu Trp Asn Ile Pro Ala Asp Lys
 210 215 220

Ser Ile Asp Glu Phe Trp Gln Gly Tyr Phe Val Tyr Ala Asn Arg Thr
 225 230 235 240

Leu Val Glu Asn Thr Ile Gly Glu Asn Lys Phe Val Arg Leu Leu Ile
 245 250 255

Val Asn Asp Ser Gln Ser Cys Leu Ile Ala Gly Lys Pro Asp Glu Cys
 260 265 270

Gln Lys Val Ile Glu Lys Leu His Leu Lys Leu Pro Ala Val Pro Val
 275 280 285

Thr Gln Gly Met Ile Gly His Cys Pro Glu Ala Ile Pro Tyr Leu Asp
 290 295 300

Gln Ile Ser His Ile His Glu Met Leu Glu Ile Pro Lys Pro Glu Asn
 305 310 315 320

Val Lys Leu Phe Thr Thr Ser Glu Asn Arg Glu Leu Val Ser Met Lys
 325 330 335

Asp Ser Val Ser Lys Leu Val Ala Glu Ile Tyr Gln His Val Ala Asp
 340 345 350

Phe Pro Asn Ile Val Asn Lys Val Lys Glu Thr Cys Lys Thr Asp Ile
 355 360 365

Phe Ile Glu Leu Gly Ser Asn Asn Tyr Arg Ser Gly Ala Val Lys Thr
 370 375 380

Ile Leu Gly Pro Glu Ile Val Ser Val Ala Ile Asp Arg Gln Asn Glu
 385 390 395 400

Thr Ala Trp Gly Gln Leu Met Lys Met Val Ala Ser Leu Ile Ser His
 405 410 415

Arg Val Pro Gly Val Glu Leu Lys Lys Leu Tyr His Pro Glu Leu Leu
420 425 430

Lys Phe Asp Pro Gln Ala Lys Pro Asn Arg Phe Ile Arg Asn Ile Glu
435 440 445

Leu Asn Gly
450

<210> 108
<211> 1203
<212> ADN
<213> *Thraustochytrium* sp.

<400> 108

```

ctcaaaacat atgagggttga ctatcctttg tacacagggtg ccatgggctaa aggaattgcg      60
tctgctgatt tggttattgc tgctggtaaa tcaaagatct tggcatcatt tggagctggg      120
gggttggcct tacaagtggg agaagatgcc attaaacaaa ttaaagctga attggggaac      180
ggtccgtttg ctgtaaattt gattcattca ccattcgatc ctagcttgga gaagggtaac      240
gttgatcttt ttctaaaata taacgttcga tttgttgaag tatccgcatt tatgtcatta      300
accctcagg ttgtacgata cagagccgct gggttggcca aagcaagaga tggatctgtg      360
aaaattcaaa atcgtattat tgccaaaatt tcaagaacag agttagcgga actgttcttg      420
aaaccagcac ccaaaaatat tttagatgca ttggttgcgg atggatctat tagtcaagaa      480
caagcccaac ttgcattact tgtgccaatg gctgatgata ttactgtgga agctgattct      540
ggtgggcata ctgacaatcg accaattcat gttttgttac ctttgataat tcagcaaaga      600
aatagaattt gtaaacaata cccaaaacat ttaaaagttc gaatcggagc agctgggtgg      660
attggatgcc cgaaggcagc atttgctgcg tttgagatgg gtgctgcata cattgcaact      720
ggaacggtaa atcaactttc aaaggaagca ggtacttgtg actatgtacg taaagtattg      780
aataaagcta catattcgga tgttaccatg gctccagccg cagatatgtt cgatcatggg      840
gttgaattac aagttttgaa gaaagggtact atgtttcctt cacgtgctaa aaaactatac      900
gatttgttca aaaaatacaa atcgattgag gaattaccag cagatgaggt gaaaaaactt      960
gagcaaaaag ttttcaaaaa gtcgtttgat gaagtatggg atgagaccaa gaattactat     1020
attaatogtt tacattctcc cgaaaaaatt gaacgtgctg aaagagatgc aaaacttaaa     1080
atgtogttat gttttcgttg gtatttgtcg aagtcttcca gatgggctaa taccggtgaa     1140
tctggaagag tgcaggatta tcaaatttgg tgtgggtccag caattgggtc atataatgat     1200
ttt                                                                1203

```

<210> 109
<211> 401
<212> PRT
<213> *Thraustochytrium* sp.

5

<400> 109

ES 2 623 610 T3

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

Arg Lys Val Leu Asn Lys Ala Thr Tyr Ser Asp Val Thr Met Ala Pro
260 265 270

Ala Ala Asp Met Phe Asp His Gly Val Glu Leu Gln Val Leu Lys Lys
275 280 285

Gly Thr Met Phe Pro Ser Arg Ala Lys Lys Leu Tyr Asp Leu Phe Lys
290 295 300

Lys Tyr Lys Ser Ile Glu Glu Leu Pro Ala Asp Glu Val Lys Lys Leu
305 310 315 320

Glu Gln Lys Val Phe Lys Lys Ser Phe Asp Glu Val Trp Asp Glu Thr
325 330 335

Lys Asn Tyr Tyr Ile Asn Arg Leu His Ser Pro Glu Lys Ile Glu Arg
340 345 350

Ala Glu Arg Asp Ala Lys Leu Lys Met Ser Leu Cys Phe Arg Trp Tyr
355 360 365

Leu Ser Lys Ser Ser Arg Trp Ala Asn Thr Gly Glu Ser Gly Arg Val
370 375 380

Gln Asp Tyr Gln Ile Trp Cys Gly Pro Ala Ile Gly Ser Tyr Asn Asp
385 390 395 400

Phe

<210> 110

<211> 1350

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 110

atgggttggtt tacaaatgaa aaagaaacca gtatgggaga tgagtaagga agaacaaaagt	60
tctggaaaga atgttgtatt tgactatgat gaattgttg aatttgctga aggtgatatt	120
ggtaaagtct ttggacctaa gtttgatatt atcgataagt atagtcgacg tgtacgttta	180
cctgcgagag aatatcttct agttaccaga gttactttga tggatgctga agttgggaat	240
ttcagagttg gatctagaat ggttactgaa tatgatgttc cagtaaatgg tgaactttca	300
caaggtggtg atgttccatg ggctgttctt gttgaatctg gacaatgtga tcttatgtta	360
atatcttata tgggtattga ttttcaatgt aaaggtgac gtgtctatcg attattaaat	420
actacgttga cgttttacgg tgttgcctcat gagggtgaaa cactagtata cgatattcgt	480
gtaactggat ttgcaaaagg tatgcacggg gaaatctcca tgtttttttt tgaatatgat	540
tgttatgtga atggacgatt attaatcgaa atgagagatg gttgtgcggg attttttact	600
gatgaagaac ttgcagcagg taaaggagtt attaaaactg ttgctgaact tcataaaaaga	660
aaatctattg ttccaaaatc cattaaacct tttgctctaa atccagcagt acacaaaaca	720
atgttttctg aaaatgatat ggaaaaattg tgtgagcgtc aatgggaaaa tgtattgggt	780
agtggacttc aaggtattga ctacaagtta tgtgcacgga aaatgcttat gattgatcgt	840
attactaaaa tacaacataa tgggtggtgca tatggtcttg gattattggg tggcgaaaaa	900
attcttgaac gtgatcattg gtattttcca tgccattttg taaaggatca agttatggct	960
ggctcacttg ttagtgatgg ttgcagtcag ctactaaaac tttatatgtt atgggtgggt	1020
ttacatgatg tggttccaga ttttcaattt cgtccagttc ctggacaacc aaataaagtt	1080
cgttgccgtg gacaaattag tccacatcgt ggtaaacttg tttatgttat ggaaataaga	1140
gaaatgggat tcaatgaatc aactggacaa ccatatgeta ttgctgatgt tgatattatt	1200
gatgtaaaact atgaacttgg tcaatcattt gatatggctg atattgatag ttatggacgt	1260
ggtaatttgt caaagaaaat tgtggttgat tttaaaggaa ttgctttgca aatggaaggt	1320
accgtgaaat catcaaatat cattgattct	1350

<210> 111

<211> 450

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 111

ES 2 623 610 T3

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

Ser Gly Gln Cys Asp Leu Met Leu Ile Ser Tyr Met Gly Ile Asp Phe
 115 120 125

Gln Cys Lys Gly Asp Arg Val Tyr Arg Leu Leu Asn Thr Thr Leu Thr
 130 135 140

Phe Tyr Gly Val Ala His Glu Gly Glu Thr Leu Val Tyr Asp Ile Arg
 145 150 155 160

Val Thr Gly Phe Ala Lys Gly Met His Gly Glu Ile Ser Met Phe Phe
 165 170 175

Phe Glu Tyr Asp Cys Tyr Val Asn Gly Arg Leu Leu Ile Glu Met Arg
 180 185 190

Asp Gly Cys Ala Gly Phe Phe Thr Asp Glu Glu Leu Ala Ala Gly Lys
 195 200 205

Gly Val Ile Lys Thr Val Ala Glu Leu His Lys Arg Lys Ser Ile Val
 210 215 220

Pro Lys Ser Ile Lys Pro Phe Ala Leu Asn Pro Ala Val His Lys Thr
 225 230 235 240

Met Phe Ser Glu Asn Asp Met Glu Lys Leu Cys Glu Arg Gln Trp Glu
 245 250 255

Asn Val Leu Gly Ser Gly Leu Gln Gly Ile Asp Tyr Lys Leu Cys Ala
 260 265 270

Arg Lys Met Leu Met Ile Asp Arg Ile Thr Lys Ile Gln His Asn Gly
 275 280 285

Gly Ala Tyr Gly Leu Gly Leu Leu Val Gly Glu Lys Ile Leu Glu Arg
 290 295 300

Asp His Trp Tyr Phe Pro Cys His Phe Val Lys Asp Gln Val Met Ala
 305 310 315 320

Gly Ser Leu Val Ser Asp Gly Cys Ser Gln Leu Leu Lys Leu Tyr Met
 325 330 335

Leu Trp Leu Gly Leu His Asp Val Val Pro Asp Phe Gln Phe Arg Pro
 340 345 350

Val Pro Gly Gln Pro Asn Lys Val Arg Cys Arg Gly Gln Ile Ser Pro
 355 360 365

His Arg Gly Lys Leu Val Tyr Val Met Glu Ile Arg Glu Met Gly Phe

370		375		380
Asn Glu Ser Thr Gly Gln Pro Tyr Ala Ile Ala Asp Val Asp Ile Ile				
385		390		395
Asp Val Asn Tyr Glu Leu Gly Gln Ser Phe Asp Met Ala Asp Ile Asp				
	405		410	415
Ser Tyr Gly Arg Gly Asn Leu Ser Lys Lys Ile Val Val Asp Phe Lys				
	420		425	430
Gly Ile Ala Leu Gln Met Glu Gly Thr Val Lys Ser Ser Asn Ile Ile				
	435		440	445
Asp Ser				
450				

<210> 112
 <211> 1200
 <212> ADN
 <213> *Thraustochytrium* sp.

<400> 112

tgcttcaaac catttcctgg taatccttta gataacgatac atacacotgg taaaatgcct	60
ttaacatggt ttaatatgtc cgagtttatg tgtggtaaag tatcaaattg tcttggacca	120
gaatttaaga gatttgataa ctctaaaaca tccagaagtc ctgcctttga tcttgcaactt	180
gttacacgtg ttgtgagtg atcagatatg gaatttaaac ctcatTTaaa tattgatggt	240
aatccaagta aggggtacaat gatagggtgaa tttgattgcc ctgcagatgc gtgggtttttt	300
caaggatcat gtaacgatgg tcatatgcog tattctattg ttatggaaat tgctcttcaa	360
acttctggtg tattaacttc agttttgaaa gcacctttga ctatggataa agatgatatt	420
cttttccgca atttgatgc cactgctgaa atgggttcgaa gtgatggtga ttgtagaggt	480
aaaactatca aaaactttac tcaatgtacc gggttacagta tgctcggaaa aatgggaatt	540
catagattca catttgaatt atctgttgat gatgtagttt tctacaaagg atcaacatct	600
tttggttggt tcaccctga agtattcgag tcacaagttg gtcttgataa tggtaaaaaa	660
gtacaaccat ggtatttgga acaaaaatca tctaattgtag taacttatga cgttgcgtcc	720
actgctggca aggataagtt attttcaaag attggatcta aggatgcaca agttcaaaga	780
agaaatacac aatgtgagtt tctagatact atgcatatta ttccaaatac tggaaagtac	840
aacaaagggt atgctcatgg agaaaagaaa gttaatccaa acgactgggt ctttctctgt	900
catttctggt ttgatcctgt gatgcctggt tcattaggta ttgaaagtat gtttcaactc	960
attgaagcat tttcaattga tcaaggaatc gcttcaaaac atgggtattgt gaatccaact	1020
tttgctcatt ccaatggaaa aacttcttgg aaatacagag gtcaattgaa taacaaaggt	1080
aaacgaatgg atagtgaat tcatatcaaa gatattgtca aaaatgctga tggtagtgtt	1140
gatttgattg ctgatggatt tttattgggt gattcactaa gagtatactc tgcagatgat	1200

<210> 113

<211> 400

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 113

ES 2 623 610 T3

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

210	215	220
Tyr Leu Glu Gln Lys Ser Ser Asn Val Val Thr Tyr Asp Val Ala Ser		
225	230	235 240
Thr Ala Gly Lys Asp Lys Leu Phe Ser Lys Ile Gly Ser Lys Asp Ala		
	245	250 255
Gln Val Gln Arg Arg Asn Thr Gln Cys Glu Phe Leu Asp Thr Met His		
	260	265 270
Ile Ile Pro Asn Thr Gly Lys Tyr Asn Lys Gly Tyr Ala His Gly Glu		
	275	280 285
Lys Lys Val Asn Pro Asn Asp Trp Phe Phe Ser Cys His Phe Trp Phe		
	290	295 300
Asp Pro Val Met Pro Gly Ser Leu Gly Ile Glu Ser Met Phe Gln Leu		
305	310	315 320
Ile Glu Ala Phe Ser Ile Asp Gln Gly Ile Ala Ser Lys His Gly Ile		
	325	330 335
Val Asn Pro Thr Phe Ala His Ser Asn Gly Lys Thr Ser Trp Lys Tyr		
	340	345 350
Arg Gly Gln Leu Asn Asn Lys Gly Lys Arg Met Asp Ser Glu Ile His		
	355	360 365
Ile Lys Asp Ile Val Lys Asn Ala Asp Gly Thr Val Asp Leu Ile Ala		
	370	375 380
Asp Gly Phe Leu Leu Val Asp Ser Leu Arg Val Tyr Ser Ala Asp Asp		
385	390	395 400

<210> 114
 <211> 1729
 <212> ADN
 <213> *Thraustochytrium* sp.
 <400> 114

```

aggaatcgct tcaaaacatg gtattgtgaa tccaactttt gtcattcca atggaaaaac      60
ttcttgaaa tacagaggtc aattgaataa caaaggtaaa cgaatggata gtgaaattca      120
tatcaaagat attgtcaaaa atgctgatgg tactgttgat ttgattgctg atggattttt      180
attggttgat tcaactaagag tatactctgc agatgatctt cgcgtaaaaa ttgtaccggg      240
aaccaaagct gcacctaaat cagtagctgc tgctccaaga catgttgcaa caccaattcc      300
aggagtgcct tcgaatacaa gcagtgttga aatcagtttg gaatctttga agaaagaatt      360

gttaaactctt gagaaacat tgtatcttga aacttccaat catattgtaa aacaattcgg      420
tgacgttaac aatggccaag catccgttat tccaccatgc accatcaatg atttggttga      480
gcgtagtttt atggaaacat acaatgttgt tgcaccactt tacactggag ccatggctaa      540
aggatttga tctgctgatt tggtaattgc agctggtaaa agaaaaattt tgggttcttt      600
tggcgctgga ggcttaccaa tgcaactggg tctgcttctt gttgaaaaaa tccaagccgc      660
acttccagaa ggtccatacg ctgtcaactt gattcatagt ccattcgact caaatcttga      720
aaagggaaat gtagatctat ttttgaaaaa aggtgttcat gttgttgaag catctgcatt      780
cactgctctg accactcaag tagttcgta cctgtcatgt ggtttatctc gggctaaaga      840
cggatctgta ttgatcaaaa atagaatcat cggtaaagtt tcaagaaccg aattggctga      900
aatgtttttc agacctgcac cacaaaactt gcttgacaag cttattgcta gtggagaaat      960
cactaaagaa caagcttcat tggctttgga agtaccatg gctgatgatg tagctgttga     1020
agctgatagc ggtggacata ctgataatag accaattcat gtaatcctac ctttgattat     1080
caatctacga aatagaattc ataaagaatg tggttttcct gctgctttga gagttcgcgt     1140
tggtgctggt ggtggaattg gttgtccaag tgctgcagtt gctgcattca atatgggagc     1200
tgcattcttg attactggca gcgtcaacca agttagcaaa caatctggta cgtgtgatat     1260
cgtagaaaag caattatctg aagcttcgta ttcagatatt accatggcac cagcggtga     1320
tatgtttgat caaggagtcg agcttcaagt attaaaaaaa ggaactatgt tccatctcg     1380
tgcaaagaaa ttgtatgaat tattctgtat gtacaactca tttgatgaca tgccaaaaag     1440
cgaacttcaa agactagaga agcgaatttt tcaaaaatcg cttgcggaag tttgggaaga     1500
aactaaagat ttttatatca atcgtttgaa taatcctgag aagattgaac atgctgagaa     1560
gaaagatcca aagttgaaga tgtcattatg ctttagatgg tatttgggtt taagttcatt     1620
ttgggcaaac aatggaatta aagaaagatc aatggactat caaatgttgt gtggtccagc     1680
gattggttca tacaatgatt ttgtaaaagg aacttatttg gatcctgca                       1729

```

<210> 115

<211> 455

<212> PRT

<213> *Thraustochytrium* sp.

<400> 115

ES 2 623 610 T3

Asn Leu Glu Lys Pro Leu Tyr Leu Glu Thr Ser Asn His Ile Val Lys
1 5 10 15

Gln Phe Gly Asp Val Asn Asn Gly Gln Ala Ser Val Ile Pro Pro Cys
20 25 30

Thr Ile Asn Asp Leu Gly Glu Arg Ser Phe Met Glu Thr Tyr Asn Val
35 40 45

ES 2 623 610 T3

Val 50	Ala	Pro	Leu	Tyr	Thr	Gly 55	Ala	Met	Ala	Lys	Gly 60	Ile	Ala	Ser	Ala
Asp 65	Leu	Val	Ile	Ala	Ala 70	Gly	Lys	Arg	Lys	Ile 75	Leu	Gly	Ser	Phe	Gly 80
Ala	Gly	Gly	Leu	Pro 85	Met	His	Leu	Val	Arg 90	Ala	Ser	Val	Glu	Lys 95	Ile
Gln	Ala	Ala	Leu 100	Pro	Glu	Gly	Pro	Tyr 105	Ala	Val	Asn	Leu	Ile 110	His	Ser
Pro	Phe	Asp 115	Ser	Asn	Leu	Glu	Lys 120	Gly	Asn	Val	Asp	Leu 125	Phe	Leu	Glu
Lys	Gly 130	Val	His	Val	Val	Glu 135	Ala	Ser	Ala	Phe	Thr 140	Ala	Leu	Thr	Thr
Gln 145	Val	Val	Arg	Tyr	Arg 150	Ala	Cys	Gly	Leu	Ser 155	Arg	Ala	Lys	Asp	Gly 160
Ser	Val	Leu	Ile	Lys 165	Asn	Arg	Ile	Ile	Gly 170	Lys	Val	Ser	Arg	Thr 175	Glu
Leu	Ala	Glu 180	Met	Phe	Phe	Arg	Pro	Ala 185	Pro	Gln	Asn	Leu	Leu 190	Asp	Lys
Leu	Ile 195	Ala	Ser	Gly	Glu	Ile	Thr 200	Lys	Glu	Gln	Ala	Ser 205	Leu	Ala	Leu
Glu 210	Val	Pro	Met	Ala	Asp	Asp 215	Val	Ala	Val	Glu	Ala 220	Asp	Ser	Gly	Gly
His 225	Thr	Asp	Asn	Arg	Pro 230	Ile	His	Val	Ile	Leu 235	Pro	Leu	Ile	Ile	Asn 240
Leu	Arg	Asn	Arg	Ile 245	His	Lys	Glu	Cys	Gly 250	Phe	Pro	Ala	Ala	Leu 255	Arg
Val	Arg	Val	Gly 260	Ala	Gly	Gly	Gly	Ile 265	Gly	Cys	Pro	Ser	Ala 270	Ala	Val
Ala	Ala	Phe 275	Asn	Met	Gly	Ala	Ala 280	Phe	Leu	Ile	Thr	Gly 285	Ser	Val	Asn
Gln 290	Val	Ser	Lys	Gln	Ser	Gly 295	Thr	Cys	Asp	Ile 300	Val	Arg	Lys	Gln	Leu

Ser	Glu	Ala	Ser	Tyr	Ser	Asp	Ile	Thr	Met	Ala	Pro	Ala	Ala	Asp	Met
305					310					315					320
Phe	Asp	Gln	Gly	Val	Glu	Leu	Gln	Val	Leu	Lys	Lys	Gly	Thr	Met	Phe
				325					330					335	
Pro	Ser	Arg	Ala	Lys	Lys	Leu	Tyr	Glu	Leu	Phe	Cys	Met	Tyr	Asn	Ser
			340					345					350		
Phe	Asp	Asp	Met	Pro	Lys	Ser	Glu	Leu	Gln	Arg	Leu	Glu	Lys	Arg	Ile
			355				360					365			
Phe	Gln	Lys	Ser	Leu	Ala	Glu	Val	Trp	Glu	Glu	Thr	Lys	Asp	Phe	Tyr
	370					375					380				
Ile	Asn	Arg	Leu	Asn	Asn	Pro	Glu	Lys	Ile	Glu	His	Ala	Glu	Lys	Lys
385					390					395					400
Asp	Pro	Lys	Leu	Lys	Met	Ser	Leu	Cys	Phe	Arg	Trp	Tyr	Leu	Gly	Leu
				405					410					415	
Ser	Ser	Phe	Trp	Ala	Asn	Asn	Gly	Ile	Lys	Glu	Arg	Ser	Met	Asp	Tyr
			420					425					430		
Gln	Ile	Trp	Cys	Gly	Pro	Ala	Ile	Gly	Ser	Tyr	Asn	Asp	Phe	Val	Lys
		435					440					445			
Gly	Thr	Tyr	Leu	Asp	Pro	Ala									
	450					455									

<210> 116
 <211> 903
 <212> ADN
 <213> *Thraustochytrium* sp.

<400> 116

ES 2 623 610 T3

```

tccattcggg aatctggtta cacgattagc ggagaaagat tcacaactga agctcacaaa    60
ttggttactg gaaagcctca tgctccgatt aagaagaagg atgctttcct agtatctggt    120
ggtgctcgtg gtattactcc actttgtatt cgtgaaattg ctaaagcagt gaaagggtggc    180
acttacattht tgatgggtcg atcagctttg gctgatgaac ccttgtgggc taatggtaaa    240
tccggaaaag atttagataa agctggtttg gcatttttga aggaagagtt tgcagctggg    300
cgtggtagta aaccaactcc aaaagttcac aaatctttga ttgataaagt gctcgggtatt    360
agggagggtta gagcatctat tgcaaatata gaagcccatg gagcaaaagc tatatatattg    420
tcttgcgatg tatcttccgc tgagaaagta aaggctgcag tgcaaaaagt tgaaaaggag    480
catctagttc gtattactgg tattgtgcat gcatcaggcg ttttgaggga taaattgggt    540
gagaacaaaa ctttgatga tttcaacgca gtatatggaa ccaaagtaac tggactagta    600

aacttgctgt cagcagtga catgaattht gttcgctcatt tggttatggt tagttctttg    660
gctggatatc atggaaatgt tggccaatct gattatgcaa tggctaacga atcacttaac    720
aagattgggt ttagattggg tgcagcttat tctcaattgt gtgttaaadc tatttgthtt    780
ggaccttggg atgggtggaat ggtaactcca gctttgaaaa aacaatttca atcaatgggt    840
gtccagatta ttctcgtga aggtggcgcg gagactgthg caagaatagt cttatcttca    900
aat                                                    903

```

<210> 117

<211> 201

5 <212> PRT

<213> *Thraustochytrium* sp.

<400> 117

```

Ser Ile Arg Glu Ser Gly Tyr Thr Ile Ser Gly Glu Arg Phe Thr Thr
1          5          10          15

Glu Ala His Lys Leu Val Thr Gly Lys Pro His Ala Pro Ile Lys Lys
20          25          30

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

Cys Ile Arg Glu Ile Ala Lys Ala Val Lys Gly Gly Thr Tyr Ile Leu
50          55          60

Met Gly Arg Ser Ala Leu Ala Asp Glu Pro Leu Trp Ala Asn Gly Lys
65          70          75          80

Ser Gly Lys Asp Leu Asp Lys Ala Gly Leu Ala Phe Leu Lys Glu Glu
85          90          95

Phe Ala Ala Gly Arg Gly Ser Lys Pro Thr Pro Lys Val His Lys Ser
100         105         110

Leu Ile Asp Lys Val Leu Gly Ile Arg Glu Val Arg Ala Ser Ile Ala
115         120         125

Asn Ile Glu Ala His Gly Ala Lys Ala Ile Tyr Leu Ser Cys Asp Val
130         135         140

Ser Ser Ala Glu Lys Val Lys Ala Ala Val Gln Lys Val Glu Lys Glu
145         150         155         160

His Leu Val Arg Ile Thr Gly Ile Val His Ala Ser Gly Val Leu Arg
165         170         175

Asp Lys Leu Val Glu Asn Lys Thr Leu Asp Asp Phe Asn Ala Val Tyr
180         185         190

Gly Thr Lys Val Thr Gly Leu Val Asn
195         200

```

<210> 118

<211> 39

5 <212> ADN

<213> *Thraustochytrium* sp.

<400> 118

ctaaagtctc atcaaattca tggtaaaaat gtttgcct 39

10

<210> 119

<211> 13

<212> PRT
<213> *Thraustochytrium* sp.

<400> 119

5

Leu Lys Ser His Gln Ile His Gly Lys Asn Val Leu Pro
1 5 10

<210> 120
<211> 8976
<212> ADN
<213> Secuencia artificial

10

<220>
<223> PFA1 con codones optimizados

15

<400> 120

atggaggacc agcgtattgc gatcgttggc cttagcgcca tccttccttc gggcgagAAC	60
gtccgcgagt cgtgggaggc gatccgtgac ggccctcaact gcctttccga cctgcccgcc	120
gaccgcgttg acgtcactgc ctactacaac cccaagaagg gcgtcaagga caagatctac	180
tgcaagcgtg gtggcttcat ccccgagtac gagtttgact cgcgcgagtt cggccctcaac	240
atgcttcaga tggaggactc ggacgccaac cagaccctca ccctgctcaa ggttaaggag	300
gccctcgacg acgccaacat tcccgcgttt accaaccgaga agaagaacat cggttgcgtc	360
ctcgggtattg gcgggtggta gaaggcctcg catgagttct acagccgcct caactacgtc	420
gtcgtggata aggtcctccg caagatgggc ctcccggacg aggacgtcga gactgctgtc	480
gagaagtcca aggcacaact tcccagatgg cgccttgact ccttccccgg ctttctcggg	540
aacgtcactg cgggccgctg caccaacacc ttcaacatgg agggcatgaa ctgcgtggtc	600
gatgccgcct gcgcctcgtc cctcatcgct atcaaggctg ccatcgatga gctgctccac	660
ggcgattgog acgcgatgat tgctggcgcg acgtgcaccg acaacgccct tggcatgtac	720
atggcctttt ccaagacccc cgtcttttcc acggaccaga gctgcctcgc ctacgacgag	780
aaaaccaagg gtatgctcat tggcgagggt tccgccatgt tcgtccttaa gcgctacgcc	840

gacgcgcgtcc gcgatggcga caccgtccac gccgtcatcc gctcgtgctc gtcctcctcc 900
 gacggcaagg cgtcgggtat ctacaccccg accatctcgg gccaggagga ggccatcctt 960
 cgcgcctacc gtcgtgccgg cgtgagcccg aacacgatca ccttctgtga ggcccatggc 1020
 accggcaccc cgtcggcga caagatcgag ctgaccgccc tcgcaacgt ctttgacaag 1080
 gcctacggcc ctggccacaa ggaggaggtc gctgtgggct ccatcaagtc gcagatcggc 1140
 cacctcaagg cgtcggcgg ctgagctggc ctgctcaagc tcgtgatggc tctcaagcat 1200
 aagacgctcc cgcagtcctat caacgtcgag aaccgcgcc acctcgtcga tggcactgtc 1260
 atctcggaca ccacgctcta catcaacacc atgaaccgcc cgtggatcac caagccgggc 1320
 gtcccccgtc gtgcccggcat ctccagcttc ggctttggcg gcgctaacta ccacgctgtc 1380
 cttgaggagt tcgagcccga gcagaccaag cctaccgcc tgaacgtttc ggcccagccg 1440
 atgtcctcc acgccgtcaa cgcgaactcg ctccagaagc tctgcgagga ccagctcaag 1500
 ctctcaagg agtcccgcga gaagtgcgtc aacacgaaga acaccgacta cgtcgtttt 1560
 tccaagtttc aggactcctt taagctcaag ggctccgtcc ccagccagca cgtcgcgtg 1620
 ggctttgctt ccaagagcat cgaggacacg atttccattc ttagcgccat tgtcaaccgc 1680
 ttccagaagg acatcacgac caccagctgg gcgtcccga aggaggggcg catctttcgc 1740
 agcacgcccc tcatcaacga caacaagtcc gtggccgcc tgttctcggg tcagggcgct 1800
 cagtaacccc acatgttcaa cgacgtcgcg atgcagtggc cgcagttccg cctctgcgtt 1860
 aacgatatgg agaaggccca ggaggagggt atcaacgaca agtcgggttaa gcgcattagc 1920
 caggtcatgt tccccgcaa gccctacgcg cgcgagagcc cctcgcaca caaggagatc 1980
 agcaagaccg agtactcgca gacgacgacc gtgcctcgt ccgtcggcct ctttgagatt 2040
 ttccgcgacg ccggctttgc ccggctttt gttgcgggcc actcgtcgg tgagttctcc 2100
 gccctttacg ccgtcggcct catcgaccgc gaggacctct ttaagctcgt gtgcaaccgc 2160
 gccatggcta tgcgcgacgc ccccaagaag tcgctgacg gcgccatggc tgcgctcatc 2220
 ggtccgaacg cctcgtccat caagctctcg gctcccagg tttgggtcgc gaacaacaac 2280
 tcgcccctgc agaccgtcat cactggtgcc aacagcggcg tccaggccga gacttcgaag 2340
 ctcaagacgc agggtttccg cgtgggtccac ctgcctcgc agggcgcgtt tcacagcccg 2400
 cacatggaga acgcgagaa gcagtttcag aaggccctct cggccgtcaa gttcaacaag 2460
 cccaccggct cgtcccccaa gattttcagc aacgtcaccg gcggtgtctt taccgatcct 2520
 aagacggccc tctcccgcga catgactagc tcgggtccagt ttctcaccca gatcaagaac 2580
 atgtacgccg ctggcgcccg cgttttcacg gagttcggcc ccaagcaggt cctctcgaag 2640
 ctgctcaacg agattttccc gggcgacacc agcgtcctca ctgttagcgt gaacctgcc 2700
 tcgccaagg actcggacat ccagctccgc caggcggccg tgcagatggc ggtcgtcggc 2760
 gtcgctctca ccgactttga taagtgggag cttaaggacc cgaccgcgat gaaggagttc 2820

cctcgcaaga aaacgacct caccctctcc gccgctacct acgttagcaa gaaaacgctc 2880
 caggagcgcg agcgtatcat gaacgacggt cgcactgtca gctgcgtgca gcgcacgcag 2940
 aacacgaaca cgggagagct tgagaagctc aagaagcagc tccaggacaa ggagaacgag 3000
 gttgtccgcg tccaggccct tgccaccag gccagcgccg accttcagaa caccaaggct 3060
 gagcttcaga aggtcaggc caccaagtgc tcgaacgctg cctcggacgc cgtcgtcgcc 3120
 aagcacaagg ccatcctcct cgctatgctg gaggagctgg agactggcaa ggccgtcgat 3180
 tactccagct tttccaaggg tcaggttgcc tcccctgcga ccgttcgtgt cgtgtcggt 3240
 cccgtgcagg ctgccgcacc ggttcaggtc agcgccctcc tggactcggg cctgctcgcg 3300
 aaggcggagc aggtcgtgct tgaggtcctc gcctccaaga cgggctacga gactgagctt 3360
 atcgagctgg acatggagct tgagactgag cttggtatcg attcgatcaa gcgcgtcgag 3420
 attctttcgg aggtccaggc ccagctcaac gtggaggcca aggacgttga cgcctgtcg 3480
 cgcacccgta cggtcggcga ggtcatcgat gccatgaagg cggagattgc cggcggtcag 3540
 cctgctgccc ccgtccaggc cgctgcgccg acgcaggctg tcgccccggt ccaggcctcc 3600
 gcgcctgtcg atageggcct cctcgccaag gcggagcagg tcgtccttga ggtgctcgct 3660
 tccaagactg gttacgagac tgagcttatt gagcttgaca tggagctgga gactgagctt 3720
 ggcatgact ccatcaagcg cgtggagatt ctgagcgagg tccaggccca gctcagcgtg 3780
 gaggccaagg atgtcgatgc cctctcccgt acgcgcaccg tcggcgaggt cattgacgcg 3840
 atgaaggccg agatcgcggg tggtcagccg gccgcccccg tccaggctgc tgcacctacg 3900
 caggctgtcg ctcccgtcca ggccagcgct cccgtcgact cgggccttct tgetaaggcc 3960
 gagcaggctg tccctgaggt ccttgccagc aagactggct acgagactga gcttattgag 4020
 cttgacatgg agcttgagac tgagcttggc atcgactcga ttaagcgctg cgagatcctc 4080
 agcgaggctc agggccagct ctccgtcgag gctaaggatg tggatgctct cagccgcacg 4140
 cgcacggtgg gcgaggctcat tgatgccatg aaggcggaga tttccggcgg tcagcccgt 4200
 gccccgtcc aggtcgtgct tccgaccag atcgtcgccc cggtcagggt ttcggtccg 4260
 gtggacagcg gctccttgc caaggccgag caggctgtcc ttgaggtcct cgccagcaag 4320
 accggctacg agactgagct gatcgagctt gacatggagc ttgagactga gctgggcatc 4380
 gattccatta agcgcgtcga gatcctctcg gaggtccagg cccagctcag cgtggaggcc 4440
 aaggatgtcg atgccctctc gcgtaccgt accgtcggcg aggttatcga tgcctatgaag 4500
 gccgagatca gcggcggtca gccacggcg cccgttcagg tcgctgcccc tacgcagatc 4560
 gttgcccctg tccaggctcag cgctcccgtg gacagcggcc tcctcgctaa ggctgagcag 4620
 gtgggtgctg aggtcctggc ctccaagacc ggctacgaga ctgagcttat cgagcttgac 4680
 atggagcttg agactgagct tggcattgac agcatcaagc gtgtcgagat cctctccgag 4740

gtgcaggccc agctcagcgt ggaggccaag gacgttgacg cgctcagccg tacgcgcacc 4800
 gttggcgagg tgatcgacgc catgaaggcc gagattagcg gtggtcagcc cgctgccccg 4860
 gttcaggtag ctgcccctac gcagatcgtc gcccccgtag aagcttcgcg ccctgtggac 4920
 agcggccttc tcgccaaggc cgagcaggtc gtccttgagg tgctggcctc caagaccggc 4980
 tacgagactg agctgatcga gcttgacatg gagctggaga ctgagcttgg catcgactcg 5040
 atcaagcgcg tggagattct ctcgaggtc caggcccagc tctcggtcga ggccaaggac 5100
 gtcgatgcgc tctcccgcac cgcaccgtg ggogaggtca tcgacgctat gaagcgggag 5160
 atcagcggcg gtcagccggc gggccctgtg caggtaggccc ctccgaccca gatcgctcgt 5220
 cctgtccagg tttccgcccc ggtggactcg ggccctctgg ctaaggccga gcaggtcgtc 5280
 cttgaggtcc tcgcttccaa gaccggctac gagactgagc tgatcgagct ggacatggag 5340
 cttgagactg agctgggcat cgattcgatc aagcgcgtcg agattctctc ggaggtccag 5400
 gccagctca acgttgaggc caaggacgtg gacgccctct cgcgtactcg caccgttggc 5460
 gaggttattg atgctatgaa ggccgagatc gccggtaggtc agccggtcgc ccctgttcag 5520
 gttgctgccc ctgcgcgggt ggtcgccccg gtccaggtgt ccaccccggt tgacagcggc 5580
 ctccctgcca aggccgagca ggttgtagctg gaggtcctcg cctgcaagac gggctacgag 5640
 actgagctta tcgagcttga catggagctg gagactgagc ttggcatcga ctccatcaaa 5700
 cgcgtcgaga ttctttcgga ggtccaggcc cagctgtcgg tggaggctaa ggatgtcgat 5760
 gccctcagcc gcacgcgcac ggtcggtag gtcactcgat ctatgaaggc cgagatttcg 5820
 ggcggtcagc ccaccgcccc cgtgcaggtc gcgcgcacca ccaggtcgt ggccccggtc 5880
 aaggtttcca cggccgtgga ctcgggcctt ctgcgaagg ccgagcaggt cgtgctggag 5940
 gttctcgctt ccaagacggg ttacgagact gagctgattg agcttgacat ggagctggag 6000
 actgagctgg gcattgactc catcaagcgc gtcgagatcc tctcgagggt ccaggcccag 6060
 ctcaacgtcg aggccaaagga cgtcgatgcc ctctcgcgca ccgcacccgt cggcgaggtc 6120
 attgatgcca tgaaggccga gatcgctggc gatcagcctg ccccggtgtt ggtcccgttg 6180
 caggccaagt cgggtgtcgc gaaccccgcc ctctcgcca aggcggagca ggtcgtgctg 6240
 gaggtcctgg ccagcaagac gggctacgag actgagctta tcgagcttga catggagctt 6300
 gagactgagc ttggtattga ctcgattaa cgcgttgaga tcctttccga ggtccaggcc 6360
 gagctgtccg tggaggccaa ggatgtcgat gcgctctccc gcacccgcac ggtggcgag 6420
 gtcactgacg ctatgaaggc cgagattgcc ggctccgcgg tcaactgtcg tacccttgac 6480
 gactcgacca ttatggagga gactgacgac gaggacgagg actttatcct gtacgaccac 6540
 gtctacggct ccgagtgcga ggatctctcg ctctcgttct cgtcgggtcaa gtcgattcct 6600
 cgcgcggaca agctcctgct ggacaacatt gccgagcgcc ccattgtcat tgtcgattgc 6660
 ggcacgaagc tcacgaccga gctggcgaag gccatcgggc agcgcgctgt cgttgccacg 6720

ttctcggccc	agtcgctcgt	gtcccgtggc	ttcgtgggca	agagcttcac	cctcggcaac	6780
accgaggagt	cggagatcga	gaagatggtg	tcctccatcg	agtcgtccta	cggcaagatc	6840
ggcggctttg	tctaccagca	ctttcatgac	agcgactacg	gtatgcagct	cggctgggct	6900
ctcatggcgg	cgaagcacct	caaggagtcc	ctcaacgacc	cgatcaagaa	cggccgcacc	6960
tttttcctgg	ctgtcgcccg	catgaacggc	aagctcggca	tggacaacgc	ctccgtccac	7020
gaccagggca	tcgtcgagag	ctgcggtatc	gctgagcgtg	gtgccatctt	tggcctctgc	7080
aagaccctgg	acctggagtg	gcctaacgtg	tttgcgcgcg	gtgtggacat	cgcggagggg	7140
atgtcctact	ccctcgcggc	cagagctgatc	gtcgatgaga	tcagctgcgc	caacctttcg	7200
atccgcgaga	gcggctacac	tattagcggc	gagcgcttca	ccacggaggc	gcacaagctc	7260
gtcacgggca	agcctcacgc	gcccatcaag	aagaaggacg	cctttctcgt	gtcgggtggt	7320
gctcgcggca	tcacgcccct	gtgcattcgc	gagattgcca	aggccgtcaa	gggtggcacc	7380
tacattctca	tgggcecgctc	ggcgctcgcg	gacgagcccc	tctgggctaa	cggcaagagc	7440
ggcaaggacc	tcgacaaggc	cggcctcgcc	ttccttaagg	aggagtctgc	tgccggccgt	7500
ggctcgaagc	ccacccccaa	ggtccacaag	tcgctcatcg	acaaggctct	cggcatccgc	7560
gaggttcgcg	cgtccatcgc	caacatcgag	gcgcacggcg	ctaaggccat	ctacctctcg	7620
tgcgatgtgt	cgaagccga	gaagggtcaag	gctgccgtcc	agaagggtcga	gaaggagcat	7680
ctcgtccgca	tcacgggcat	cgtgcacgcc	tcgggcgtcc	tgcgcgacaa	gctcgtcgag	7740
aacaagaccc	tcgacgactt	taacgctgtg	tacggcacga	aggtcacggg	cctcgtcaac	7800
ctccttagcg	ccgtcaacat	gaacttcgtc	cgcacacctg	tgatgttctc	gtcgtcgtct	7860
ggttaccacg	gcaacgtcgg	ccagtcggac	tacgctatgg	ccaacgagag	ccttaacaag	7920
atcggtctcc	gtcttggtgc	cgcgtactcc	cagctctgcg	tcaagtccat	ctgcttcggc	7980
ccttgggatg	gcggcatggt	gacgccggcg	ctcaagaagc	agttccagtc	catgggcggt	8040
cagatcatcc	ctcgcgaggg	tggcgccgag	actgtcgtc	gcattgtgct	ctcgtccaac	8100
cccagccagg	tcctcgtcgg	caactggggc	gtcccgcgcc	tcagccccct	ctccaagtgc	8160
gccaccatcg	tcagacctt	taccctgag	cttaaccct	tcctcaagtc	ccaccagatc	8220
cacggcaaga	acgtcctgcc	catgacggtc	gccattgggt	acctcgccca	cctcgtgaag	8280
aacttttatg	ccggccacca	cctctggggc	gtggaggacg	cgcagctctt	ctccggcgctc	8340
gtcatcgacc	acgcctgca	ggcccaggtc	aagctcactg	agcagagcct	ggatgacgat	8400
ggcaagggtca	aggtccaggc	ggtgctcacc	gcctcgaacg	acaacggcaa	gatgggtgccg	8460
gcctacaagg	ccgtcatcgt	gctcggcaag	acttcccgtc	cggccttcat	cctcaaggac	8520
ttttcgtctc	aggagtccaa	ctcgcgctcg	gccgacgagc	tgtacgacgg	caagaccctg	8580
ttccacggcc	cgtgttccg	tggcatcacc	aagctcctca	acgtgtccga	cactagcctc	8640

ES 2 623 610 T3

```

acgacccagt gcaccaacat cgatctcacc gccactgagc ggggccagtt tgccgacatc 8700
gagccggtca accctttcat ggcggacgcc gccttccagg ccatgctcgt ctgggtccgc 8760
aacctccgta actccgccag ccttccgaac aactgcgagc gcgtcgatat ctacaagccc 8820
atcgcgcccc gcgagaagta ctacaccacg ctgcaggccc tcggcaaacac ctccggctcg 8880
gttctcaagt ccgttttcta catgcatgac gagcagggcg aggtgttctt ctcgggccgc 8940
gccagcgtcg tggtaacga taagatggaa ttctaa 8976

```

<210> 121
 <211> 5880
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PFA2 con codones optimizados

<400> 121

atggtgaagc tttccgttgg tgacaacatt tgccacgata agcgcgtcgc cgtgggtcggc	60
atggccgtca tgtacgccgg ctgccagaac cagcacgagt tttggcagag cctccagggt	120
aagaacatga acagcaagag catcagccag aaccgcctgg gctccgagta ccgcgaggag	180
cactttaagc cggagcgcct gaagtacagc gacaccttct gcaacgagcg ttacggctgc	240
atcgacgaga acgtccagag cgagcatgag ctctctctga agctcgctaa ggacgcgata	300
gccgatacca agggcagcat cgaccttaac aagaccggca ttgtctccgg ctgcctctcg	360
ttccctatgg ataacctcca gggcgacctt ctcaacctct accagtgcc tattgagaag	420
aagatcggcc cgaacgcctt caaggatgtc aacctctggc cgaagcgcac gaccaacggc	480
aaggacgata agaaggccta ctctgatccc gccagcttcg tcgctgagca gcttgacatg	540
ggccccctcc actactcgtc cgacgtgcc tgcgcctccg ctctctacgt cctccgcctc	600
gcccaggacc acctcctcag cggtgccgcc gacaccatgc tctgcccgcg ctctgtgcctc	660
ccggagccct ttttcatact ttccggcttt tcgaccttcc accccatgcc cctttcgggt	720
gacgtgtcgg cccctcttca caagacgagc cagggcctca ctccgggcga gggcggtgct	780
atcatggtcc tgaagcgctt caacgatgcc attcgcgacg gcgaccgcat ctacggcacg	840
ctcctggggc ccgagctttc caacgcgggt tgcggcctcc cgctctccc gcacatgccg	900
tccgagttcg actgcatgga gaaggccctc cagcgcgttc accgcctccc gtctctcatc	960
cagtacgtgg agtgccacgc cactggcacc ccgcagggcg acaaggtcga gatcgacgcc	1020
atgacgaagt gcttcggcga gcatctgcct cgcttcggct ccaccaaggg taacttcggc	1080
cacaccctcg tggtgctgg ctttgccggc atgtgcaagg tctctctctc gatgcagtac	1140
ggtgagattc ctctacgcc tggcctggag aaccccgaca acattatgca cgatcttgtc	1200
gttaccgaga ctattccctg gccgaacacc aacggcgatc ttaagcgtgc gtgcctcagc	1260
gcctttggct ttggcggtac taacgcccac gccgtgttcg aggagtaccg cagcgacctt	1320

caggccaaca agacccttga gaacgagagc aagteccacg agatcttttc ctccctttaag 1380
 attgccattg ttggcatgga gtccgagttt ggcaactctca agggcctcca ggagttcgag 1440
 cgtgccatct acaacggcgg ccacggcgcg tgcgaccttc cggagaaccg ctggcgcttt 1500
 ctcggtgagg acaaggagtt tctccaggcc tgcggcctcc agaagctccc gcgtggctgc 1560
 tacatcaagg aggtcgagac tgactttaag cgccttogcc tccccatgat ccaggaggac 1620
 atcctccgcc ccctccagct cctcgccgtg tcgatcatcg accgcgcct caacgccage 1680
 ggcgtaagc ccaacggcaa ggtcgccgtc ctcggtgggc tcggcaccga tcttgagctc 1740
 taccgccacc gcgctcgcg tgcctgaag gagcgccctc agaccgccgt caaggaggac 1800
 atccccctgc tggagaagct catgaactac gtgaacgacc gcggcacctc cagctcctac 1860
 acctcgta ca tccgcaacct cgttgcgacc cgcgtcagct cgctctgggg ctccaccggc 1920
 cctagcttca cgatcacgga gggcgagaac tcggtttacc gttgcctcga cctcgccgc 1980
 tggttcctcg ccaacgggtga ggtcgatgcc gtggttgctg ctggcggtga tctctcgggc 2040
 tcggccgaga acctgttcgt caagtgcgc cgcctcaagg tgtccacca gaacgagccc 2100
 tttgctaact ttgagtcgaa cgcgcacggc tacttcgcgc gcgacggctg cggtgccctc 2160
 gttctcaagc gcctttcgga ctgcactgac tccaccgaga agatctacgc gaccgtggac 2220
 agcattgctg tcggcgacga ggtgggccc actattaagc aggcctgaa gaacgcctcg 2280
 atcgccgca aggacatcga gctcgcgag ctctccgct ccagcggcaa gcaccactcc 2340
 ggccgcatca cctgcgagga cgagcttaac gagctcggc agatcttcaa cgagggcatt 2400
 cagcgcgtgg ccatcggcag cgtcaaggcc aacgtcggc acgtcggtc cgcctccggt 2460
 gctgccagcc tcatcaagac ggccctctgc ctctacaacc gctacctccc caagctcccc 2520
 aactggaaca agccgaccaa ggacgtcgag tggtcgaaga gcttctttgt ctgcgagcac 2580
 tcgcgcgcct ggctcaagaa cgtggacgag aacggccacg cggtcgtgag cggcgtctgc 2640
 gagaacggct cctgctacgg catcgatcag agcgacgtcc agggccacca tgaggagtgc 2700
 aacctcggtg ccctcgataa gaacgagccc aaggtgctcg gtatctacgg cgattccgtg 2760
 gacgatattc tggccagct gaacaagtac ctggagaagt tccttcagga gactggcact 2820
 gctgcggctg cgcagaagggt gaagagccct accattgaca tcgactcgaa cgtctttgcc 2880
 gagatgctga accttcccc ggacaagaac aagaagtttg ccgtcgctct ggtcacgacc 2940
 cccaacaagc tccagcgca gattgagctc gccgttaagg gcatccctcg ctgcgtgaag 3000
 gccaaagcgc actggtgctc cccctccggc agcatctttg cgtgcaacc gctcaagtcg 3060
 gacaacattg cctttatgta cggcgagggc cgcctgcctt acgcggcct cggctacgat 3120
 ctccaccgca tctggcccat gttcacgag ctcgtaaca accgcacgac tgagctgtgg 3180
 gaccaggggtg actcggtgta cctgccgcgc agctcctccg tggccgagaa ggagaaggtc 3240

tttggcgact	tcgacaagaa	ccagatcgag	atgttccgcc	tcggtatatt	cgtcagcatg	3300
tgctttaccg	acatggcgac	ggagctcctc	ggccttaagc	cgaaggccgc	tttcggcctc	3360
tccttcggcg	agatcagcat	gctctttgct	ttctcgaaga	agaacaccaa	gctctccaag	3420
gagcttactc	gccgcctcaa	ggaggccaag	gtgtgggcgt	cgcagctggc	cgtcgagttc	3480
gccgccatcc	gcgacctttg	gaacatcccg	gccgacaagt	ccatcgatga	gttctggcag	3540
ggttacttcg	tttacgccaa	ccgtacgctc	gtggagaaca	ccattggcga	gaacaagttc	3600
gtccgcctcc	ttatcgtaaa	cgactcccag	tcctgcctca	ttgccggtaa	gcccgatgag	3660
tgccagaagg	tcacgcagaa	gctccacctt	aagctccccg	ccgtccccgt	caccaggggc	3720
atgattggcc	actgcccgga	ggccattccc	tacctcgacc	agatcagcca	catccacgag	3780
atgcttgaga	tcccgaagcc	tgagaacgtc	aagctcttca	cgacgtccga	gaaccgcgag	3840
cttgtctcga	tgaaggactc	cgttagcaag	ctcgtcggcg	agatctacca	gcacgtcgct	3900
gacttcccca	acattgtcaa	caagggtcaag	gagacttgca	agacggacat	tttcatcgag	3960
ctgggcagca	acaactaccg	ttccgggtgc	gtcaagacta	tcctcgggtc	ggagatcggt	4020
agcgttgcca	tcgaccgtca	gaacgagact	gcctggggcc	agctcatgaa	gatggtcgcc	4080
agcctgatct	cccaccggt	ccccggcgtc	gagctcaaga	agctgtacca	tcgggagctc	4140
ctgaagttcg	atccccaggc	caagcccaac	cgttttatcc	gcaacatcga	gtcacaacggc	4200
tttttcgacc	gcacgaacat	catcgctgat	aagcagcttt	cccctgcgga	cccgaagctc	4260
gccgagatcg	tcaacaaccg	caacatgccg	aaggataacg	tgtacgtccc	cattgagcgc	4320
gtcaagacga	tgatcaaggc	cgagcccgt	aacctccagg	tgcccgctcg	ctcgaagccc	4380
gtggtcaccg	agcgtatctc	gtcggacgac	aacctctttg	agaagctctc	ggagatcact	4440
aagtccttcg	acggtgtcaa	cgcctgcacc	gaggccatgc	tcggcgattc	gggctttctc	4500
aagacgtacg	aggttgacta	cccgtctctc	accggcgcta	tggccaaggg	tatcgctccc	4560
gccgacctcg	tcattgcggc	gggtaagtcg	aagatccttg	cgctcctttg	tgctggcggc	4620
ctcgtctctc	aggtggctga	ggatgccatt	aagcagatca	aggctgagct	tggcaacggt	4680
ccctttgccg	tcaacctcat	ccactcgctt	ttcgacccct	cgtttgagaa	gggcaacggt	4740
gaccttttcc	tcaagtacaa	cgtccgcttt	gtcagaggtga	gcgcgttcat	gagcctcacc	4800
ccccaggtcg	ttcgtctaccg	cgtcgccggc	cttgccaagg	cccgtgacgg	ctcgggtcaag	4860
attcagaacc	gcacatcgcg	caagatttcg	cgcacggagc	tggccgagct	cttctctcaag	4920
cccgtctccga	agaacatcct	cgatgccttc	gttgccgacg	gctcgatttc	ccaggagcag	4980
gctcagctcg	cgtcctctgt	ccctatggcc	gatgacatca	ccgttgaggc	cgactccggt	5040
ggccacaccg	acaaccgccc	cattcatgtg	ctcctcccc	tcacatcca	gcagcgcaac	5100
cgcatttgca	agcagtaccc	gaagcacctc	aaggctccga	tcggcgctgc	cggtggcatc	5160
ggttgcccta	aggcggtctt	tgccgccttt	gagatgggtg	cggcctacat	cgccacgggc	5220

```

accgttaacc agctctcgaa ggaggccggc acctgcgact acgtgcgcaa ggtgctcaac 5280
aaggccacct actccgacgt cacgatggct cccgctgccg acatgttcga ccacggtgtc 5340
gagctccagg ttctcaagaa gggcaccatg ttccgctcgc gcgccaagaa gctctacgac 5400
ctctttaaga agtacaagtc gatcgaggag ctccctgccg acgaggtcaa gaagctggag 5460
cagaagggtt ttaagaagtc gttcgacgag gtctgggacg agactaagaa ctactacatt 5520
aaccgcctcc actcccctga gaagatcgag cgcgcggagc gtgacgcaa gctgaagatg 5580
tcgctctgct ttcgttggtg cctgagcaag tcgtcccgtc gggccaacac cggcgagtgc 5640
ggccgtgtcc aggactacca gatctggtgc ggccccgcca tcggctcgta caacgacttc 5700
gcgaagggct cgcctgcct tgaccctgag atccttggtc cgttcccgtc ggttgtccag 5760
atcaacaagc atattctgcg cggcgcttgc ttctaccagc gtctttcgca gctcaagtac 5820
cttaacttca actacgagga gctcgatacg ctcacctaca gcgctagcaa ctttatctaa 5880

```

<210> 122
 <211> 4368
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PFA3 con codones optimizados

<400> 122

atgggtggcc tgcagatgaa gaagaagcct gtgtgggaga tgtcgaagga ggagcagtcg	60
tccggcaaga acgtcgtctt tgactacgac gagctcctcg aattcgcgga gggtgacatc	120
ggcaaggtgt tcggcccca gtttgacatc atcgacaagt acagccgccg tgtgcgcctc	180
ccggcccgcg agtacctcct cgtcacccgt gtcacgctca tggatgccga ggtcggcaac	240
ttccgcgtcg gctcgcgcct ggtcacccag tacgacgtcc cggatgaacgg cgagctttcc	300
cagggcgggc acgttcctct ggccgtcctc gtcgagtcgg gccagtgcga cctcatgctt	360
atctcgtaca tgggcattga ctttcagtgc aagggtgacc gcgtttaccg ccttctcaac	420
acgacctca cgttctacgg tgtcgcccac gagggcgaga ctctcgttta cgacatccgc	480
gtcactggtt tcgccaaggc catgcacggc gagattagca tgttcttctt cgagtacgac	540
tgctacgtca acggccgcct gctcatcgag atgcgcgacg gttgcgctgg cttcttcacg	600
gacgaggagc tcgccgcggg caagggcgtc atcaagaccg tcgctgagct ccacaagcgc	660
aagtcgattg tgcccaagtc gatcaagcct tttgccctca accccgccgt ccacaagacg	720
atgttcagcg agaacgacat ggagaagctt tgcgagcgcc agtgggagaa cgtcctcggc	780
tccggcctcc agggcatcga ctacaagctg tgcgcccgca agatgctcat gatcgaccgc	840
atcacgaaga tccagcacia cggcggtgcg tacggcctcg gcctcctcgt tggcgagaag	900
attcttgagc gcgaccattg gtacttcctt tgccacttcg tcaaggacca ggtgatggcg	960

ggctccctcg ttagcgacgg ctgctcgag ctgctcaagc ttacatgct ttggctcggc 1020
 ctccacgacg tggccccga tttccagttc cgtcctgtcc ctggccagcc caacaagggtg 1080
 cgctgccgtg gccagatcag ccccatcgt ggcaagctcg tgtacgtgat ggagattcgc 1140
 gagatgggtt tcaacgagtc caccggccag ccctacgcca tcgctgacgt tgacattatc 1200
 gatgtgaact acgagctcgg ccagtccttt gacatggccg acatcgactc gtacggccgt 1260
 ggcaacctct ccaagaagat tgtgtcgat ttcaagggca ttgcgtcca gatggagggc 1320
 accgtcaaga gctccaacat catcgattcg tccccaaagt ccacgattat ccagccgcg 1380
 cccaactgcc tccgcgccga tctctcgc ccctcgcagg tcacctggca cccgatggcc 1440
 ggtgtcaacg gcgccccgc cccctccttc agcccgctcg attaccctcc tcgtgccgtt 1500
 tgetttaagc ccttccttg caacccctc gacaacgac atacgcggg caagatgccg 1560
 ctgacctggg ttaacatgtc ggagtttatg tgcggcaagg tcagcaactg ccttggccct 1620
 gagtttaagc gcttcgacaa ctccaagacg agccgctccc cggccttcga cctggccctg 1680
 gttacgcgcg tgggtgtcgg cagcgatatg gagttcaagc cccacctcaa catcgacgtc 1740
 aaccgctcga agggcacgat gattggcgag ttcgactgcc ccgctgacgc ctggttcttt 1800
 cagggtcct gcaacgacgg ccacatgccg tacagcatcg tcatggagat cgccttcag 1860
 accagcggtg tctcaccctc cgtcctcaag gccccgctca ctatggacaa ggacgacatt 1920
 ctctttegca acctcgacgc caccgcccag atggtccgtt ccgacgtcga ttgccgcgg 1980
 aagaccatca agaacttcac ccagtgcacc ggctacagca tgcttgcaa gatgggcatc 2040
 caccgcttca cttttgagct ctcggtcgat gacgtcgtgt tttacaaggg ctcgaccagc 2100
 tttggttggg tcacgcggga ggtgtttgag tcgcaggtcg gcctcgataa cggcaagaag 2160
 gtccagccgt ggtatctgga gcagaagtcg tcgaacgtgg tgacgtacga tgtcgccctc 2220
 accgcccggc aggacaagct cttctcgaag atcggtcga aggacgtca ggtccagcgt 2280
 cgcaacaccc agtgcgagtt tctcgacacg atgcacatta ttccgaacac cggcaagtac 2340
 aacaagggct acgcgcacgg tgagaagaag gtcaaccca acgactgggt cttctcctgc 2400
 cacttttggg tcgaccgggt gatgccggc tccctcggtt ttgagtccat gttccagctc 2460
 atcgaggcct tttcgattga ccagggtatc gcgtccaagc atggcatcgt gaacctacc 2520
 ttgcgcact cgaacggcaa gacctcgtgg aagtaccgc gccagctcaa caacaagggc 2580
 aagcgcattg acagcgagat tcacatcaag gatattgtca agaacgccga cggtaactgtc 2640
 gatctcatcg ccgatggtt tcttctcgtg gactcgcttc gcgtttacag cgcgatgac 2700
 ctccgcgtca agatcgtccc cggcactaag gctgctcca agagcgtcgc ggcgcgtccg 2760
 cgccatgtgg ccactccgat ccccgcgctc cctccaaca cctcctcggg ggagatctcg 2820
 cttgagtccc ttaagaagga gctcctcaac ctggagaagc ccctctacct tgagacttcc 2880
 aaccacatcg tcaagcagtt cggcgacgtt aacaacggcc aggcctccgt catccctccg 2940

tgcaccatta acgatctcgg tgagcgctcg tttatggaga ettacaacgt cgtcgctccc 3000
ctctacaccg gcgcgatggc gaagggcatc gcttcggggg acctcgatcat cgctgccggg 3060
aagcgcaaga tccctcggcag cttcggcgcc ggtggcctcc cgatgcacct cgtgcgcgcc 3120
tcggtcgaga agatccaggc cgccctcccg gagggcccg acgcgggtcaa cctcatccac 3180
tcgcctttcg actcgaacct tgagaagggg aacgtggacc tctttctgga gaagggcgtc 3240
cacgtggteg aggcctccgc ctttaccgcc ctcaacgacc aggtcgttcg ctaccgcgcc 3300
tgccggcctct cgcgtgctaa ggacgggttc gtgctgatta agaaccgcat catcggttaag 3360
gtcagccgca cggagcttgc cgagatgttc tttcgccctg cccccagaa cctcctcgat 3420
aagctcatcg ccagcggcga gatcaccaag gagcaggcgt cctcgctct tgaggttcc 3480
atggccgacg atgtcgctgt tgaggccgac tccggcggcc acaccgataa ccgtcccatc 3540
cacgtcatec tcccgctgat tattaacctc cgcaaccgta tccacaagga gtgcggcctt 3600
cctgccgctc tcccgctccg cgtcggcgct ggtgggtggc tccggttccc ctccggcgt 3660
gtcgccgctc tcaacatggg cgcggccttc ctgatcaccg gctccgttaa ccaggtgagc 3720
aagcagtcgg gcacgtgcga cattgtgcgc aagcagctta gcgaggccag ctactccgac 3780
atcacgatgg ctcccgcgc tgacatgttc gaccagggcg tggagctcca ggtcctcaag 3840
aagggtagca tgtttccctc gcgcgccaaag aagctctacg agctcttttg catgtacaac 3900
agctttgacg acatgccgaa gtccgagctc cagcgccctgg agaagcgcat tttccagaag 3960
agcctcgccg aggtctggga ggagactaag gacttttaca tcaaccgct caacaaccg 4020
gagaagatcg agcacgccga gaagaaggac cccaagctca agatgtccct ttgctttcgc 4080
tggtatctcg gcctttcgag cttttgggcc aacaacggca tcaaggagcg cagcatggat 4140
taccagattt ggtgcggccc ggccattggc agctacaacg acttcgtgaa gggcacctac 4200
ctcgaccccg ccgtcgccgg ttcgtaaccc tgcgtgggtc agatcaacat gcagatcctc 4260
cgcggtgctg gcttctcca gcgcgtccgc gccattaagc acgaccgcg cctcgatatc 4320
gacgttgatg aggacgtctt tacctaccgc cccgagagca ccctctaa 4368

<210> 123

<211> 33

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador sintético pDS233

10

<220>

<221> misc_feature

<222> (25)..(25)

<223> n e s a, c, g o t

15

<220>

<221> misc_feature

	<222> (28)..(28)	
	<223> n es a, c, g o t	
	<220>	
5	<221> misc_feature	
	<222> (31)..(31)	
	<223> y es c o t	
	<400> 123	
10	tgatatggga ggaatgaatt gtgtngtnga ygc	33
	<210> 124	
	<211> 36	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> Cebador sintético pDS235	
20	<220>	
	<221> misc_feature	
	<222> (28)..(28)	
	<223> n es a, c, g o t	
25	<220>	
	<221> misc_feature	
	<222> (31)..(31)	
	<223> r es a o g	
30	<220>	
	<221> misc_feature	
	<222> (34)..(34)	
	<223> n es a, c, g o t	
35	<400> 124	
	ttcataaca aaatgataat tagctccncc raancc	36
	<210> 125	
	<211> 23	
40	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador sintético prDS183	
45	<220>	
	<221> misc_feature	
	<222> (15)..(15)	
	<223> y es c o t	
50	<220>	
	<221> misc_feature	
	<222> (19)..(19)	
	<223> m es a o c	
55	<220>	
	<221> misc_feature	
	<222> (21)..(21)	
	<223> n es a, c, g o t	
60	<400> 125	
	ggcggccaca ccgayaaymg ncc	23
	<210> 126	
65	<211> 22	
	<212> ADN	

<213> Secuencia artificial
 <220>
 <223> Cebador sintético prDS184
 5
 <220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n es a, c, g o t
 10
 <220>
 <221> misc_feature
 <222> (20)..(20)
 <223> r es a o g
 15
 <400> 126
 cggggccgca ccanaytgr ta 22
 20
 <210> 127
 <211> 17
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Cebador sintético prDS181
 <220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n es a, c, g o t
 30
 <220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n es a, c, g o t
 35
 <220>
 <221> misc_feature
 <222> (14)..(14)
 <223> s es c o g
 40
 <220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n es a, c, g o t
 45
 <400> 127
 tccttcggng cngsngg 17

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada seleccionada entre el grupo que consiste en:

- (a) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos al menos un 80 % idéntica a la SEQ ID NO: 5, en la que la secuencia de polinucleótidos codifica un polipéptido que tiene actividad de ácido graso poliinsaturado (PUFA) sintasa que consiste en actividad de β -hidroxiacil-ACP deshidrasa similar a FabA (DH) y actividad de enoil-ACP reductasa (ER);
- (b) una molécula de ácido nucleico que comprende un polinucleótido que tiene la secuencia que se establece en la SEQ ID NO: 5;
- (c) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifica un polipéptido, en la que el polipéptido comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de PUFA sintasa que consiste en actividad de DH y actividad de ER; y
- (d) una molécula de ácido nucleico que comprende una secuencia de polinucleótidos que codifican un polipéptido que comprende la secuencia de aminoácidos que se establece en la SEQ ID NO: 6.

2. Una molécula de ácido nucleico aislada que comprende una secuencia de polinucleótidos que es totalmente complementaria a la secuencia de polinucleótidos de la reivindicación 1.

3. Una molécula de ácido nucleico recombinante que comprende la molécula de ácido nucleico de la reivindicación 1, o combinaciones de las mismas, y una secuencia de control de la transcripción.

4. Una molécula de ácido nucleico recombinante de la reivindicación 3 en la que la molécula de ácido nucleico recombinante es un vector recombinante.

5. Una célula hospedadora que expresa la molécula de ácido nucleico de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas.

6. La célula hospedadora de la reivindicación 5, en la que la célula hospedadora se selecciona entre el grupo que consiste en una célula vegetal, una célula microbiana y una célula animal.

7. La célula hospedadora de la reivindicación 6 en la que la célula hospedadora es una célula microbiana.

8. La célula hospedadora de la reivindicación 7 en la que la célula microbiana es una bacteria o un traustóquitrido.

9. La célula hospedadora de la reivindicación 8 en la que el traustóquitrido es un *Schizochytrium* o un *Thraustochytrium*.

10. Un método para producir al menos un ácido graso poliinsaturado (PUFA), que comprende:

expresar un gen de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir PUFA, en el que el gen de PUFA sintasa comprende la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas, y en el que se produce al menos un PUFA.

11. El método de la reivindicación 10 en el que el al menos un PUFA comprende ácido docosahexaenoico (DHA).

12. Un método para producir lípidos enriquecidos con ácido docosahexaenoico (DHA), que comprende:

expresar un gen de PUFA sintasa en una célula hospedadora en condiciones eficaces para producir lípidos, en el que el gen de PUFA sintasa comprende la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas en la célula hospedadora, y en el que se producen lípidos enriquecidos con DHA.

13. Un método para preparar un vector recombinante que comprende insertar la molécula de ácido nucleico aislada de la reivindicación 1 en un vector.

14. Un método para preparar una célula hospedadora recombinante que comprende introducir el vector recombinante de la reivindicación 4 en una célula hospedadora.

15. Un polipéptido aislado codificado por la secuencia de polinucleótidos de la reivindicación 1.

16. Un polipéptido aislado seleccionado entre el grupo que consiste en:

- (a) un polipéptido que comprende una secuencia de aminoácidos al menos un 80 % idéntica a la SEQ ID NO: 6, en la que el polipéptido tiene una actividad de ácido graso poliinsaturado (PUFA) sintasa que consiste en actividad de β -hidroxiacil-ACP deshidrasa similar a FabA (DH) y actividad de enoil-ACP reductasa (ER); y
- (b) un polipéptido que comprende una secuencia de aminoácidos como se establece en la SEQ ID NO: 6.

5 17. El polipéptido aislado de la reivindicación 15 o la reivindicación 16, en las que el polipéptido es un polipéptido de fusión.

10 18. Una composición que comprende el polipéptido de una cualquiera de las reivindicaciones 15-17 y un vehículo biológicamente aceptable.

19. Un método para aumentar la producción de DHA en un organismo que tiene actividad de ácido graso poliinsaturado (PUFA) sintasa, que comprende:

15 expresar la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas, en el organismo en condiciones eficaces para producir DHA, y en el que la producción de DHA en el organismo aumenta.

20 20. Un método para aislar lípidos a partir de una célula hospedadora, que comprende:

- (a) expresar un gen de ácido graso poliinsaturado (PUFA) sintasa en la célula hospedadora en condiciones eficaces para producir lípidos, en el que el sistema de poliquétido sintasa (PKS) comprende la molécula de ácido nucleico aislada de la reivindicación 1, la molécula de ácido nucleico recombinante de la reivindicación 3, o combinaciones de las mismas, en la célula hospedadora, y
- 25 (b) aislar lípidos a partir de la célula hospedadora.

21. El método de la reivindicación 20 en el que los lípidos comprenden ácido docosaheptaenoico (DHA).

30 22. El método de una cualquiera de las reivindicaciones 10, 14 y 20, en el que la célula hospedadora se selecciona entre el grupo que consiste en una célula vegetal, una célula animal aislada y una célula microbiana.

Figura 1. Arquitectura genética de la cepa PTA-9695 de la ATCC de *Schizochytrium* sp.

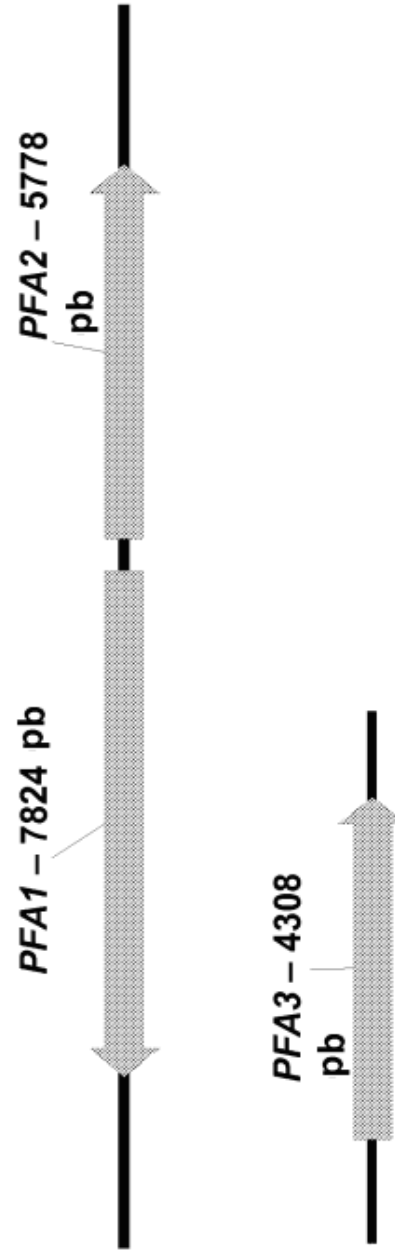


Figura 2. Arquitectura genética de la cepa PTA-10212 de la ATCC de *Thraustochytrium sp.*

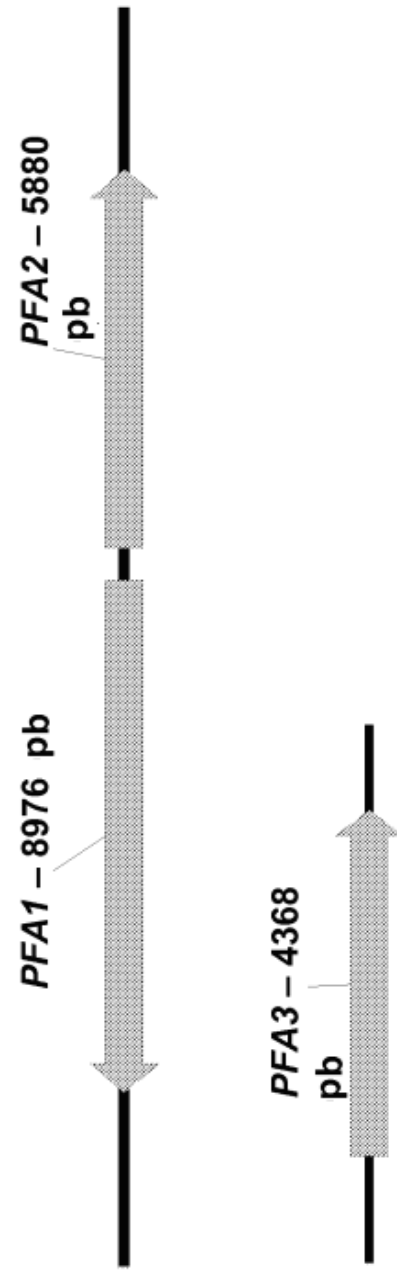


Figura 3. Arquitectura de dominio:

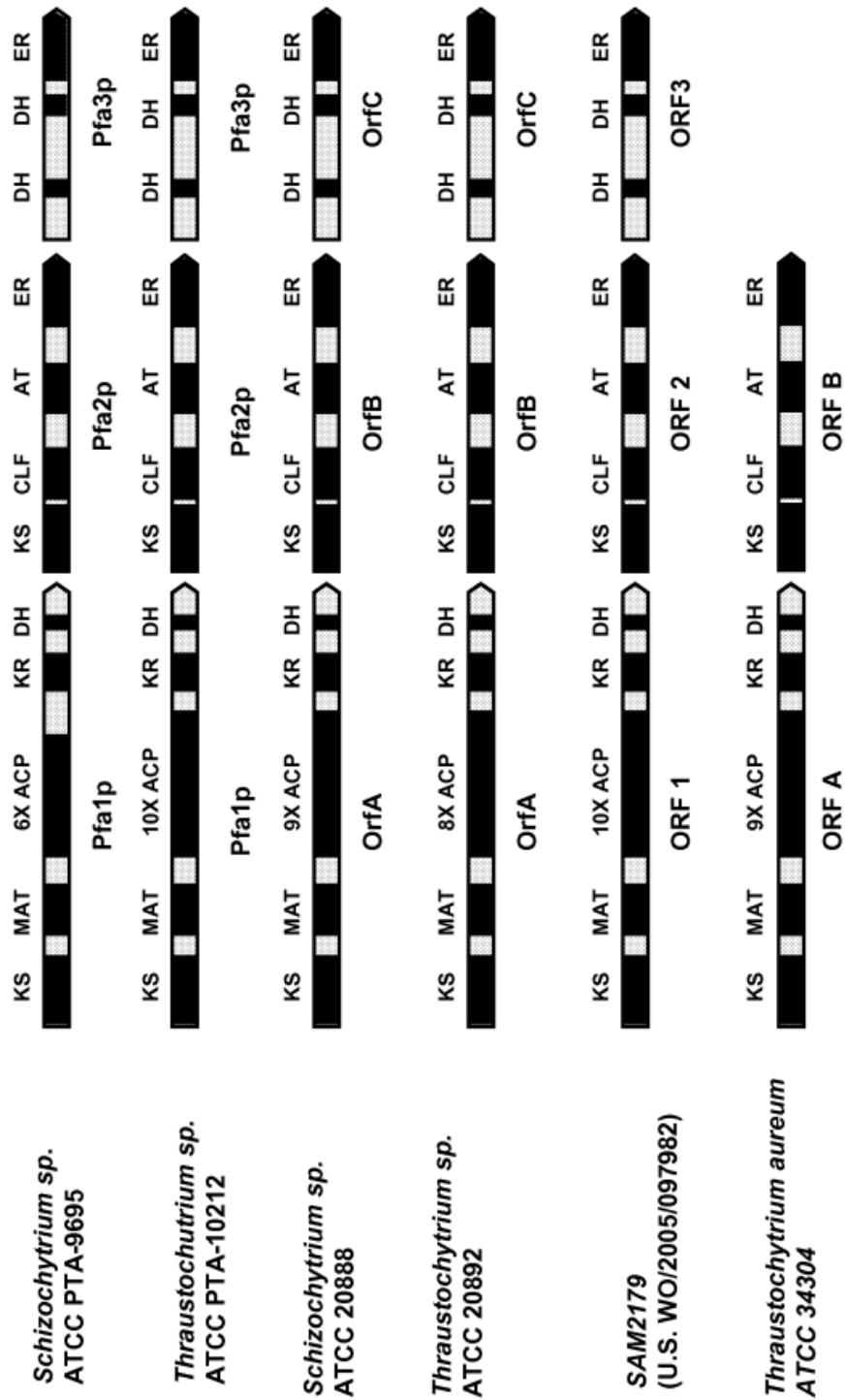


FIG. 4

	1	50
Sch. sp. 9695 Pfalp	(1)	-----
Thr. aureum ORF1	(1)	RKCIRPSLGHWHAIIGVLGRALRIVRPIRYEATNLRRLPRSGWLVALGLF
Sch. sp. 20888 OrfA	(1)	-----
Thr. sp. 10212 Pfalp	(1)	-----
Thr. sp. 20892 OrfA	(1)	-----
	51	100
Sch. sp. 9695 Pfalp	(1)	-----M
Thr. aureum ORF1	(51)	CDLSSCAGKLDLQTRDTAKDPCCCKRKWSASRAPPRPRAEADKASNEMETK
Sch. sp. 20888 OrfA	(1)	-----MAARLQEQKGGEM
Thr. sp. 10212 Pfalp	(1)	-----ME
Thr. sp. 20892 OrfA	(1)	-----MKDME
	101	150
Sch. sp. 9695 Pfalp	(2)	DTRIAIVGMSAILPSCENVRESWEAIRDGLDCLSDLPADRVDTVAYYHPE
Thr. aureum ORF1	(101)	DDRNAIVGMSAILPSCENVRESWEAIRDGLDCLSDLPADRVDTVAYYDPN
Sch. sp. 20888 OrfA	(14)	DTRIAIVGMSAILPSCGTTVRESWEAIRDGLDCLSDLPADRVDTVAYYDPV
Thr. sp. 10212 Pfalp	(3)	DQRIAIIVGMSAILPSCENVRESWEAIRDGLDCLSDLPADRVDTVAYYHPT
Thr. sp. 20892 OrfA	(6)	DRRNAIVGMSAHLPCGTDTVRESWEAIRDGLDCLSDLPADRVDTVAYYHNP
	151	200
Sch. sp. 9695 Pfalp	(52)	KTTKDKIYCKRGGFIPEYDFDAREFGLNMFQMEDSDANQTSLLKVKKEAL
Thr. aureum ORF1	(151)	KTTKDKIYCKRGGFIPEYDFDAREFGLNMFQMEDSDANQTSLLKVKKEAL
Sch. sp. 20888 OrfA	(64)	KTTKDKIYCKRGGFIPEYDFDAREFGLNMFQMEDSDANQTSLLKVKKEAL
Thr. sp. 10212 Pfalp	(53)	KGVKDKIYCKRGGFIPEYDFDAREFGLNMFQMEDSDANQTSLLKVKKEAL
Thr. sp. 20892 OrfA	(56)	KATKDKIYCKRGGFIPEYDFDAREFGLNMFQMEDSDANQTSLLKVKQKAL
	201	250
Sch. sp. 9695 Pfalp	(102)	TDANIPATPSGKKNIGCVLGIGGGQKASHEFYSSLNLYVVVEKVLKRMGLP
Thr. aureum ORF1	(201)	QDAGVAPETPKKKNIGCVLGIGGGQKASHEFYSSLNLYVVVEKVLKRMNLP
Sch. sp. 20888 OrfA	(114)	QDAGVAPETPKKKNIGCVLGIGGGQKASHEFYSSLNLYVVVEKVLKRMGNP
Thr. sp. 10212 Pfalp	(103)	TDANIPATPSGKKNIGCVLGIGGGQKASHEFYSSLNLYVVVEKVLKRMGLP
Thr. sp. 20892 OrfA	(106)	QDASVAPETPKKKNIGCVLGIGGGQKASHEFYSSLNLYVVVEKVLKRMGLP
	251	300
Sch. sp. 9695 Pfalp	(152)	EDVAAAVNKKYKASFPPEWRDLSFPGFLGNVTAGRCNTFFNMEGMNCVVDA
Thr. aureum ORF1	(251)	EDVVEAAVEKKYKAMFPEWRDLSFPGFLGNVTAGRCNTFFNMEGMNCVVDA
Sch. sp. 20888 OrfA	(164)	EDVVKVAVEKKYKAMFPEWRDLSFPGFLGNVTAGRCNTFFNMEGMNCVVDA
Thr. sp. 10212 Pfalp	(153)	EDDVETAVEKKYKAMFPEWRDLSFPGFLGNVTAGRCNTFFNMEGMNCVVDA
Thr. sp. 20892 OrfA	(156)	DAIVVEAAVEKKYKAMFPEWRDLSFPGFLGNVTAGRCNTFFNMEGMNCVVDA
	301	350
Sch. sp. 9695 Pfalp	(202)	ACASSLIAKVAITDELLIGDCDAMIAGATCTDNTSGMYMAFSKTPVFSTD
Thr. aureum ORF1	(301)	ACASSLIAKVAITDELLIGDCDAMIAGATCTDNTSGMYMAFSKTPVFSTD
Sch. sp. 20888 OrfA	(214)	ACASSLIAKVAITDELLIGDCDAMIAGATCTDNTSGMYMAFSKTPVFSTD
Thr. sp. 10212 Pfalp	(203)	ACASSLIAKVAITDELLIGDCDAMIAGATCTDNTSGMYMAFSKTPVFSTD
Thr. sp. 20892 OrfA	(206)	ACASSLIAKVAITDELLIGDCDAMIAGATCTDNTSGMYMAFSKTPVFSTD
	351	400
Sch. sp. 9695 Pfalp	(252)	QSVRAYDAATKGMLIGECSAMFVLKRYADAVRDGDTTHAVIRSCASSSDG
Thr. aureum ORF1	(351)	QSVRAYDAATKGMLIGECSAMFVLKRYADAVRDGDTTHAVIRSCASSSDG
Sch. sp. 20888 OrfA	(264)	QSVRAYDEKTKGMLIGECSAMFVLKRYADAVRDGDTTHAVIRSCASSSDG
Thr. sp. 10212 Pfalp	(253)	QSLAYDEKTKGMLIGECSAMFVLKRYADAVRDGDTTHAVIRSCASSSDG
Thr. sp. 20892 OrfA	(256)	QSVRAYDEKTKGMLIGECSAMFVLKRYADAVRDGDTTHAVIRSCASSSDG
	401	450
Sch. sp. 9695 Pfalp	(302)	KAAGIYPTTISGQEEALRAYARRANVDPSTVTLVEGHGTGTPVGDRIELT
Thr. aureum ORF1	(401)	KAAGIYAPTISGQEEALRAYARRAGVDPSTVTLVEGHGTGTPVGDRIELT
Sch. sp. 20888 OrfA	(314)	KAAGIYPTTISGQEEALRAYARRANVDPSTVTLVEGHGTGTPVGDRIELT
Thr. sp. 10212 Pfalp	(303)	KAAGIYPTTISGQEEALRAYARRAGVSPNTVTLVEGHGTGTPVGDRIELT
Thr. sp. 20892 OrfA	(306)	KAAGIYPTTISGQEEALRAYARRAGVCPSTVTLVEGHGTGTPVGDRIELT
	451	500
Sch. sp. 9695 Pfalp	(352)	ALSNLFSAKASANGGGAEEAEQAVGSIKSIQIHLKAVAGLAGLVKVVIA
Thr. aureum ORF1	(451)	ALSNLFDAANKGR-----KETAVGSIKSIQIHLKAVAGFAGLVKVVIA
Sch. sp. 20888 OrfA	(364)	ALSNLFDAKAVGEG-----NTEKAVGSIKSIQIHLKAVAGLAGLVKVVIA
Thr. sp. 10212 Pfalp	(353)	ALSNLFDAKAVGPG-----HKEKAVGSIKSIQIHLKAVAGCAGLVKVVIA
Thr. sp. 20892 OrfA	(356)	ALSNLFDAKAVGSK-----KEQAVGSIKSIQIHLKAVAGFAGLVKVVIA

FIG. 4 (cont)

	501	550
Sch. sp. 9695 Pfalp	(402) LKHKTLPQINVDKPPSLVDGTFPIQOSPLYNTMNRPWETPVGVPRRAGI	
Thr. aureum ORF1	(495) LKHKTLPQINVDHPPALHDGSPIDSSLYINTMNRPWETAPGVPRRAGI	
Sch. sp. 20888 OrfA	(409) LKHKTLPQINVDNPNLYDNTPINESSLYINTMNRPWETPPGVPRRAGI	
Thr. sp. 10212 Pfalp	(398) LKHKTLPQINVDNPNLYDNTPINESSLYINTMNRPWETKPGVPRRAGI	
Thr. sp. 20892 OrfA	(400) LKHKTLPQINVDQPLLDGTOIODSSLYINKTNRPWETONKLPRRAGI	
	551	600
Sch. sp. 9695 Pfalp	(452) SSFGFGGANYHAVLEEFEPHESESAYRYNNLPQVALLHAGDVATLAATVRA	
Thr. aureum ORF1	(545) SSFGFGGANYHAVLEEAEPHEAPYRINNOVPPQVLLHASSASALASICDA	
Sch. sp. 20888 OrfA	(459) SSFGFGGANYHAVLEEAEPHTTAYRLNKRPPQVLLMAATPAALOSLCEA	
Thr. sp. 10212 Pfalp	(448) SSFGFGGANYHAVLEEFEPQOTFPYRLNVSAPMLHAVNANSLOKICED	
Thr. sp. 20892 OrfA	(450) SSFGFGGANYHAVLEEFEPHEKPYRLNTVGHVLLYAPSEVALKVLND	
	601	650
Sch. sp. 9695 Pfalp	(502) KIALATAEQEEARVVKADYIAYHFFLDECKLPGAVPOAHARVGLLVRLDL	
Thr. aureum ORF1	(595) QADALQAAVSP-EASKHADYRAIVAFHEAFKLPACVPAGHARIGFVSGSA	
Sch. sp. 20888 OrfA	(509) OLKEFEAAIKENETVKNTAYIKCVKFGEOQKFPGLPATNARLGELVKDA	
Thr. sp. 10212 Pfalp	(498) OLKLLKESREKCVNTKNTDYVAFSKFQDSFKLKGSVPSOHARVGFASKSI	
Thr. sp. 20892 OrfA	(500) QLAELTIALEAKTHKNVDKVCYKFIDEFQOQSCPPENPRVGLFLATLP	
	651	700
Sch. sp. 9695 Pfalp	(552) SLIIVLEAAAKLAGESATENTVSVATGEAAFRVRGVATEENVAALFS	
Thr. aureum ORF1	(644) AATLAVLRAASAKLKQSSATLENTLLRE--GVTYRSAAMHTPGSVAALFA	
Sch. sp. 20888 OrfA	(559) EDACSTLRAICQFAKQVKEANRLPRE-GVSFRAKGIATN-GAVAALFS	
Thr. sp. 10212 Pfalp	(548) EDTISILSAIVNRFQKDIITTSWALPKE-GAIFRSTALINDNKSVAALFS	
Thr. sp. 20892 OrfA	(550) TSNIIVALKAILAQLDAKPDAKKWDLPKKKFGATFSSSVKGSVAALFA	
	701	750
Sch. sp. 9695 Pfalp	(602) GOGAQYTHMFSDVAMNWPFRESVAAMDRAQRERFG--SPAKRVSSVLYP	
Thr. aureum ORF1	(692) GOGAQYTHMFADVAMNWPFERSVQEMDAAQVTAAP----KRLSEVLYP	
Sch. sp. 20888 OrfA	(607) GOGAQYTHMFSDVAMNWPFROSLAAMDAAQSKVAGSDDFERVSQVLYP	
Thr. sp. 10212 Pfalp	(597) GOGAQYTHMFNDVAMNWPFRLCVNDMEKAQEEVIN-DSVKRISQVMEP	
Thr. sp. 20892 OrfA	(600) GOGTOYLNMFSDVAMNWPFRRDSIVAMEAQTEVFEG--QVEPISKVLEP	
	751	800
Sch. sp. 9695 Pfalp	(650) RKPYGDEPRDHEKISQTRYSQPATLACSVGAFDIFKAAGLAPSTFAAGHS	
Thr. aureum ORF1	(738) RKPYAAEPEQDNKATSMNTINSQPALMACAAGAFVFRQAGLAPDHVAGHS	
Sch. sp. 20888 OrfA	(657) RKPYEREPEQNPKKISITAYSQPSLACALGFRIIFKEAGFTPDFAAGHS	
Thr. sp. 10212 Pfalp	(646) RKPYARESPLDNKEISKTEYSQTTTVAASSVGLFELIFRDAGFAPAFVAGHS	
Thr. sp. 20892 OrfA	(648) RERYASESEQGNELLCITEYSQPTTIAAVGAFDIFKAAGFKPDMVGGHS	
	801	850
Sch. sp. 9695 Pfalp	(700) LGEFALYAAGSLDRDAVFDLVCARAKAMSDFTAQASSSGGAMAAVIGAK	
Thr. aureum ORF1	(788) LGEFGALLAAGCASREELFLVCSRAKAMQDVPGQ-----	
Sch. sp. 20888 OrfA	(707) LGEFALYAAGCVDRDELFLVCCRARIIMGGKDAP-ATPKGCMAAVIGEN	
Thr. sp. 10212 Pfalp	(696) LGEFSALYAAGLIDREDLFLVCNRAMMRDAPK--KQADGAMAAVIGEN	
Thr. sp. 20892 OrfA	(698) LGEFALYAAGSLSRBDLYLVCKRAKAMANASDG-----AMAAVIGPD	
	851	900
Sch. sp. 9695 Pfalp	(750) ADQSLGGAPDVHLANSNSPSQTVITGTAEVAAASDKLRCSGNFRVVEL	
Thr. aureum ORF1	(823) -----DG-----AWLANCNSPSQVVISGDKTNVERESSRLAGLG-FRIIPL	
Sch. sp. 20888 OrfA	(756) ENIKVQ--AANVVLGNSNSPSQTVITGSEVGIOAESARLQKEG-FRVVHL	
Thr. sp. 10212 Pfalp	(744) ASSIKLS-APEVAVANNNSPSQTVITGANSQVQETSCLKTQG-FRVVHL	
Thr. sp. 20892 OrfA	(742) ARLVTPQ-NSDVIVANFNSTATQVVISGTVOGVKEESKLLISKG-FRVVHL	
	901	950
Sch. sp. 9695 Pfalp	(800) ACEFAFHSPPHMRGAEQTFASALAQAPVSAPAAARFYNSVNTGGAAVTSPAD	
Thr. aureum ORF1	(863) ACEGAFHSPPHMTAAQATFOALDSIKISTPTNGARLYNNVSGKTCRSLGE	
Sch. sp. 20888 OrfA	(804) ACESAFHSPPQMNASSAFKDVISKVSFRTPKNE--TKLFSNVSGETYPTD	
Thr. sp. 10212 Pfalp	(792) ACEGAFHSPPHMTAAKQFOKALSAAKFNKPTQS--SPKIFSNVTGGVFTD	
Thr. sp. 20892 OrfA	(790) KCQGAHSPMLGPSHDSFKSLVETCTISPPKNVK-FFCNVSGKESP---N	
	951	1000
Sch. sp. 9695 Pfalp	(850) VRTNLGKHMTSPVQFVQOVRAMHAAGARVFVEFGPKQVLSPLVKETLGHA	
Thr. aureum ORF1	(913) LRDCLGKHMTSPVLFOAQVENMYAAGARVFVEFGPKQVLSKLVGETLADK	
Sch. sp. 20888 OrfA	(852) AREMLTQHMTSSVKFTQVRNMHQAAGARVFVEFGPKQVLSKLVSETLKDD	
Thr. sp. 10212 Pfalp	(840) PKTALSPPHMTSSVQFETQIKNMYAAGARVFVEFGPKQVLSKLVNEIFPGD	
Thr. sp. 20892 OrfA	(836) PKOTLKSMTSSVQFEEQIRNMYDAGARVFVEFGPKQVLSKLVLAENFP--	

FIG. 4 (cont)

		1001		1050
Sch. sp. 9695 Pfalp	(900)	GDVTVAVNEDSAKDSDTQLROAAETLAVAGUPLKDFDHWOLPDATRLLEP		
Thr. aureum ORF1	(963)	SDFVTVAVNSSSSKDSQOLREAAAKLAVLGVPANFDPWELCDARLRE		
Sch. sp. 20888 OrfA	(902)	PSVTVSVNPASGTDSDQLRDAAMQVAVAGUNLQGFDFWDAPDATRMOA		
Thr. sp. 10212 Pfalp	(890)	TSVLTVSVNPASAKDSDTQLRQAAVOMAVAGVALTDFDKWELKDPTRMKE		
Thr. sp. 20892 OrfA	(884)	-SCTAISVNPASSGDSQOLRLAAMKFAVSGAALSTFDPPWEYRKPDILLI		
		1051		1100
Sch. sp. 9695 Pfalp	(950)	VKKKKTTLRLSAATYVSAKTLRQREAVINDGYTVSGAT-----AV		
Thr. aureum ORF1	(1013)	CPRSKTTTLRLSAATYVSNKTLAAREKVMEDNCFSSIFASG-----P		
Sch. sp. 20888 OrfA	(952)	IKKKTTTLRLSAATYVSDKTKKVRDAAMNDGRCVTYIKGAAPLIKAPEPV		
Thr. sp. 10212 Pfalp	(940)	FPRKKTTLRLSAATYVSKKTLQERERIMNDGRTVSCVQ-----R		
Thr. sp. 20892 OrfA	(933)	RKPKKTALVLSAATYVSPKTLAERKKAMEDIKLVSIITPR-----D		
		1101		1150
Sch. sp. 9695 Pfalp	(990)	VKEVDTANEEPLVHOADLOROLAEASTAAQAAOSKVAELERTIQLERK		
Thr. aureum ORF1	(1055)	ASQEMERRIANIRAELEAAROLDTAKTQLARKQVQDPTADRQRDMIAKH		
Sch. sp. 20888 OrfA	(1002)	VDEAAKREAEKLOKELQDAAROLDKRAAAEANSKLAAKEEAKTANAS		
Thr. sp. 10212 Pfalp	(979)	LENTNTGELERLKKLODKENEVVRVQALATQASADLONTKAELOKQOAI		
Thr. sp. 20892 OrfA	(973)	SMVSGIAIAQEVRTAKQPLETEIRRLNKELEHLKRELAAKASVKSASKS		
		1151		1200
Sch. sp. 9695 Pfalp	(1040)	-----VQQQQQE-----		
Thr. aureum ORF1	(1105)	RSTLAMVKEFEALAS-----GSPCAVPPAPVVDTAEDVPFADKV		
Sch. sp. 20888 OrfA	(1052)	-AKPAVDTAIVEKHRAILKSMLELDGYGSVDASSLQQQQQQQTAPAPVK		
Thr. sp. 10212 Pfalp	(1029)	KSSNAASDAVIAKKHAILLAMLEELTGKAVDISSFSKGOVASPATVRVY		
Thr. sp. 20892 OrfA	(1023)	-----SKERSVLKSHRALLONILQYDLRVVPPAVRSVAVDNTAPYADQV		
		1201		1250
Sch. sp. 9695 Pfalp	(1047)	-----K-GEN--S-----		
Thr. aureum ORF1	(1146)	STPPPO-----VTSAPINELARAEAVVMEVLAAKTGYEVDHIEADM		
Sch. sp. 20888 OrfA	(1101)	AAAPAAP---VASAPAPAVSNELLEKAETVMEVLAAKTGYETDMIEADM		
Thr. sp. 10212 Pfalp	(1079)	SAPVQAA---APVQVSASVDSGLLAKAEQVLEVLASKTGYETELIELDM		
Thr. sp. 20892 OrfA	(1069)	STPASERSASPLFEKRSSVSSARLAEEAAAVLSVLADKTGYDSSMIEMDM		
		1251		1300
Sch. sp. 9695 Pfalp	(1052)	-----DSNAAAEVLRFHKEILLQRMLODCDEQAVPVAT		
Thr. aureum ORF1	(1187)	LLDAELGIDSVKRIEILAAVQAQLGVEAKDVDALSRTRTVGEVVDAMKAE		
Sch. sp. 20888 OrfA	(1148)	ELETELGIDSIKRVEILSEVOAMLNVEAKDVDALSRTRTVGEVVDAMKAE		
Thr. sp. 10212 Pfalp	(1126)	ELETELGIDSIKRVEILSEVOAQLNVEAKDVDALSRTRTVGEVIDAMKAE		
Thr. sp. 20892 OrfA	(1119)	DLESELGVDSIKRVEIMSEVQTLISVEVSDVDALSRTIKTVGDVIEAMKLE		
		1301		1350
Sch. sp. 9695 Pfalp	(1084)	VVPTPTSSPTPTSSPVSGN-----SKSTRGSADLOAILAKAETV		
Thr. aureum ORF1	(1237)	IGGOATSAPSMAQPOASAPSPSPASVLPKEVALPASVDPAKLARAEAV		
Sch. sp. 20888 OrfA	(1198)	IAGSSAPAPAAAAPAPAKA-----APAAAAPAVSNELLEKAETV		
Thr. sp. 10212 Pfalp	(1176)	IAGGQPAAPVQVAAPTQVV-----APVQASAPVDSGLLAKAEQV		
Thr. sp. 20892 OrfA	(1169)	LGGPQGQTLTAESIROPVSEPAVPTS---SSSIANVSSARLAEEAA		
		1351		1400
Sch. sp. 9695 Pfalp	(1123)	VMAVLAAKTGYEADMEVADMLEAELGIDSIKRVEILSEVOQLGVEAKD		
Thr. aureum ORF1	(1287)	VMEVLAAKTGYEVDHIEADMILLEAELGIDSVKRIEILAAVQAQLGVEAKD		
Sch. sp. 20888 OrfA	(1237)	VMEVLAAKTGYETDMIESDMLETELGLIDSIKRVEILSEVOAMLNVEAKD		
Thr. sp. 10212 Pfalp	(1215)	VLEVLAASKTGYETELIELDMLETELGLIDSIKRVEILSEVOAQLSVEAKD		
Thr. sp. 20892 OrfA	(1215)	VLSVLADKTGYDSSMIEMDMLESELGVDSIKRVEIMSEVQTLISVEVSD		
		1401		1450
Sch. sp. 9695 Pfalp	(1173)	VDALSRTRTVGEVVDAMKAEIVAA-----		
Thr. aureum ORF1	(1337)	VDALSRTRTVGEVVDAMKAEIGGOATSAPASVAQPOASAPSPSATTASVL		
Sch. sp. 20888 OrfA	(1287)	VDALSRTRTVGEVVDAMKAEIAGSSAPAPAAAAPGPAAN-----		
Thr. sp. 10212 Pfalp	(1265)	VDALSRTRTVGEVIDAMKAEIAGGQPAAPVQVAAPTQVVA-----		
Thr. sp. 20892 OrfA	(1265)	VDALSRTRTVGDVIEAMKLELGGPQGQTLTAESIROPVSEPAVPTS---		
		1451		1500
Sch. sp. 9695 Pfalp	(1198)	-----		
Thr. aureum ORF1	(1387)	PKPVAPTSADPAKLARAEAVVMEVLAAKTGYEVDHIEADMILLEAELGID		
Sch. sp. 20888 OrfA	(1327)	PAPAPAAPAVSNELLEKAETVMEVLAAKTGYETDMIESDMLETELGLID		
Thr. sp. 10212 Pfalp	(1305)	PVQASAP--VDSGLLAKAEQVLEVLAASKTGYETELIELDMLETELGLID		
Thr. sp. 20892 OrfA	(1312)	-S-SSSIANTSSARLAEEAAVLSVLADKTGYDSSMIEMDMLESELGVDS		

FIG. 4 (cont)

		1501		1550
Sch. sp. 9695 Pfallp	(1198)	-----	-----	-----GGSAPAV
Thr. aureum ORF1	(1437)	SVKPIEILAAVOAQLGVEAKDVDALSRTTRTVGEVVEAMKAEIIGGOATSAP		
Sch. sp. 20888 OrfA	(1377)	SIKRVEILSEVQAMLNVEAKDVDALSRTTRTVGEVVDAMKAEIAGGSAPAP		
Thr. sp. 10212 Pfallp	(1353)	SIKRVEILSEVQQLSVEAKDVDALSRTTRTVGEVIDAMKAEISGGQPAAP		
Thr. sp. 20892 OrfA	(1360)	SIKRVEIMSEVOTLLSVEVSDVDALSRTTRTVGEVIDAMKLELGGPQGOTL		
		1551		1600
Sch. sp. 9695 Pfallp	(1205)	PSAPAAASAAPTPAASTAPSA-----DLQALLSKAETVVMVLAIAKTGYEA		
Thr. aureum ORF1	(1487)	ASVAPQOISVSEPTPLAASPS-----ADPAKLARAEAVVMEVLAIAKTGYEV		
Sch. sp. 20888 OrfA	(1427)	AAAPAPAAAPAPAPAPAPAP-----VSSELLEKAETVVMVLAIAKTGYET		
Thr. sp. 10212 Pfallp	(1403)	VQVAAPTQIVAPVQVSAP-----VDSGLLAKAEQVLEVLASKTGYET		
Thr. sp. 20892 OrfA	(1410)	TAEIRQPPVSEPAVPTSSSSSIANVLSARLAEAEAAVLSVLADKTGYDS		
		1601		1650
Sch. sp. 9695 Pfallp	(1250)	DMVEADMILEAELGIDSIKRVEILSEVQQLGVEAKDVDALSRTTRTVGEV		
Thr. aureum ORF1	(1532)	DMIEADMLLDAELGIDSVKPIEILAAVOAQLGVEAKDVDALSRTTRTVGEV		
Sch. sp. 20888 OrfA	(1472)	DMIEADMELETELGIDSIKRVEILSEVQAMLNVEAKDVDALSRTTRTVGEV		
Thr. sp. 10212 Pfallp	(1446)	ELIELDMELETELGIDSIKRVEILSEVQQLSVEAKDVDALSRTTRTVGEV		
Thr. sp. 20892 OrfA	(1460)	SMEIEMDMLESELGVDSIKRVEIMSEVOTLLSVEVSDVDALSRTTRTVGDV		
		1651		1700
Sch. sp. 9695 Pfallp	(1300)	VDAMKAEIVAASAGSAP-----APAVPSAPAAASAAPTPAASTAPSAADLOAL		
Thr. aureum ORF1	(1582)	VDAMKAEIIGGOATSAPASVAPQOASAPSPSATASVLEKIVAPTSADPAK		
Sch. sp. 20888 OrfA	(1522)	VDAMKAEIAGGSAPAPAP-----AAAPAPAAAPAPAPAPAPAPAVSSEL		
Thr. sp. 10212 Pfallp	(1496)	IDAMKAEISGGQP-----TAPVQVAAPTQIVAPVQVSAPVDSGL		
Thr. sp. 20892 OrfA	(1510)	IEAMKLELGGPQG-Q---TLTAEIRQPPVSEPAVPTSSSSSIANVSSAR		
		1701		1750
Sch. sp. 9695 Pfallp	(1346)	LSKAETVVMVLAIAKTGYEADMVEADMILEAELGIDSIKRVEILSEVQGO		
Thr. aureum ORF1	(1632)	LARAEAVVMEVLAIAKTGYEADMIEADMLLDAELGIDSVKPIEILAAVOAQ		
Sch. sp. 20888 OrfA	(1568)	LEKAETVVMVLAIAKTGYETDMIESDMIELETELGIDSIKRVEILSEVQAM		
Thr. sp. 10212 Pfallp	(1535)	LAKAEQVLEVLASKTGYETELIELDMELETELGIDSIKRVEILSEVQAO		
Thr. sp. 20892 OrfA	(1556)	LAEAEAAVLSVLADKTGYDSMEIEMDMLESELGVDSIKRVEIMSEVOTL		
		1751		1800
Sch. sp. 9695 Pfallp	(1396)	LGVEAKDVDALSRTTRTVGEVVDAMKAEIVAASGGSAPAPAVPSAPAAASAA		
Thr. aureum ORF1	(1682)	LGVEAKDVDALSRTTRTVGEVVEAMKAEIIGGOAT-----SAPASMAQPOIS		
Sch. sp. 20888 OrfA	(1618)	LNVEAKDVDALSRTTRTVGEVVDAMKAEIAGSS-----ASAPAAAPAPAPAA		
Thr. sp. 10212 Pfallp	(1585)	LSVEAKDVDALSRTTRTVGEVIDAMKAEISGGQ-----PAAPVQVAAPTQI		
Thr. sp. 20892 OrfA	(1606)	LSVEVSDVDALSRTTRTVGEVIEAMKLELGGPQGOTLTSEPIHQPPVSEPA		
		1801		1850
Sch. sp. 9695 Pfallp	(1446)	PTPAATAPSAADLOALLAKAETVVMVLAIAKTGYEADMVEADMILEAELG		
Thr. aureum ORF1	(1727)	VSEPTPLAASPSADPAKLARAEAVVMEVLAIAKTGYEADMIEADMLLDAELG		
Sch. sp. 20888 OrfA	(1663)	APAPAPAAAPAVSNELLEKAETVVMVLAIAKTGYETDMIESDMIELETELG		
Thr. sp. 10212 Pfallp	(1630)	VAPVQASAP--VDSGLLAKAEQVLEVLASKTGYETELIELDMELETELG		
Thr. sp. 20892 OrfA	(1656)	VPTSSSSSIANVSSARLAEAEAAVLSVLADKTGYDSMEIEMDMLESELG		
		1851		1900
Sch. sp. 9695 Pfallp	(1496)	IDSIKRVEILSEVQQLGVEAKDVDALSRTTRTVGEVVDAMKAEIVAASAG		
Thr. aureum ORF1	(1777)	IDSVKPIEILAAVOAQLGVEAKDVDALSRTTRTVGEVVDAMKAEIIGGOATS		
Sch. sp. 20888 OrfA	(1713)	IDSIKRVEILSEVQAMLNVEAKDVDALSRTTRTVGEVVDAMKAEIAGGSAP		
Thr. sp. 10212 Pfallp	(1678)	IDSIKRVEILSEVQQLSVEAKDVDALSRTTRTVGEVIDAMKAEISGGQPA		
Thr. sp. 20892 OrfA	(1706)	VDSIKRVEIMSEVOTLLSVEVSDVDALSRTTRTVGEVIEAMKLELGGPQGO		
		1901		1950
Sch. sp. 9695 Pfallp	(1546)	SAPAPAVE-----		
Thr. aureum ORF1	(1827)	APASVAQF-----		
Sch. sp. 20888 OrfA	(1763)	APAAAP-----APAAAPAVSNELLEKAETVVMVLAIAKTGYETDMIES		
Thr. sp. 10212 Pfallp	(1728)	APVQVAAPTQIVAPVQVSAPVDSGLLAKAEQVLEVLASKTGYETELIEL		
Thr. sp. 20892 OrfA	(1756)	TLTAEIR-----		
		1951		2000
Sch. sp. 9695 Pfallp	(1554)	-----		
Thr. aureum ORF1	(1835)	-----QASAP-----		
Sch. sp. 20888 OrfA	(1808)	DMELETELGIDSIKRVEILSEVQAMLNVEAKDVDALSRTTRTVGEVVDAMK		
Thr. sp. 10212 Pfallp	(1778)	DMELETELGIDSIKRVEILSEVQQLNVEAKDVDALSRTTRTVGEVIDAMK		
Thr. sp. 20892 OrfA	(1764)	-----		

FIG. 4 (cont)

	2001	2050
Sch. sp. 9695 Pfallp (1554)	-----SAPASAPTPPASTAPSDLOALLSKAETVVMVAVLAA	
Thr. aureum ORF1 (1840)	-----SPSATASAPVTPLAPASVDPKLAARAAVMEVLAA	
Sch. sp. 20888 OrfA (1858)	AEIAGSSAPAPAAAPAPAAAPAPAAAPAVSEILLEKAETVVMVAVLAA	
Thr. sp. 10212 Pfallp (1828)	AEIAGGQP--AAPVQVAPAPVVRVQVSTPVDSEGLAKAEQVLEVLAC	
Thr. sp. 20892 OrfA (1764)	-----QPPVSEPAVPTSSSSSIINVSARLAEAEAAVLSVLAD	
	2051	2100
Sch. sp. 9695 Pfallp (1592)	KTGYEADMVEADMILEAELGIDSIKRVEILSEVOGQOLGVEAKDVDALSRT	
Thr. aureum ORF1 (1877)	KTGYEADMVEADMILEAELGIDSIKRVEILSEVOGQOLGVEAKDVDALSRT	
Sch. sp. 20888 OrfA (1908)	KTGYETDMIESDMLETELGLDSEIKRVEILSEVOAMLNVEAKDVDALSRT	
Thr. sp. 10212 Pfallp (1876)	KTGYETELIELDMLETELGLDSEIKRVEILSEVOQOLSVEAKDVDALSRT	
Thr. sp. 20892 OrfA (1802)	KTGYDSSMIEMDMLESELGVDSEIKRVEIMSEVOALLSVEVSDVDALSRT	
	2101	2150
Sch. sp. 9695 Pfallp (1642)	RTVGEVVDAMKAEIVASGGGAPAAVPSAPAAASA-----APTPATAPS	
Thr. aureum ORF1 (1927)	RTVGEVVDAMKAEIGGQTSAPASVQPOASAPSPSATASVLPKPVASPA	
Sch. sp. 20888 OrfA (1958)	RTVGEVVDAMKAEIAGGS-----APAPAAAP-----APAAAP	
Thr. sp. 10212 Pfallp (1926)	RTVGEVVDAMKAEISGGQ-----PTAPVQVAPPTQ-----VVAHVKVST	
Thr. sp. 20892 OrfA (1852)	KTVDVIEAMKMEIGGPOG----QTLTAESIREPPVSEPAVPTSSSSSI	
	2151	2200
Sch. sp. 9695 Pfallp (1686)	ADLOALLSKAETVVMVAVLAAKTGYEADMVEADMILEAELGIDSIKRVEIL	
Thr. aureum ORF1 (1977)	SVDPAKLARAAVMEVLAAKTGYEADMIDADMILLDAELGIDSIKRVEIL	
Sch. sp. 20888 OrfA (1992)	AVSNELLEKAETVVMVAVLAAKTGYETDMIESDMLETELGLDSEIKRVEIL	
Thr. sp. 10212 Pfallp (1965)	PVDSEGLAKAEQVLEVLASKTGYETELIELDMLETELGLDSEIKRVEIL	
Thr. sp. 20892 OrfA (1898)	NVSSARLAEAEAAVLSVLADKTGYDSSMIEMDMLESELGVDSEIKRVEIM	
	2201	2250
Sch. sp. 9695 Pfallp (1736)	SEVOGQOLGVEAKDVDALSRTRTVGEVVDAMKAEIVASGGGAPAAVPSAPA	
Thr. aureum ORF1 (2027)	AAVOAQLGVEAKDVDALSRTRTVGEVVEAMKAEIAGAGPNDQAAS----	
Sch. sp. 20888 OrfA (2042)	SEVOAMLNVEAKDVDALSRTRTVGEVVDAMKAEIAGGSAPAPAAAPASA	
Thr. sp. 10212 Pfallp (2015)	SEVOAQLNVEAKDVDALSRTRTVGEVIDAMKAEIAGDQAPAPVVPVQAKS	
Thr. sp. 20892 OrfA (1948)	SEVOTLLSVEVSDVDALSRTKTVDVIEAMKLELGESSIETLNCTEVEH	
	2251	2300
Sch. sp. 9695 Pfallp (1786)	LLPT-----	
Thr. aureum ORF1 (2073)	-----	
Sch. sp. 20888 OrfA (2092)	GAAP-----	
Thr. sp. 10212 Pfallp (2065)	GVANPALLAKAEQVLEVLASKTGYETELIELDMLETELGLDSEIKRVEI	
Thr. sp. 20892 OrfA (1998)	TSYKS-----	
	2301	2350
Sch. sp. 9695 Pfallp (1790)	-----	
Thr. aureum ORF1 (2073)	-----G-----	
Sch. sp. 20888 OrfA (2096)	-----	
Thr. sp. 10212 Pfallp (2115)	LSEVOAELSVEAKDVDALSRTRTVGEVIDAMKAEIAGSAVTATLDDSTI	
Thr. sp. 20892 OrfA (2003)	-----	
	2351	2400
Sch. sp. 9695 Pfallp (1790)	-----LPGSECEDLSTFPVITTEPLPAELVDAEGGARP	
Thr. aureum ORF1 (2074)	-----HLPGTGCEDLSLCSASVVEIARCSSELALERPMDRP	
Sch. sp. 20888 OrfA (2096)	-----AVKIDSVHGADCDLSLMHAKVVDIRRPDELALERPENRP	
Thr. sp. 10212 Pfallp (2165)	MEETDDEDEDLILYDHVYGSECEDLSTFSVKSIPRADKLILNIAERP	
Thr. sp. 20892 OrfA (2003)	-----VKASGCENVDTRFAKVVOISLPSKLSKSTVSHDRP	
	2401	2450
Sch. sp. 9695 Pfallp (1824)	VVVVDGSGALTSLLVSSIGDRAVLLOVQSSSACSPR-STTHKLVTADRS	
Thr. aureum ORF1 (2109)	ILIVSDGSGALPAALASRLGSCAVILITAGETDQSVR---STKHVDMEGWG	
Sch. sp. 20888 OrfA (2136)	VVVVDGSELTLALVVLGACAVLTTEGLQLAQAAGAAAIHVLAKDLS	
Thr. sp. 10212 Pfallp (2215)	IVIVDCGKLTTELAKAIGERAVVATESAQSLSVSPG--FVGKSFTLGNT	
Thr. sp. 20892 OrfA (2037)	VVVVDGSPLTTELCKILGGNIVVLSYQKPGAPPG-----VEVPDLS	
	2451	2500
Sch. sp. 9695 Pfallp (1873)	FAALQAALTSVEAQFGKVGGFVFQFG-DDDVQAQLGMAALLAKHLKTSLS	
Thr. aureum ORF1 (2156)	EADLVRALEAVESREGVPGGVVLERASETARDQLGFAILLAKHSSKALN	
Sch. sp. 20888 OrfA (2186)	AESAELAKAEAFREGALGGFISQAERFEPAEILGETIMCAKFAKASLC	
Thr. sp. 10212 Pfallp (2263)	ESEIERMSSIESSEYKIGGFVYQHFDSDYGMQLMALMAAKHLKESLN	
Thr. sp. 20892 OrfA (2080)	EEALIQALLALIRSTYGVPIGFICQVSNVSTKAQLCWALLAAKHLKDLN	

FIG. 4 (cont)

		2501		2550
Sch. sp. 9695 Pfalp	(1922)	EQTEGGRTFFVAVARLDGQLGLSG--K---STTATVDLSRAOQGSVFGLC		
Thr. aureum ORF1	(2206)	QOIPGGRACFVGVSRIDGKLGLSG--CAKKGWAAAEIAOQAVAGLC		
Sch. sp. 20888 OrfA	(2236)	TAVACGRPAFIGVARLDGRLGFT-----SOGTSDALKRAORGAIFGLC		
Thr. sp. 10212 Pfalp	(2313)	DPTEKNGRTFFLAARMNGHGLGMDN--ASVHDOQIVESCGLAERGAIFGLC		
Thr. sp. 20892 OrfA	(2130)	AVLPDSRSEFVGVVRLNGELGTFENISDFSKFDLTKALDYGORGELGLC		
		2551		2600
Sch. sp. 9695 Pfalp	(1967)	KTLDLEWFA--VFCRGIDLAADI DAAQAARCLLGELSDFDVAVRESGYSA		
Thr. aureum ORF1	(2254)	KTLDLEWPH--VFARSIDIELGANEETAAQAFEEELSCFDLTVREAGYTK		
Sch. sp. 20888 OrfA	(2279)	KTIGLEWSESDVFSRGVDIAQGNHPEDAATAVIREMACADIRIREVIGGA		
Thr. sp. 10212 Pfalp	(2361)	KTLDLEWPN--VFARGVDIAEGMSYSLAELVDEISCANISIRESGYTI		
Thr. sp. 20892 OrfA	(2180)	KSLDLEWEQ--VFCRGIDLACDIMPLOAARI RNELQCENMRIREVGYDI		
		2601		2650
Sch. sp. 9695 Pfalp	(2015)	SGQRCTTTTKSLTTGKPHQPTSSDLELVSGGARGITPLCVRELAQRGG		
Thr. aureum ORF1	(2302)	DGKRNTTEARFVGLGPKPKQAIRSSDVELVSGGARGITPVCVRELAKISG		
Sch. sp. 20888 OrfA	(2329)	NQQRCTIRAKLETGNPQRIAKDDVLLVSGGARGITPLCIREITRQIAG		
Thr. sp. 10212 Pfalp	(2409)	SGERTTEAHKLVGKPHAPKKKDAFLVSGGARGITPLCIREIAKAVKG		
Thr. sp. 20892 OrfA	(2228)	SGARTISTDDLLCGPSKAKVETADLELVSGGARGITPHCVREIASRSPG		
		2651		2700
Sch. sp. 9695 Pfalp	(2065)	GTIVLIGRSELPTTEPAWAVGVESGKPLEKAALAFKAEFAAGRGAKPTP		
Thr. aureum ORF1	(2352)	GTIVLIGRS--PLADSPAWACGV--EEANIGTAAMAHKAEFAAGRGPKPTP		
Sch. sp. 20888 OrfA	(2379)	GKYLIGRSKVSASEPAWCAGITDEKAVOKAATQELKRAFSAGEGPKPTP		
Thr. sp. 10212 Pfalp	(2459)	GTIVLIGRS--ALADEPLWANGK-SGKDLKAGLAFLKEEFAAGRGSKPTP		
Thr. sp. 20892 OrfA	(2278)	TTFVLIGRS--EMSEPDWAVGH--YNKDLDOSTMKHLKATHAAG-GVKPTP		
		2701		2750
Sch. sp. 9695 Pfalp	(2115)	MLHKLVGAVVGAREVPASLAETAGATAYESCDVSSAKVREMYVERV		
Thr. aureum ORF1	(2400)	KAHKAIVGSVLGAREVLGSLESIRAQGARAEYSCDVSCAEVKAIVDDI		
Sch. sp. 20888 OrfA	(2429)	RAVTKLVGSVLGAREVPSIAAIEALGGKATYSSCDVNSAADVAKAVRDA		
Thr. sp. 10212 Pfalp	(2507)	KVHKSILDKVLGIREVPASTANIEAHGAKATYSCDVSSAEKVAAVQKV		
Thr. sp. 20892 OrfA	(2325)	KAHKAIVNRVTGSRREVRESIRAIQEGANVEYIACDVSDENKVBQVQKV		
		2751		2800
Sch. sp. 9695 Pfalp	(2165)	QQQGRRVSGVGHASGVLDRDKLVENKSLADFSAVYDTKVGGLINLLACVD		
Thr. aureum ORF1	(2450)	ERRVG-AVTGVGHASGVLDRDKSVRLLEADFEVVGTKVDGLINLLQAVD		
Sch. sp. 20888 OrfA	(2479)	ESQLGARVSGIVHASGVLDRDLIEKKLPDEFDAVFGTKVTGLENLLAAVD		
Thr. sp. 10212 Pfalp	(2557)	EKEHLVRITGIVHASGVLDRDKLVENKTLDDFNAVYGTGKVTGLNLLSAVN		
Thr. sp. 20892 OrfA	(2375)	EOKYGCETIGIWHASGVLDRDKLVEQKTTDDFAVFGTKVTGLNIVSQVN		
		2801		2850
Sch. sp. 9695 Pfalp	(2215)	LNOLRHVLVLFSSLAGFHGNYGOSDYAMANEALNKIAAHL SAVHPOLCAHS		
Thr. aureum ORF1	(2499)	RPKLRLHLVLFSSLAGFHGNTGQAVYAMANEALNKNAFHLETAMFGLSVKT		
Sch. sp. 20888 OrfA	(2529)	RANIKHMLVLFSSLAGFHGNYGOSDYAMANEALNKGLELAKDVS---VKS		
Thr. sp. 10212 Pfalp	(2607)	KNFVRLHMLVLFSSLAGYHGNVGOSDYAMANEALNKIGFRLGAAYSOLCVKS		
Thr. sp. 20892 OrfA	(2425)	MSKLRHFLVLFSSLAGFHGNKGOTDYATANEALNKIAHTLSAFLPKINAKV		
		2851		2900
Sch. sp. 9695 Pfalp	(2265)	ICFGPWDG-GMVTIPALKANFIRMGTOTIIPROGGAQTVANMLVSSSPGOIL		
Thr. aureum ORF1	(2549)	IGFGPWDG-GMVNDALKAHFASMGVQIIPLDGGAETVSRITIGACSEPTOVL		
Sch. sp. 20888 OrfA	(2576)	ICFGPWDG-GMVTIPQLKQFOEMGVQIIPREGGAETVARIVLGSSPAEIL		
Thr. sp. 10212 Pfalp	(2657)	ICFGPWDG-GMVTIPALKKQFQSMGVQIIPREGGAETVARIVLSSNPSONL		
Thr. sp. 20892 OrfA	(2475)	LDFGPWVGSGMVTETLEKHFKAHMGVOTIPLEPGARTVAQIILASSPPOS		
		2901		2950
Sch. sp. 9695 Pfalp	(2314)	VGNWGVPPVPSATEHTVLOTLRQS-DNPFILDSHVIQGRVLPMTLAVGY		
Thr. aureum ORF1	(2598)	VGNWGLPPVPSASVHKITVRLGGESANPFLSSHTIQGRVLPMTLALGL		
Sch. sp. 20888 OrfA	(2625)	VGNWRTPSKKVGSDEITLHRKISAK-SNPFLEDHVIQGRVLPMTLALGS		
Thr. sp. 10212 Pfalp	(2706)	VGNWGVPPVPSLKSATVOTFTPE-LNPFILSHQIHGKNVLPMTLAVGY		
Thr. sp. 20892 OrfA	(2525)	LGNWGFPAKPLQRSNVITGTLSP-EIEFTADHKIQGRVLPMTLMAAIG		
		2951		3000
Sch. sp. 9695 Pfalp	(2363)	MAHQQSTVAGHQLWAVEDAQLFKGLAIDNGADVPRVELSRRKEQEDA		
Thr. aureum ORF1	(2648)	LAEARGLYVGHQVVGIEDAQVFGVLDKGTACEVQLRRESSTASP---		
Sch. sp. 20888 OrfA	(2674)	LAETCLGLFPGLSLWAIDBAQLFKGVTVDGDNCEVTLTPSPAPSG----		
Thr. sp. 10212 Pfalp	(2755)	LAHLVKNFYAGHHLNGVEDAQLFSGVVIDHVAQAVKLTEQSLDDG---		
Thr. sp. 20892 OrfA	(2574)	MASIEGLYPGYNLQGVNAQLFOGLTINQETKFOITLIEHNSEN---		

FIG. 4 (cont)

		3001		3050
Sch. sp. 9695 Pfalp	(2413)	GKVKKQVLLKKSQVN-GKSVPAYKATVVL-S-PAPPPSVITDFDLT---		
Thr. aureum ORF1	(2695)	---SEVVLASLNVFAGKVVPAYRAHVVLGASGPRTGGOLELKDLGVD		
Sch. sp. 20888 OrfA	(2720)	---KVNVCITIKTFSS-GKLVPAVRAVIVLSNQGAPPANATMQPPSL---		
Thr. sp. 10212 Pfalp	(2802)	---KVKVQVLTASNDNGKVPAYKAVIVLG-KTSRPAFLLKDFSLQ---		
Thr. sp. 20892 OrfA	(2621)	----LDVLTSLGVMLESGKVLPAVRCVCLNTTQOQPKLSPIILNLEVDP		
		3051		3100
Sch. sp. 9695 Pfalp	(2458)	--PDPACTEHDLYDGKTLFHGKAFQSGIEQVLSATPKQLTAKCRNEPLTPE		
Thr. aureum ORF1	(2742)	ADPACSVGKGALYDGKTLFHGPAFQYNDDEVLRCSPAELAVRCRVVPSAAQ		
Sch. sp. 20888 OrfA	(2763)	--DADPALQGSYIDGKTLFHGPAFRGIDDVLSCTKSQLVAKCSAVPGSDA		
Thr. sp. 10212 Pfalp	(2845)	--ESNRSADFLYDGKTLFHGGLFRGKTLNVS DTSLTQCTNEDLTAT		
Thr. sp. 20892 OrfA	(2667)	-----ACEVNPYDGKSLFHGGLLQFVQOVLHSS TKGLVAKCRALPIKEA		
		3101		3150
Sch. sp. 9695 Pfalp	(2506)	QRGQFVVNLSQODPFQADLAFOAMLVWARMLRQSAALPNNCERFDFTKPM		
Thr. aureum ORF1	(2792)	DRGQFVSRGVLYDPFINDTVFOALLVWARLVADSASLPSNVERISFHGQP		
Sch. sp. 20888 OrfA	(2811)	ARGEFATDTDADDPFVNDLAFOAMLVWVRRTLGOALPNSLORIVQHRPV		
Thr. sp. 10212 Pfalp	(2893)	ER-GQFADIEPVNPFMADAAFOAMLVWVRNLNSASLPNNCERVDIYKPI		
Thr. sp. 20892 OrfA	(2711)	IR-GPFIQTLADPIEDDIFOLMLVWCNRLGASLPNRIEKMSYFGNV		
		3151		3200
Sch. sp. 9695 Pfalp	(2556)	AP-GATYYTSVALASASP-LVDSVCKCTVAMHDEQGEVFSARASVVLNK		
Thr. aureum ORF1	(2842)	PSEGVEFYITLKLDSASGPLDPIAKAQFFLHRACGAVFASGRASVVLNK		
Sch. sp. 20888 OrfA	(2861)	PQ-DKPFYITLESNOGSG---GHSQHKHALQFHNEQGDLLIDVQASVQATD		
Thr. sp. 10212 Pfalp	(2942)	AP-GEKYITTLQALGNT---SGSVLKSVEFYMHDEQGEVELSGRASVVLND		
Thr. sp. 20892 OrfA	(2760)	SE-GSTEEASVTPVGPR-VPKDPVIKMOFLLODESGNTFSSGEGSVVLSD		
		3201		
Sch. sp. 9695 Pfalp	(2604)	TLTY-		
Thr. aureum ORF1	(2892)	ALSF-		
Sch. sp. 20888 OrfA	(2907)	SLAF-		
Thr. sp. 10212 Pfalp	(2988)	KMEF-		
Thr. sp. 20892 OrfA	(2808)	ELVF-		

FIG. 5

	1	50
Sch. sp. 9695 Pfa2p	(1)	-----MPCDN
Sch. sp. 20888 OrfB	(1)	-----MAARNVSAAHMHDKPT
Thr. sp. 10212 Pfa2p	(1)	-----MVKLSVGDNICH
Thr. aureum ORF2	(1)	QAIGHRAARWSCRSKSKARGHKAQKEMNQGGRNDEGVSVARADPCF
Thr. sp. 20892 OrfB	(1)	-----MQLPPAHSADNPI
	51	100
Sch. sp. 9695 Pfa2p	(7)	AVVGMAVOYAGCNDQEFWNTLMRKEINSSPISAERLCNRPFLHHPOR
Sch. sp. 20888 OrfB	(19)	AVVGMAVOYAGCNDQEFWNTLMRKEINSSPISAERLCNRPFLHHPOR
Thr. sp. 10212 Pfa2p	(17)	AVVGMAVYAGCNDQEFWNTLMRKEINSSPISAERLCNRPFLHHPOR
Thr. aureum ORF2	(51)	AVVGMAVEYAGCNDQEFWNTLMRKEINSSPISAERLCNRPFLHHPOR
Thr. sp. 20892 OrfB	(15)	AVVGMAVKYAGCNDQEFWNTLMRKEINSSPISAERLCNRPFLHHPOR
	101	150
Sch. sp. 9695 Pfa2p	(57)	SKYADTFCNERYGCTDASVDN-EHLLIADLANRAILLDGIN--LDDASTT
Sch. sp. 20888 OrfB	(69)	SKYADTFCNERYGCTDASVDN-EHLLIADLANRAILLDGIN--LDDASTT
Thr. sp. 10212 Pfa2p	(67)	SKYADTFCNERYGCTDASVDN-EHLLIADLANRAILLDGIN--LDDASTT
Thr. aureum ORF2	(101)	SKYADTFCNERYGCTDASVDN-EHLLIADLANRAILLDGIN--LDDASTT
Thr. sp. 20892 OrfB	(65)	SKYADTFCNERYGCTDASVDN-EHLLIADLANRAILLDGIN--LDDASTT
	151	200
Sch. sp. 9695 Pfa2p	(104)	ANRDFGIVSGCLSPMDNLOGELLNLYQVHYENVGAQRFEDNR-PWSE
Sch. sp. 20888 OrfB	(109)	KDSTKCGIVSGCLSPMDNLOGELLNLYQVHYENVGAQRFEDNR-PWSE
Thr. sp. 10212 Pfa2p	(107)	IDLNTGIVSGCLSPMDNLOGELLNLYQVHYENVGAQRFEDNR-PWSE
Thr. aureum ORF2	(150)	AOLKTCGIVSGCLSPMDNLOGELLNLYQVHYENVGAQRFEDNR-PWST
Thr. sp. 20892 OrfB	(113)	FDLENTGIVSGCLSPMDNLOGELLNLYQVHYENVGAQRFEDNR-PWST
	201	250
Sch. sp. 9695 Pfa2p	(153)	RPEVSVSEASDPVVSDFPASFVANGLGLGPRYSILDAACASALYVRLAQ
Sch. sp. 20888 OrfB	(158)	REQNKEAGDRIIFMDPASFVAELNLGALHYSVDAACATYVRLAQ
Thr. sp. 10212 Pfa2p	(156)	RTTNG---KDDNSAFDPASFVAELNLGALHYSVDAACATYVRLAQ
Thr. aureum ORF2	(199)	RTNHLHLPQDPFTHRDPAFVAGQLGLGPRYSILDAACASALYVRLAQ
Thr. sp. 20892 OrfB	(163)	ROKSTKAHAGDRIIFMDPASFVAELNLGALHYSVDAACATYVRLAQ
	251	300
Sch. sp. 9695 Pfa2p	(203)	DHLSGSAADMLCGATCPEPFFFILSGFSTFQAMPVQPPDDNPSNPLRQ
Sch. sp. 20888 OrfB	(208)	DHLSGSAADMLCGATCPEPFFFILSGFSTFQAMPVQPPDDNPSNPLRQ
Thr. sp. 10212 Pfa2p	(203)	DHLSGSAADMLCGATCPEPFFFILSGFSTFQAMPVQPPDDNPSNPLRQ
Thr. aureum ORF2	(249)	DHLSGSAADMLCGATCPEPFFFILSGFSTFQAMPVQPPDDNPSNPLRQ
Thr. sp. 20892 OrfB	(213)	DHLSGSAADMLCGATCPEPFFFILSGFSTFQAMPVQPPDDNPSNPLRQ
	301	350
Sch. sp. 9695 Pfa2p	(253)	GSQGLTPGEGGSMVLKRLDADRDGDRYIGTLLGTSLSNAGCGPLSPH
Sch. sp. 20888 OrfB	(256)	DSQGLTPGEGGSMVLKRLDADRDGDRYIGTLLGTSLSNAGCGPLSPH
Thr. sp. 10212 Pfa2p	(249)	TSQGLTPGEGGSMVLKRLDADRDGDRYIGTLLGTSLSNAGCGPLSPH
Thr. aureum ORF2	(296)	DSQGLTPGEGGSMVLKRLDADRDGDRYIGTLLGTSLSNAGCGPLSPH
Thr. sp. 20892 OrfB	(260)	TSQGLTPGEGGSMVLKRLDADRDGDRYIGTLLGTSLSNAGCGPLSPH
	351	400
Sch. sp. 9695 Pfa2p	(303)	LPSEKSCNEHLTSVGTDE-SEIYVECHATGTPOGDVVEVIAVSHCFRG
Sch. sp. 20888 OrfB	(306)	LPSEKSCNEHLTSVGTDE-SEIYVECHATGTPOGDVVEVIAVSHCFRG
Thr. sp. 10212 Pfa2p	(299)	NPSEFDCNEKALQVHRLP-SSIYVECHATGTPOGDVVEVIAVSHCFRG
Thr. aureum ORF2	(346)	QPSEESCNEHLTSVGTDE-SEIYVECHATGTPOGDVVEVIAVSHCFRG
Thr. sp. 20892 OrfB	(310)	LPSEESCNEHLTSVGTDE-SEIYVECHATGTPOGDVVEVIAVSHCFRG
	401	450
Sch. sp. 9695 Pfa2p	(352)	NTDHPPRRGSGTKGNFGHTLVAAGFAGMAKVLISMONGTIPPTPGNDRSN
Sch. sp. 20888 OrfB	(355)	---KTPRFGSGTKGNFGHTLVAAGFAGMAKVLISMONGTIPPTPGNDRSN
Thr. sp. 10212 Pfa2p	(348)	---HTPRFGSGTKGNFGHTLVAAGFAGMAKVLISMONGTIPPTPGNDRSN
Thr. aureum ORF2	(395)	---ASPRFGSGTKGNFGHTLVAAGFAGMAKVLISMONGTIPPTPGNDRSN
Thr. sp. 20892 OrfB	(360)	---NTPRFGSGTKGNFGHTLVAAGFAGMAKVLISMONGTIPPTPGNDRSN
	451	500
Sch. sp. 9695 Pfa2p	(401)	CDPELVVDEATNPYSSQANAGKPGDELKCASLSAFGFGGTNAHVFEF
Sch. sp. 20888 OrfB	(402)	MDPLVVSGETNPNTNPEPKK-----AGLSAFGFGGTNAHVFEF
Thr. sp. 10212 Pfa2p	(395)	INHDLVVVETNPNTNPEPKK-----AGLSAFGFGGTNAHVFEF
Thr. aureum ORF2	(441)	QIDPELVVTAENPNTNPEPKK-----AGLSAFGFGGTNAHVFEF
Thr. sp. 20892 OrfB	(406)	QASEHVVTATNPNTNPEPKK-----AGLSAFGFGGTNAHVFEF

FIG. 5 (cont)

	501	550
Sch. sp. 9695 Pfa2p	(451)	HRQIAATATASPVLP-----EVTGPPTAIIGMDATFGTLKGLDAFEQA
Sch. sp. 20888 OrfB	(443)	HDPSNACTGHDSISALSARCGGSENMFAITGMDATFGALKGLDAFERA
Thr. sp. 10212 Pfa2p	(436)	YRSDLOANKTLENE--S-KSHEIFSFKLAIVGMESEFGTLKGLQEFERA
Thr. aureum ORF2	(482)	HIPSRAAPPVAVLCQPRLG-----SGPNRKLAIIGMDATFGSLKGLSALEAA
Thr. sp. 20892 OrfB	(447)	FNAEGISYRPGKPP-----VESNRPSIVITGMDCTFGSLEGIDAFETA
	551	600
Sch. sp. 9695 Pfa2p	(494)	IYKGTDCASDLPSKRWRFLGADTDFLTAMGLDAVPRGCVYRDVDVDYKRL
Sch. sp. 20888 OrfB	(493)	IYTGANGAIPLEKRWFLGKDKDFLDLCGVKATPHGCIYEDVEVDFORL
Thr. sp. 10212 Pfa2p	(483)	IYNGGHGACDLPENRWFLGEDKEFLQACGLQKLPBGCIYKEVETDFKRL
Thr. aureum ORF2	(527)	IYERPHARPLPAKRWFLGGDESFLHEITGLECSPHGCIYEDVDVDFKRL
Thr. sp. 20892 OrfB	(491)	IYEGRDAARDLPKRWFLGEDLFLRAIRLKEKPGCEVESVDVDFKRL
	601	650
Sch. sp. 9695 Pfa2p	(544)	RSPMIPEDVLRPOQLLAVATMDRAIDAGLGGKQVAVLVGLGTDTELYR
Sch. sp. 20888 OrfB	(543)	RTPMTPEDMRLPOQLLAVETIDRAIDSGMKKGGNVAVFVGLGTDTELYR
Thr. sp. 10212 Pfa2p	(533)	RLPMIQEDILRPLQLLAVSIIDRAINASGVKPNQKQVAVLVGLGTDTELYR
Thr. aureum ORF2	(577)	RTPMYPEDILRPOQLLAVSTIDRAIDSGGLAKGGNVAVFVGLGTDTELYR
Thr. sp. 20892 OrfB	(541)	KIPITPEDMRLRPOQLLAVSTMDRAIDAGLKGQHVAVLVGLGTDTELYR
	651	700
Sch. sp. 9695 Pfa2p	(594)	HRARVTLKERLDP---AAFSPEQVQEMMDYINDCGTSTSYTSYIGNLVAT
Sch. sp. 20888 OrfB	(593)	HRARVALKERVR----PEASKKLNMDYINDCGTSTSYTSYIGNLVAT
Thr. sp. 10212 Pfa2p	(583)	HRARVALKERLQTA---KEDIPLLEKLMYINDRGTSYTSYIGNLVAT
Thr. aureum ORF2	(627)	HRARVALKERLQGLVRSAGGALTSRLMYINDSGTSTSYTSYIGNLVAT
Thr. sp. 20892 OrfB	(591)	HRARVALKEVLHPSL--KSDTATLOKIMYINDAGTSTSYTSYIGNLVAT
	701	750
Sch. sp. 9695 Pfa2p	(641)	RVSSQWNGFTGPSFTVTEGANSVYRCLELKGKFIIDTHQVDAVVVAGVDLCG
Sch. sp. 20888 OrfB	(638)	RVSSQWNGFTGPSFTITEGNNSVYRCLELKGKFIILETGEVDGVVVAGVDLCG
Thr. sp. 10212 Pfa2p	(631)	RVSSQWNGFTGPSFTITEGANSVYRCLELQWFLANGEVDAVVVAGVDLCG
Thr. aureum ORF2	(677)	RVSSQWNGFTGPSFTVTEGANSVYRCLELQAKYMLERGEVDAVVVAGVDLCG
Thr. sp. 20892 OrfB	(639)	RVSSQWNGFTGPSFTVTEGNNSVYRCLELQAKDMLQVNRVDAVVVAGVDLCG
	751	800
Sch. sp. 9695 Pfa2p	(691)	SAENLYLKARRSAISRQDHPRANFEASADGYFAGEGSGALVLRQAIVGS
Sch. sp. 20888 OrfB	(688)	SAENLYVKSRRFVSTSDTPRASFDAAADGYFVGECCGAFVLRKETSCTK
Thr. sp. 10212 Pfa2p	(681)	SAENLYVKSRRSKVSTQNEPFANFESNADGYFAGDGCAGLVLRKLSQTD
Thr. aureum ORF2	(727)	SAEAFVVRSRMQISKSORPAAPFDRADGFFAGEGCGALVFKRLIDCVS
Thr. sp. 20892 OrfB	(689)	SAESFFVRANRQKISKLSHPCASFDRDADGFFAGEGCGALVFKRLIDCAP
	801	850
Sch. sp. 9695 Pfa2p	(741)	-DDKYASVAGITCAQPAEAVSPLILQVHNDNNEKRVVEMVELAADSSE
Sch. sp. 20888 OrfB	(738)	-DDRIYACMDAIVPQNVPSACLREALDQARVKPGD---IEMLELSADSAR
Thr. sp. 10212 Pfa2p	(731)	STEKIYATVDSIAVGVDEVGPTIKQALKNASTAAKD---IELAELSASSSK
Thr. aureum ORF2	(777)	-GERIYASLDSVVVATTPRALRAAGSNRVDPAS---IDMVLSADSHR
Thr. sp. 20892 OrfB	(739)	-QEKIYASIDSIADKPTSSAVKAYQSDSSLSL---IELLEISGDSKR
	851	900
Sch. sp. 9695 Pfa2p	(790)	HAPHLANSPLSAESQLEQVSKLLAHQVPG----SVAIGSVRANVGDVGYA
Sch. sp. 20888 OrfB	(784)	HLKDPSVLPKELTAEIEIGGLQTIIRDDDKLPRNVATGSVKATVGDVGYA
Thr. sp. 10212 Pfa2p	(778)	HHSGRITCEDELNELGEIFNEG-----IQRVAIGSVKANVGDVGYA
Thr. aureum ORF2	(823)	FVRAPGTVAQPLTAEVEVGAVREVI GTAGRGRSVAAGSVRANVGDAGEA
Thr. sp. 20892 OrfB	(785)	FAAFEGAVEIQSSVEAQLKGLSKVLEPAK--GGQVAVGSTRATVGDVGYA
	901	950
Sch. sp. 9695 Pfa2p	(836)	SGAASLIKALCLNRYLP---AMPPOWERPVAPVSEALETCPRSRWLKN
Sch. sp. 20888 OrfB	(834)	SGAASLIKAALCTYNRYLPNSGDDNDEPAEPAPDSTLEACQTSRAWLKN
Thr. sp. 10212 Pfa2p	(819)	SGAASLIKALCLNRYLP-KLPNRNKPTKDVERSKSFFVCEHSRAWLKN
Thr. aureum ORF2	(873)	SGAALIKALCLNRYLA-ATPGWDAPAAGVDGAGELYVCRESRAWLKN
Thr. sp. 20892 OrfB	(833)	TGAASLIKALCLNRYLP-ALANWSGICEQSAAGSNHFEVCHETRPWLKN
	951	1000
Sch. sp. 9695 Pfa2p	(883)	PGESRLAAVASASESGSCFGVLLTDEYATHESNPLSLDDAAPKLIAIRG
Sch. sp. 20888 OrfB	(884)	PGERRIARVSGVSETRSCYSVLLSEAECHYERENRISLDEEAPKLIVIRA
Thr. sp. 10212 Pfa2p	(868)	VDENRHAIVSGVCENGSCYGVVMSDVQGHHEESNLVSLDKNEPKVLGYG
Thr. aureum ORF2	(922)	AGVARHAAISGVDEGGSCYGLVLSVPGQYETGNRISLQAESPKLILLISA
Thr. sp. 20892 OrfB	(882)	QNEKRCALISGTDPSHTCFSLVLSDTG-CYEEHNRTCFDVOAPOLVLIHG

FIG. 5 (cont)

	1001		1050
Sch. sp. 9695 Pfa2p	(933)	DTVDDIMAKNNELALLRAHAETGSATDDDDPAAVAFTAHLRLRFLRLVGE	
Sch. sp. 20888 OrfB	(934)	DSHEEILGRDKIRERFLQPTGAPRESEKQARRIFLE--LIGETLAQ	
Thr. sp. 10212 Pfa2p	(918)	DSVDDILVOLNKYLEKFLQETGTAAPAAQKVKEPTIDIDSN--VFAEMINL	
Thr. aureum ORF2	(972)	PDHAALIDKVAEIAAEQADGLSAPAAAADR-----LIGESIVG	
Thr. sp. 20892 OrfB	(931)	FDGKTIIRPLEEYILLEVEGHASPS-----EYFHK--LIGOSILE	
	1051		1100
Sch. sp. 9695 Pfa2p	(983)	TVASHGATATLCLALLITPEKLEKELELAAGVPRSAKAGFNWMSPSGSA	
Sch. sp. 20888 OrfB	(982)	DAASSGSQKPLAISLVSTPSKLOREVELAAGIPRCIKMFRDWSSPAGSR	
Thr. sp. 10212 Pfa2p	(966)	PODKN---KKFAVALVTTPNKLOREIELAVEGIPRCVKAARDWCSPSGSI	
Thr. aureum ORF2	(1012)	CAAGS---GGITICLVASPASLHKEALAHGIPRCIKARRDWASPAAGSY	
Thr. sp. 20892 OrfB	(969)	NSKES---K-LTSLVLCNPNQLOKELMLAIGVQSRMLTGKDWVSPSGSC	
	1101		1150
Sch. sp. 9695 Pfa2p	(1033)	EAPTPVTSDRVAFMYGEGRSPPYGVGLDLHRIWPLHERINDETAALWEN	
Sch. sp. 20888 OrfB	(1032)	YAPEPLASDRVAFMYGEGRSPPYGITQDIHRIWPELHEVINETNRLWAE	
Thr. sp. 10212 Pfa2p	(1013)	EACNPLKSDNIAFMYGEGRSPPYAGLGYDLHRIWPLHELVNNETTELWDQ	
Thr. aureum ORF2	(1059)	EAPEPLASDRVAFMYGEGRSPPYCGVGRDLHRIWPLHERVNAKTVNLVGE	
Thr. sp. 20892 OrfB	(1015)	EAPNPLSSAKVAFMYGEGRSPPYCGVGLGLHRIWPLHENYNNKTVDLWTE	
	1151		1200
Sch. sp. 9695 Pfa2p	(1083)	GDSWLMIPRAVDASQRAVQTAFDADQIEMFRTGIFVSIQCLTDYARDVLGV	
Sch. sp. 20888 OrfB	(1082)	GDRWVMPRASFKSELESQQQEFDRNMIEMFRLGILTSIAFTNLARDVLNI	
Thr. sp. 10212 Pfa2p	(1063)	GDSWYLPRSSVAEKEKVFQDFDKNOIEMFRLGIFVSMCFTDMATELLGL	
Thr. aureum ORF2	(1109)	GDWLLPRATSABEEBQICRNFDNOVEMFRTGVYISMCLTDLARSLIGL	
Thr. sp. 20892 OrfB	(1065)	GDGWLYPRTLREBHTKAIESFNANOIEMFRAGIFISMCOTDYVMNVLVG	
	1201		1250
Sch. sp. 9695 Pfa2p	(1133)	QPKACFGLSLGEISMLFALSRRNCGLSQDOLTORLFTSEPVWSTQLAVEFOA	
Sch. sp. 20888 OrfB	(1132)	TPKAAFGLSLGEISMLFASFKNGLISDOLTDLRESDVWNKALAVEFNA	
Thr. sp. 10212 Pfa2p	(1113)	TPKAAFGLSLGEISMLFASFKNNTLSKELTPRLKAKVWASQLAVEFAA	
Thr. aureum ORF2	(1159)	GPKASFGLSLGEISMLFALSSENCRLSEEMTRRLKASPVWNSQELAVEFNA	
Thr. sp. 20892 OrfB	(1115)	QPKAGFGLSLGEISMLFAMSKENCROSOEMTNRLGSPVWNSQELAINFNA	
	1251		1300
Sch. sp. 9695 Pfa2p	(1183)	LRKLWNVEADAPVESFWQGYIVRASRAETEKAIQPDNREVRLLIIVNDSQS	
Sch. sp. 20888 OrfB	(1182)	LREAWGIFQSVKDEFWQGYIVRGTKQDIEAIAIPSKYVRLTIINDANT	
Thr. sp. 10212 Pfa2p	(1163)	LRDLWNIPADKSIDFWQGYFVYANRTLVENTIG--ENKEVRLLIIVNDSQS	
Thr. aureum ORF2	(1209)	LRKLWGWAPGAPVDSFWQGYIVRATRAQVEQAIQEDNOEVRLLIIVNDSQS	
Thr. sp. 20892 OrfB	(1165)	LRKLWKIPRGAPLESFWQGYIVHGTRREEVEHAIQLEPVRLLIIVNDSRS	
	1301		1350
Sch. sp. 9695 Pfa2p	(1233)	ALINGKPAECLRVLEPLGGRLPPMPVKQGMIGHCEVAPYTPGIARIHEI	
Sch. sp. 20888 OrfB	(1232)	ALISGKPDACKAAIAPLGGNIPALPVTOGMCGHCBEVGPYTKDIKIHAN	
Thr. sp. 10212 Pfa2p	(1212)	CLINGKPDCEQKVIENLHLKLPAPVPTQGMIGHCEAIPYLDQISHIHEN	
Thr. aureum ORF2	(1259)	VLINGKPAACEAVIARIQSIILPPOVSOGMVGHCAEVLPTTSEIGRIHNM	
Thr. sp. 20892 OrfB	(1215)	ALINGKPDACQAVISRLNSKFPPLFVKQGMIGHCEVRAIKIDIGYIHET	
	1351		1400
Sch. sp. 9695 Pfa2p	(1283)	LEIPDSPVKMYTSVTNAELRG-----GNSSITEFVQKI	
Sch. sp. 20888 OrfB	(1282)	LEFPVVDGLDLWTTINQRLV-----PRATGAKDEWAPSSFGGYAQL	
Thr. sp. 10212 Pfa2p	(1262)	LEIPKPNVNLFTISENR-----ELVSMKDSVSKLVAEI	
Thr. aureum ORF2	(1309)	LRFPQSQETGGCKMYSSVNSRIGPVEESQMGPGTELVFSPSMEDEVAQL	
Thr. sp. 20892 OrfB	(1265)	LRISNDYSDCQLFSAVTKG-----ALDSSIMEIKHFVGEV	
	1401		1450
Sch. sp. 9695 Pfa2p	(1317)	YTRIADFPFGIVDKVSRDC--HDVFVEVGPNNMRSAAVSDILGKAATPHVSV	
Sch. sp. 20888 OrfB	(1325)	YEQANFPQIVETIYKQN--YDVFEVGPNNHRSTAVRTILGPQR-NHLAG	
Thr. sp. 10212 Pfa2p	(1296)	YQHVADFPNIYNKVKETCKTDIFIELGSNNYRSAAVTKILGP---EIVSV	
Thr. aureum ORF2	(1359)	YSPVADFPATTAAYVQQC--HDVFVEVGPDDHRSAAVRSILGPTTR-RHIAV	
Thr. sp. 20892 OrfB	(1300)	YSRIADFPQIVNTVHSAG--YDVFEELGCDASRSAAVQNLGGQG-KFLST	
	1451		1500
Sch. sp. 9695 Pfa2p	(1366)	AIDRPSAESANTQILKSLALLTAHRVPLHNPTLFADLYHPTFLTAIDSAMQ	
Sch. sp. 20888 OrfB	(1373)	AIDKQNEADAWTIVKIVASLKAHLVPGVTISPLYHSKIVAEQAQACYPALC	
Thr. sp. 10212 Pfa2p	(1343)	AIDRQNETAWGQIMKMWASLISHEVPGVELKKLYHPEILKDFDPAQKPNRF	
Thr. aureum ORF2	(1407)	AMDRKGESAWSQLFMELATLASHRVPGILDLSMYHPAVVERCRLALAAQR	
Thr. sp. 20892 OrfB	(1348)	AIDKKGHSAWSQVLPATASLAHRVPGISILDLPHNPFREMCCTMATT--	

FIG. 5 (cont)

		1501		1550
Sch. sp. 9695 Pfa2p	(1416)	E-PPPKPNRFLPSVEVNGYCPDGIKQVAAASAKPSTHCMRLHPAKAV		
Sch. sp. 20888 OrfB	(1423)	KGEKPKKPKFVRKQOLNCPFNKADPISSADLASFPADPAIEAAISSRI		
Thr. sp. 10212 Pfa2p	(1393)	IRNIELNGFFDRITNIVDIQLSPADPKLAEIINNRRNMPKDNVYVPIERVK		
Thr. aureum ORF2	(1457)	SGQPEQRNKKFLRTIEVNGFYDPADATIPAVATILPATAAISPPKLGAPH		
Thr. sp. 20892 OrfB	(1396)	---PKVEDKFLRTIQINGRFEKEMIHLEDTTLSCLPAPSEANIAAIOQRS		
		1551		1600
Sch. sp. 9695 Pfa2p	(1465)	VVAAGAVVADSTPVVKAKQTSSS-----		
Sch. sp. 20888 OrfB	(1473)	MKPVPAPKFYARLNIDEQDETRDPILNKDNAPSSSSSSSSSSSSSSSPSPA		
Thr. sp. 10212 Pfa2p	(1443)	TMIKAEPAANLOVSVGSKPVVTERISSDDNLFKEKLSEITKSFDG-----		
Thr. aureum ORF2	(1507)	DSQPEAEARPVGEASVPRRATSSSKLARTLAIDACDSDVRAALLDLDAPI		
Thr. sp. 20892 OrfB	(1443)	ERSAARSGSHDCASHSHEENKDCSPKELKLDVSVVAINFDN-----		
		1601		1650
Sch. sp. 9695 Pfa2p	(1489)	-----		
Sch. sp. 20888 OrfB	(1523)	PSAPVQKKAAPAAETKAVASADALRSALLDLSMLALSSASASGNLVETA		
Thr. sp. 10212 Pfa2p	(1486)	-----VN-----		
Thr. aureum ORF2	(1557)	AVG-----G-----		
Thr. sp. 20892 OrfB	(1486)	-----		
		1651		1700
Sch. sp. 9695 Pfa2p	(1489)	-----LLGDDAFILPCYDVDPFLYMGAMAEGISSVDLVVAAAN		
Sch. sp. 20888 OrfB	(1573)	PSDASVIVPPCNADLGSRAFMKTYGVSAFLYTGMAMKGIASADLVIAAG		
Thr. sp. 10212 Pfa2p	(1488)	-----ACTEAMIGDSG---FLKTYEVDYFLYTGMAMKGIASADLVIAAG		
Thr. aureum ORF2	(1561)	--SSRAQVPPCPMSALGSAAFRAAHGVDYALYMGAMAKGVASAEMVIAAG		
Thr. sp. 20892 OrfB	(1486)	-----DDRIOQLGHAGFREMYNTRYSLYTGMAMKGIASADLVIAAG		
		1701		1750
Sch. sp. 9695 Pfa2p	(1526)	EAPMLASTGAARLPMDQVELQIREIQORTSN-AFAVNLPGPDEAAT---		
Sch. sp. 20888 OrfB	(1623)	FOGILASFGAGGLPMQVVRSEIEKIQAAALPNGPYAVNLIHSPFDNLEKG		
Thr. sp. 10212 Pfa2p	(1529)	KSKILASFGAGGLAQVVEDAIKQIKAEIIGNPPFAVNLIHSPFDPSLEKG		
Thr. aureum ORF2	(1609)	KARMLASFGAGGLPLGEVEEALDKIQAAALPEGPPFAVNLIHSPFDPNLEEG		
Thr. sp. 20892 OrfB	(1526)	KEGILASFGAGGLPLATVRKGIKIQAAALSGPYAVNLIHSPFDGNLEQG		
		1751		1800
Sch. sp. 9695 Pfa2p	(1572)	-VDALLTGVMSVEASGYTGALSADVRYRVVTGLRRTSCGASVSATHRVV		
Sch. sp. 20888 OrfB	(1673)	NVDLFLEKGVTFVEASAFM-TLTTPQVVRVRAAGLTENAD-CSVNIRNRRII		
Thr. sp. 10212 Pfa2p	(1579)	NVDLFLEKYNVRFVEVSAM-TLTTPQVVRVRAAGLTENAD-CSVKIQNRRII		
Thr. aureum ORF2	(1659)	NVELFLRGRIRLVEASAFM-SVTTPSIVRYRVAGLERGPG-CTARVLNRVI		
Thr. sp. 20892 OrfB	(1576)	NVDLFLEKNRVVAECSAFT-TLTVPVHYRAAGLVRQD-CSILIKNRRII		
		1801		1850
Sch. sp. 9695 Pfa2p	(1621)	AKVSRTEVAEHFLRPAPAAVLEALVAAKOITPEQAALASVAMADDVAVE		
Sch. sp. 20888 OrfB	(1721)	GKVSRTTELAEMFMRPAPPEHLLOKLIASGEINOEQAEIARVPVADDIAVE		
Thr. sp. 10212 Pfa2p	(1627)	AKVSRTTELAELFLMPAPKNILDALVADGSSISOQAQIALLVPMADDITVE		
Thr. aureum ORF2	(1707)	GKVSRAELAEFMPPPAATVSKLLAAGLVTEEQASIAEIVPLVDDVAIE		
Thr. sp. 20892 OrfB	(1624)	AKVSRTELAEMFLRPAPQIILEKLVAEIISSQARMAAKVPMADDIAVE		
		1851		1900
Sch. sp. 9695 Pfa2p	(1671)	ADSGGHTDNRPIHVILPLIIVAQRRN---WRHLVDTPVRVAGGGGICPR		
Sch. sp. 20888 OrfB	(1771)	ADSGGHTDNRPIHVILPLIINLRDPIHRECGYEPANLVRVAGGGGICPQ		
Thr. sp. 10212 Pfa2p	(1677)	ADSGGHTDNRPIHVILPLIITQORN--RICKQYKHLKVRIGAAGGGICPK		
Thr. aureum ORF2	(1757)	ADSGGHTDNRPIHVILPVVLAIRDPMRECKYPAANRVRVAGGGGICPA		
Thr. sp. 20892 OrfB	(1674)	ADSGGHTDNRPIHVILPLIITQLRNTILAEYGCATAFRTIGAGGGGICPS		
		1901		1950
Sch. sp. 9695 Pfa2p	(1717)	AALLAFSLGAAFVVTGSGVNQIAREAGTSDAVRLILATATYSDVAMAPGG-		
Sch. sp. 20888 OrfB	(1821)	AALATFNMGASFVVTGTGNQVAKQSGTCDNVRKQAKATYSDVCMAPAAD		
Thr. sp. 10212 Pfa2p	(1725)	AAFAAFEMGAAYIATGTGNQLSKEAGTCDYVRKVLNKATYSDVMTAPAAD		
Thr. aureum ORF2	(1807)	AARAADFMDGAFFVTGTSINQLTROAGTSDSVRAALAPATYSDVMTAPAAD		
Thr. sp. 20892 OrfB	(1724)	AALAADFMDGASFVVTGTSINOICREAGTCDTVREILANSSYSDVMTAPAAD		
		1951		2000
Sch. sp. 9695 Pfa2p	(1766)	-----VOVLKQTMFAARATMLAQLOAKGSEDAVPEPOLPKLERSVF		
Sch. sp. 20888 OrfB	(1871)	MEFEGVKLOVLKKGTMFSPRANKLYELFKYDSFESMPAELARVEKRIF		
Thr. sp. 10212 Pfa2p	(1775)	MFDHGVLEQVLKKGTMFSPRANKLYELFKYKSTEEIPADENVKKLEQVVF		
Thr. aureum ORF2	(1857)	MFDQGVKLOVLKKGTMFSPRANKLYELFTTYQSLDAIPRAELARLEKRVF		
Thr. sp. 20892 OrfB	(1774)	MFDQGVKLOVLKKGTMFSPRANKLRKLEVNYESLETLPSKELKYENIIF		

FIG. 5 (cont)

		2001	2050
Sch. sp. 9695 Pfa2p	(1809)	KQSVADVWAAA ~~REKFGV~~ DATAAS-----	~~QOERMALCVRWYMSQSSRW~~
Sch. sp. 20888 OrfB	(1921)	SRAL ~~EEVWDET~~ KNEYINRLHN ~~PEK~~ TORAERDPKLEMSLCERWYLSLS ~~SHW~~	
Thr. sp. 10212 Pfa2p	(1825)	KKSF ~~DEVWDET~~ KNYIINRLHS ~~PEK~~ TERAERDAKLEMSLCERWYLSKSS ~~SHW~~	
Thr. aureum ORF2	(1907)	RMSI ~~DEVWNET~~ KQFYETPLNNPAKVARAERDPKLEMSLCERWYLSKSS ~~SHW~~	
Thr. sp. 20892 OrfB	(1824)	KQAVDQVW ~~DET~~ KREYCEKLNNPDKTARAMKDPKLEMSLCERWYLSKSS ~~SGW~~	
		2051	2100
Sch. sp. 9695 Pfa2p	(1851)	ATEATSARKADYQIWCGPAIGSE ~~NDFVRGT~~ KLDATEAGTGEFPRVVDINQH	
Sch. sp. 20888 OrfB	(1971)	ANTGASDRVNDYQVWCGPAIGSE ~~NDFIKG~~ -TYLDPAVANEXPCVMQINKQ	
Thr. sp. 10212 Pfa2p	(1875)	ANTGESGRVQDYQIWCGPAIGSE ~~NDFAKGSPCLDPEILGSE~~ PSVVOINKH	
Thr. aureum ORF2	(1957)	ASTGOVGREIDYQVWCGPTIGAEN ~~EFVKGS~~ SLDAENCGRFPCVVRINQE	
Thr. sp. 20892 OrfB	(1874)	ANAGIKSRALDYQIWCGPAMGSE ~~NNFASG~~ -ESLDWKVGTGVEPGVAEVMMA	
		2101	2140
Sch. sp. 9695 Pfa2p	(1901)	ILLGASHYR ~~RYVQ~~ QDDV ~~EYIIV~~ -----	
Sch. sp. 20888 OrfB	(2020)	ILRGACFLRRL ~~EILRNARLS~~ GAAALVASIDDTYVPAEKL	
Thr. sp. 10212 Pfa2p	(1925)	ILRGACFYQRL ~~SQ~~ LKYLNFNYEELDTLTYSASNFI----	
Thr. aureum ORF2	(2007)	ILCGAA ~~YEQRLARF~~ MLLAGR ~~ES~~ SADALAYTVAEAR-----	
Thr. sp. 20892 OrfB	(1923)	ILDGARELA ~~AKR~~ -----	

FIG. 6

	1	50
Sch. sp. 9695 Pfa3p	(1)	---MTSSKTPVWEMSKKEELDGRVVFVYNELLEFAEGDNGQVFGPPEFD
Sch. sp. 20888 OrfC	(1)	---MTSSKTPVWEMSKKEELDGRVVFVYNELLEFAEGDNGQVFGPPEFA
Thr. sp. 20892 OrfC	(1)	---MTSSKTPVWEMSKKEELDGRVVFVYNELLEFAEGDNGQVFGPPEFK
Thr. sp. 10212 Pfa3p	(1)	---MTSSKTPVWEMSKKEELDGRVVFVYNELLEFAEGDNGQVFGPPEFD
	51	100
Sch. sp. 9695 Pfa3p	(48)	---IDKVRRRVRLPAREYLLVSRVTLMDAEVNNRVRGSRMVTEYDVPVNGEL
Sch. sp. 20888 OrfC	(51)	---IDKVRRRVRLPAREYLLVSRVTLMDAEVNNRVRGSRMVTEYDVPVNGEL
Thr. sp. 20892 OrfC	(51)	---IDKVRRRVRLPAREYLLVSRVTLMDAEVNNRVRGSRMVTEYDVPVNGEL
Thr. sp. 10212 Pfa3p	(50)	---IDKVRRRVRLPAREYLLVSRVTLMDAEVNNRVRGSRMVTEYDVPVNGEL
	101	150
Sch. sp. 9695 Pfa3p	(98)	SEGGDVPWAVLVESGQCDLMLISYMGIDFOCKGRVYRLNNTLTTFGVA
Sch. sp. 20888 OrfC	(101)	SEGGDVPWAVLVESGQCDLMLISYMGIDFQNOGRVYRLNNTLTTFGVA
Thr. sp. 20892 OrfC	(101)	SEGGDVPWAVLVESGQCDLMLISYMGIDFOCKGRVYRLNNTLTTFGVA
Thr. sp. 10212 Pfa3p	(100)	SQGGDVPWAVLVESGQCDLMLISYMGIDFOCKGRVYRLNNTLTTFGVA
	151	200
Sch. sp. 9695 Pfa3p	(148)	HEGETLYDIRVTGFAKAGGCEISMFFFEYDCYVHGGLLIEMRDGAGFF
Sch. sp. 20888 OrfC	(151)	HEGETLYDIRVTGFAKAGGCEISMFFFEYDCYVHGGLLIEMRDGAGFF
Thr. sp. 20892 OrfC	(151)	KEGETLYDIRVTGFAKAGGCEISMFFFEYDCYVHGGLLIEMRDGAGFF
Thr. sp. 10212 Pfa3p	(150)	HEGETLYDIRVTGFAKAGGCEISMFFFEYDCYVHGGLLIEMRDGAGFF
	201	250
Sch. sp. 9695 Pfa3p	(198)	TDAELAAGKGVSTVAELAPRAQIOQDIAPFAPFCSHKTSLDAFEMRL
Sch. sp. 20888 OrfC	(201)	TNEELDAGKGVSTVAELAPRAQIOQDIAPFAPFCSHKTSLDAFEMRL
Thr. sp. 20892 OrfC	(201)	TDEELAAGKGVSTVAELAPRAQIOQDIAPFAPFCSHKTSLDAFEMRL
Thr. sp. 10212 Pfa3p	(200)	TDEELAAGKGVSTVAELAPRAQIOQDIAPFAPFCSHKTSLDAFEMRL
	251	300
Sch. sp. 9695 Pfa3p	(248)	LVDKQWASVFG--SGMAQEDYKLCARKMLMIDRVTHIDPHGGAYGLGLLV
Sch. sp. 20888 OrfC	(251)	LVDKQWASVFG--SGMAQEDYKLCARKMLMIDRVTHIDPHGGAYGLGLLV
Thr. sp. 20892 OrfC	(251)	LVDKQWASVFG--SGMAQEDYKLCARKMLMIDRVTHIDPHGGAYGLGLLV
Thr. sp. 10212 Pfa3p	(250)	LVDKQWASVFG--SGMAQEDYKLCARKMLMIDRVTHIDPHGGAYGLGLLV
	301	350
Sch. sp. 9695 Pfa3p	(296)	GEKILRERHWFYFCHFVNDQVMAGSLVSDGCSOLLKLYMNLGLHDTGPA
Sch. sp. 20888 OrfC	(301)	GEKILRERHWFYFCHFVNDQVMAGSLVSDGCSOLLKLYMNLGLHDTGPA
Thr. sp. 20892 OrfC	(301)	GEKILRERHWFYFCHFVNDQVMAGSLVSDGCSOLLKLYMNLGLHDTGPA
Thr. sp. 10212 Pfa3p	(298)	GEKILRERHWFYFCHFVNDQVMAGSLVSDGCSOLLKLYMNLGLHDTGPA
	351	400
Sch. sp. 9695 Pfa3p	(346)	FDLFPVSGHANKVRCRGQISPHGKLVYVMEIKEMGFDAKIGDPAIADV
Sch. sp. 20888 OrfC	(351)	FDLFPVSGHANKVRCRGQISPHGKLVYVMEIKEMGFDAKIGDPAIADV
Thr. sp. 20892 OrfC	(351)	FDLFPVSGHANKVRCRGQISPHGKLVYVMEIKEMGFDAKIGDPAIADV
Thr. sp. 10212 Pfa3p	(348)	FDLFPVSGHANKVRCRGQISPHGKLVYVMEIKEMGFDAKIGDPAIADV
	401	450
Sch. sp. 9695 Pfa3p	(396)	DIIDVNEELGQSF--NADDSYGGNLSKKIVVDFKGIALLQEGTVKSSN
Sch. sp. 20888 OrfC	(400)	DIIDVNEELGQSF--NADDSYGGNLSKKIVVDFKGIALLQEGTVKSSN
Thr. sp. 20892 OrfC	(401)	DIIDVNEELGQSF--NADDSYGGNLSKKIVVDFKGIALLQEGTVKSSN
Thr. sp. 10212 Pfa3p	(398)	DIIDVNEELGQSF--NADDSYGGNLSKKIVVDFKGIALLQEGTVKSSN
	451	500
Sch. sp. 9695 Pfa3p	(446)	ES-----MTSTSTTTTETETIA
Sch. sp. 20888 OrfC	(449)	PSKVQPVFANGAATVGPEASKASSGASASAPAPKPAFADVLAPAPPA
Thr. sp. 20892 OrfC	(450)	SS-----SSGLNESWQCVKPKSORMEHES
Thr. sp. 10212 Pfa3p	(447)	-----IDDSH--KSTTIG
	501	550
Sch. sp. 9695 Pfa3p	(463)	PPSGGLKGDFTAPSTVTHHPMAEGNGGPGPTPSFSPSAYPPRAVCFSPFP
Sch. sp. 20888 OrfC	(499)	LPEHILKGDFTAPSTVTHHPMAEGNGGPGPTPSFSPSAYPPRAVCFSPFP
Thr. sp. 20892 OrfC	(474)	PPHCLASDFEAPSTVTHHPMAEGNGGPGPTPSFSPSAYPPRAVCFSPFP
Thr. sp. 10212 Pfa3p	(459)	PPHCLASDFEAPSTVTHHPMAEGNGGPGPTPSFSPSAYPPRAVCFSPFP
	551	600
Sch. sp. 9695 Pfa3p	(513)	NNPNDNHTTGKMPLETFNMSEFMCGKVSNCGLGPEFARFDASKTSRSPAF
Sch. sp. 20888 OrfC	(547)	NNPNDNHTTGKMPLETFNMSEFMCGKVSNCGLGPEFARFDASKTSRSPAF
Thr. sp. 20892 OrfC	(522)	NNPNDNHTTGKMPLETFNMSEFMCGKVSNCGLGPEFARFDASKTSRSPAF
Thr. sp. 10212 Pfa3p	(507)	NNPNDNHTTGKMPLETFNMSEFMCGKVSNCGLGPEFARFDASKTSRSPAF

FIG. 6 (cont)

		601		650
Sch. sp. 9695 Pfa3p	(563)	DLALVTRVTSVADMEHGPFYNDVNPQGQTMVGEFDCPADAWFFGASSRD		
Sch. sp. 20888 OrfC	(597)	DLALVTRAVSVSDLEKHVNERNIDLDPSKGTMTGEFDCPADAWFYKACND		
Thr. sp. 20892 OrfC	(572)	DLALVTRVVEVTNMEHGKFLNVDNPSKGTMTGEFDCPADAWFFDGCSCND		
Thr. sp. 10212 Pfa3p	(557)	DLALVTRVSVSDMEFEKPHLNIDVNPSSKGTMTGEFDCPADAWFFDGCSCND		
		651		700
Sch. sp. 9695 Pfa3p	(613)	DHMPYSILMEIALQTSGLVTSVLKAPLTMKDDILFRNLDAEAEEVGDAM		
Sch. sp. 20888 OrfC	(647)	DHMPYSILMEIALQTSGLVTSVLKAPLTMKDDILFRNLDAEAEEVAD		
Thr. sp. 20892 OrfC	(622)	DHMPYSILMEIALQTSGLVTSVLKAPLTMKDDILFRNLDAEAEMVR-PD		
Thr. sp. 10212 Pfa3p	(607)	DHMPYSILMEIALQTSGLVTSVLKAPLTMKDDILFRNLDAEAEMVR-SD		
		701		750
Sch. sp. 9695 Pfa3p	(663)	PDYRGKTIKNTFKCTGYSMGLGKMGHRTFELSVDGAVFYKGSTSFGWFT		
Sch. sp. 20888 OrfC	(696)	PDYRGKTIKNTFKCTGYSMGLGEMGVHRTFELYVDDVLFYKGSTSFGWFT		
Thr. sp. 20892 OrfC	(671)	PDYRGKTIKNTFKCTGYAMGLGKMGHRTFELSVDGVVVFYKGSTSFGWFT		
Thr. sp. 10212 Pfa3p	(656)	PDYRGKTIKNTFKCTGYSMGLGKMGHRTFELSVDGVVVFYKGSTSFGWFT		
		751		800
Sch. sp. 9695 Pfa3p	(713)	PEVFESQTLGDNKGKRLPWYREN-----NVAVDTLTSLAPASASSAQ		
Sch. sp. 20888 OrfC	(746)	PEVFASQGLDNGKKEPWEIENKVPASQVSEFDRPNQSGRTAIFANAP		
Thr. sp. 20892 OrfC	(721)	PEVFASQGLDNGKKEPWEIENKVPASQVSEFDRPNQSGRTAIFANAP		
Thr. sp. 10212 Pfa3p	(706)	PEVFESQVGLDNGKKEVQPWYLEQKS--SNVVTYDVAETAGKDKLFSKIGS		
		801		850
Sch. sp. 9695 Pfa3p	(753)	GOLOLORRGQAQFLDTIHLAGSGAGVHGQGYAHGKAVNKQDWFSSCHF		
Sch. sp. 20888 OrfC	(796)	SGAQLNRRIDOGYLDVIVSG-SGKSLGYAHGSKTVNPNDFWFFSCHF		
Thr. sp. 20892 OrfC	(769)	THAOVLRRSECEVLDYINLAPD-SGLHGKGYAHGKDVNPQDWFSSCHF		
Thr. sp. 10212 Pfa3p	(754)	KDAOVORRNDOCEFLDTMHIIIPN-TGKYNKGYAHGKKNVNPQDWFSSCHF		
		851		900
Sch. sp. 9695 Pfa3p	(803)	WFDFVMPGSLGIESMFQLEAFCVKOGLAARHGIAHPVEFAHAP-GATSWK		
Sch. sp. 20888 OrfC	(845)	WFDFVMPGSLGIESMFQLEAFAHEDLAGKARHCOPHLCARPRARSSWK		
Thr. sp. 20892 OrfC	(818)	WFDFVMPGSLGIESMFQLEAFVADQNPGEINVSNPTEFAHAP-GKTAWK		
Thr. sp. 10212 Pfa3p	(803)	WFDFVMPGSLGIESMFQLEAFSIDQGIASFHGIIVNPTFAHNS-GKTSWK		
		901		950
Sch. sp. 9695 Pfa3p	(852)	YRGQLTEKNDKMDSEVHIKSVAAFFSS---VVDVADGFLFVDGLRVYSAD		
Sch. sp. 20888 OrfC	(895)	YRGQLTEKSKMDSEVHIVSVDAHDG---VVDVADGFLWADSLRVYSVS		
Thr. sp. 20892 OrfC	(867)	YRGQLTEKNDKMDCEVHIVSITASPENGGVVDVADGALVVDGLRVYEAK		
Thr. sp. 10212 Pfa3p	(852)	YRGQLNNKGMKMDSEIHIKDIVKNADG--TVDLIADGFLLVDSLRVYSAD		
		951		1000
Sch. sp. 9695 Pfa3p	(899)	NLRVRIQPGAGHVEEQEVAAKATTN-----SSIADVDVADLQAL		
Sch. sp. 20888 OrfC	(942)	NLRVRIASGEAPAAASSAASVGSASSRTERTRSSPAVSSGPOTIDLKQL		
Thr. sp. 20892 OrfC	(917)	ELRVRVVSAKPNIPDVQOQPPSAKADPG-----KTGVALSPTQL		
Thr. sp. 10212 Pfa3p	(900)	DLRVKIVPQTKAAPKSVAAAPRHVAAPPGV-----PNTSSVEISLES		
		1001		1050
Sch. sp. 9695 Pfa3p	(939)	KQALLTLERPLQLLAG-----SEVPACAVSDLGDRGFM		
Sch. sp. 20888 OrfC	(992)	KTELLELAPPLYLSQDPTSGOLKKHTDVASCOATTVQPTGLDGLDRSFM		
Thr. sp. 20892 OrfC	(957)	RDVLLVBNPLYLGVENSNLVQFESKPATSSRIVSEKPCSTSDLGDKSFM		
Thr. sp. 10212 Pfa3p	(945)	KKELLNLEKPLYLETSHI--YKQFGDVNNGOASVIPPCTINDLGRSFM		
		1051		1100
Sch. sp. 9695 Pfa3p	(972)	ETYGVVAPLYSGAMAKGIASADLVIAMGORKMLGSFGAGGLPMHVRAGI		
Sch. sp. 20888 OrfC	(1042)	ETYGVVAPLYTGAMAKGIASADLVIAAGKRKILGSFGAGGLPMHVRAL		
Thr. sp. 20892 OrfC	(1007)	ETYNVSAPLYTGAMAKGIASADLVIAAGKRKILGSFGAGGLPISIVREAL		
Thr. sp. 10212 Pfa3p	(993)	ETYNVAPLYTGAMAKGIASADLVIAAGKRKILGSFGAGGLPMHVRASV		
		1101		1150
Sch. sp. 9695 Pfa3p	(1022)	EKIQOALPAGPYAVNLIHSPFDNLEKGNVDLFLEKGVTVVEASAFMTLT		
Sch. sp. 20888 OrfC	(1092)	EKIQOALPOGPYAVNLIHSPFDNLEKGNVDLFLEKGVTVVEASAFMTLT		
Thr. sp. 20892 OrfC	(1057)	EKIQOHLPHGPYAVNLIHSPFDNLEKGNVDLFLEMGVTVVECSAFMTLT		
Thr. sp. 10212 Pfa3p	(1043)	EKIQOALPEGPYAVNLIHSPFDNLEKGNVDLFLEKGVHVVEASAFALT		
		1151		1200
Sch. sp. 9695 Pfa3p	(1072)	POVVRYRAGLSRDARGGSVRTAHKIIGKVSRTELAEMFIRPAPQALLOK		
Sch. sp. 20888 OrfC	(1142)	POVVRYRAGLSRNADG-SVNIRNRIIGKVSRTELAEMFIRPAPAHLEK		
Thr. sp. 20892 OrfC	(1107)	AOVVRYRAGLSKSNADG-SIRIAHRIIGKVSRTELAEMFIRPAPOHLOK		
Thr. sp. 10212 Pfa3p	(1093)	TOVVRYRACGLSPAKDG-SVLIKRIIGKVSRTELAEMFFRPAPONLLOK		

FIG. 6 (cont)

		1201		1250
Sch. sp. 9695 Pfa3p	(1122)	LVASGEITPEQAALALEVPMADDIAVEADSGGHTDNRPIHVILPLIISLR		
Sch. sp. 20888 OrfC	(1191)	LVASGEITQEQAEIARRVPVADDIAVEADSGGHTDNRPIHVILPLIINLR		
Thr. sp. 20892 OrfC	(1156)	LVASGEITAEQAELATQVPVADDIAVEADSGGHTDNRPIHVILPLIINLR		
Thr. sp. 10212 Pfa3p	(1142)	LVASGEITKEQASLAEVPMADDIAVEADSGGHTDNRPIHVILPLIINLR		
		1251		1300
Sch. sp. 9695 Pfa3p	(1172)	NRLQREIKYPARHRVRVGAGGGIGCPQAAALGAFHMGAAFFVTGTVNQLSR		
Sch. sp. 20888 OrfC	(1241)	NRLHRECGYPAHLRVRVGAGGGIGCPQAAAAALTMGAAFFVTGTVNQVAK		
Thr. sp. 20892 OrfC	(1206)	NRLHRELDYPSHLRVRVGAGGGIGCPQAAALAFQMGAAFLITGTVNQLAR		
Thr. sp. 10212 Pfa3p	(1192)	NRLHRECGYPAHLRVRVGAGGGIGCPQAAALAFQMGAAFLITGTVNQVSK		
		1301		1350
Sch. sp. 9695 Pfa3p	(1222)	QSGTCDNVRKQLSEATYSDDITMAPAADMFEGVLEQVLKKGTMFPSRAKK		
Sch. sp. 20888 OrfC	(1291)	QSGTCDNVRKQLSOATYSDDICMAPAADMFEGVLEQVLKKGTMFPSRAKK		
Thr. sp. 20892 OrfC	(1256)	ESGTCDNVRKQLSEATYSDDICMAPAADMFEGVLEQVLKKGTMFPSRAKK		
Thr. sp. 10212 Pfa3p	(1242)	QSGTCDIVRKQLSEATYSDDITMAPAADMFEGVLEQVLKKGTMFPSRAKK		
		1351		1400
Sch. sp. 9695 Pfa3p	(1272)	LYELFHKYDSFEAMPDELARVEKRIFSKSLAEVWAETKDFYITRLNPE		
Sch. sp. 20888 OrfC	(1341)	LYELFCKYDSFDSMPPELERIEKRIFKRALQEVWEETKDFYINGLKNPE		
Thr. sp. 20892 OrfC	(1306)	LYELFCKYDSFEAMPPEELQREKRIFOKSLAEVWQETSDFYIHRKNPE		
Thr. sp. 10212 Pfa3p	(1292)	LYELFCMYNSFDDMPKSELOREKRIFOKSLAEVWEETKDFYINRLNPE		
		1401		1450
Sch. sp. 9695 Pfa3p	(1322)	KIRKAENEDPKLKMSLCFRWYLGLSSEWANNGLADRTMDYQIWCGPAIGA		
Sch. sp. 20888 OrfC	(1391)	-KIQPAEHDPKLLKMSLCFRWYLGLSRWANMGAPDRVMDYQIWCGPAIGA		
Thr. sp. 20892 OrfC	(1356)	-KINPAASDGKLLKMSLCFRWYLGLSSEWANSQAQDRVMDYQIWCGPAIGA		
Thr. sp. 10212 Pfa3p	(1342)	KIEHAEEKDDPKLKMSLCFRWYLGLSSEWANNGLKERSMDYQIWCGPAIGS		
		1451		1500
Sch. sp. 9695 Pfa3p	(1372)	ENDFIADSYLDPAVSGEFPDVQINLOILSGAAYLQRIILSVLA--PRID		
Sch. sp. 20888 OrfC	(1440)	ENDFIKGTLYLDPVSNEXPCVQINLOILRGACYLRRINALRND--PRID		
Thr. sp. 20892 OrfC	(1405)	ENDFTKGTLYLDTVAKSYPCVQINLOILQGAAYLKRIQVIRFDRMLLQA		
Thr. sp. 10212 Pfa3p	(1392)	ENDFVKGTYLDPVAGSYPCVQINLOILRGACFLQRIIRAIKHD--PRID		
		1501		1517
Sch. sp. 9695 Pfa3p	(1420)	VDTEDDLFTYRDPHAL-		
Sch. sp. 20888 OrfC	(1488)	LETEDAAFFVYEPTNAL-		
Thr. sp. 20892 OrfC	(1455)	VDIDDPVFTYVPTQPL-		
Thr. sp. 10212 Pfa3p	(1440)	IDVDEDVFTYRPESTL-		

FIG. 8

MDTRIAVGMSSAILPSGENVRESWEAIRDGLDCLSDLPADRVDVTAYYNPEKTTTKDKIYCKRGFIPEYDFDAREFGLNMFQMEDSDANQTSILLKV
 KEALTDANIPAFSSGKKNIGCVLGIGGGQKASHEFYSRNLNVVVDKVLKMGLEPEEDVAAAADKYSKASPEWRLDSFGFLGNVTAGRCNCTENM
 EGMNCVVDAACASSLIAVKVAIEELLYGDCDAMIAGATCTDINSIGMYMAFSKTPVFSTDPVKAYDAATKGLIGESAMLVLKRYADAVRDGD
 TVHAVIKGCASSSDGKAAAGIYTPTISGQEEALRRAYARANVDPATVTLVEGHGTGTPVGDKIELTALSNIKSKAFSANGGGAEEAEQVAVGSIKSQI
 GHLKAVAGLAGLVKVV LALKHKTLPTQINVDKPPSLVDGTPIQOSPPLYVNTMNRPFVTPVGVPRRAGVSSFGGGANYHAVLEEFEPHEHESAYRY
 NNLPQVALLHAGDVATLAATVRAKLALATAEQEEARVVKNADYIAHFRFLDECKLRGAVPQAHARVGLLVRDLSSLIAVLEAAAKLAGEESAT
 EWTVSATGEAAFRVRGVATEANVAALFSGGQAQYTHMFSDVAMNWPFPRESVAMDRQAQREFRGPAKRVS SVLYPRKPYGDEPRQDHKEIS
 QTRYSQPATLACSVGAFDIFKAAAGLAPSAAGHSLGEFAALYAAAGSLDRDAVFDLV CARAKAMSDFTAQASSGGGMAAVIGAKADQLSLGGAPD
 VWLANSNPSQTVITGTAEAVAAASDKLRCSGNFRVPLACEAAAFHSPHMRGAEQTFASALAQAPVPSAPAAARFYSNVTGGAAV TSPADVKTNLG
 KHMTPVQVQQVRAMHAAAGARVFEFGPKQVLSRLVKEITLGEAGDVTVAVNPDSAKDSDTQLRQAALTLAVAGVPLKDFDRWQLPDAITRLE
 PVKKKKTTLRLSAAITYVSAKTLRQREAVLNDGYTVSGATAVVKEVDTANEERLVRQAQDLQRQLAEASTAAQAAQSKVAELERTIQDLERKVVQQ
 QQQEKGENSDSNAAEVLRRHKELLQRMLODCDEQAVPVATVPTPTSSPTTSSPVSGNSKSTRGSADLQALLAKAETVVMVMAVLAAKTGYEAD
 MVEADMIDLEAELGIDSIKRVEILSEVQQLGVEAKDVSALSRTRTVGEVVDAMKAEIV AASGGSAPAVPSAPAAASAPTAASTAPSADLQALLSK
 AETVVMVLAAKTGYEADMVEADMIDLEAELGIDSIKRVEILSEVQQLGVEAKDVSALSRTRTVGEVVDAMKAEIV AASGGSAPAVPSAPAA
 SAAPTAASTAPSADLQALLSKAETVVMVLAAKTGYEADMVEADMIDLEAELGIDSIKRVEILSEVQQLGVEAKDVSALSRTRTVGEVVDAMK
 AEIV AASGGSAPAVPSAPAAASAPTAASTAPSADLQALLAKAETVVMVLAAKTGYEADMVEADMIDLEAELGIDSIKRVEILSEVQQLGVE
 AKDVSALSRTRTVGEVVDAMKAEIV AASGGSAPAVPSAPAAASAPTAASTAPSADLQALLSKAETVVMVLAAKTGYEADMVEADMIDLEA
 ELGIDSIKRVEILSEVQQLGVEAKDVSALSRTRTVGEVVDAMKAEIV AASGGSAPAAAVPSAPAAASAPTAASTAPSADLQALLSKAETVVMVLA
 AKTGYEADMVEADMIDLEAELGIDSIKRVEILSEVQQLGVEAKDVSALSRTRTVGEVVDAMKAEIV AASGGSAPAAVPSAPALLPTLFGSECEDSL
 TFPVITTLPLPAELVLAEGGARPVVVDGGSALTSSLVSSLDGRAVLLQVQSSSACSPRSTTHKLVTVADRSEAAALQAALTSVEAQFGKVGGFVFQF
 GDDDVQAQLGWALLAAKHLKTSLSSEQIEGGRTFFVAVARLDGQLSGKSTTATVDLSRAQQGSVFGICKTLDLEWPAVFCRGIDLAADLDAAQ
 AARCLLGELSDPDVAVRESGY SASGQRCTTTTTSKLTGKPHQPISSDDLFLVSGGARGITPLCVREL AQRVGGGTYYLIGRSELPTTEPAWAVGVESG
 KPLEKAAALFLKAEFAAGRGAKPTPMLHKKLVGAVVVGAREVRSALAEITAQGATAVYESCDVSSAAKVREMVVVQQGGRRVSGVFHASGVL
 RDKLVENKSLADFSAVYDTKVGGLINLLACVDLAQLRHLVLFSSLAGFHGNVGGSDYAMANEALNKLAAHLSAVHPQLCARSCICFGPWDGGMVT
 PALKANFIRMGIQIIPRQGAQTVANMLVSSSPGQLLVGNVGPVPSATEHTVLTQLRQSDNPFLDSHVIGQRRVLPMTLAVGYMAHQASIIYA
 GHQLWAVEDAQLFKGIAIDNGADVVRVELSRRKEEQEDAGKVVKVQVLLKSQVNGKSVPAKYKATVVLSPAPRPSVITRDFDLTPDPACTEHDIL
 YDGKTLFHGKAFAQGIEQVLSATPKQLTAKCRNLPLTPEQRGQFVNL SQQDPFQADIAFQAMLVWARMLRQSAAALPNNCERFDFFYKPMAPGATY
 YTSVKLASASPLVDSVCKCTVAMHDEQGEVYFSARASVVLNKLTY

(SEQ ID NO:2)

FIG. 9

[illegible]

FIG. 10

MPCDNIAVVGMAVQYAGCKNQDEFWDTLMRKEINSSPISAERLGTTRYRDLHFHPQRSKYADTFENDRYGCVDSVDNEHDLADLARRALLDAG
 INLDDASTTANLRFGLVSGCLSPFMDNLQGELLNLYQVHVENVVGAQRFRDSRPWSEPRVAVSEASDPVYSDPASFVANQLGLGPVRYSLDAA
 CASALYCLKLASDHLLSRSDVMLCGATCFDPFFILSGFSTFQAMPLGGDDNPLSVPLRQGSQGLTPGEGGAIMVLKRLEDVARDGDRJYGTLLG
 TSLSNAGCGLPLSPHLPSEKSCMEDLYTSVGIDPSEVQYVECHATGTPQGDVVVEALRHCFRGNTDHPPRMGSTKGNFGHTLVAAAGFAGMAKVL
 LSMQHGTIPPTPGVDRSNCIDPLVDEAIPWPYSSAQAARAGKPGDELK CASLAFGFGGTNAHCVFREHRQIAATATASPVLPEVTPGPIAIGMDAT
 FGTCLKGLDAFEQAIYKGTDGASDLPSKRWRFLGADTDLTAMGLDAVPRGCYVRDVVDYKRLRSPMIPEDVLRPQQLAVATMDRALQDAGM
 ATGGKAVLVGLGTDTELYRHRARVTLKERLDPAAFSEQVQEMMDYINDCGTSTSYIGNLVATRVSSQWGFTGPSFTVTEGANSVYRCLEL
 GKFLDTHQVDAVVVAGVDLCAENLYLKARRSAISRQDHPRANFEASADGYFAGEGSGALVLKRQADVGSDDKVYASVAGLTCAAQPAEAV
 SPLLLQVHNDNDNEKRVVEMVELAADSGRHAPHLANSPLSAESQLEQVSKLLAHQVPGSVAIGSVRANVGDVGYASGAASLIKTALCLHNRYLPAN
 PPOWERPVAPVSEALFTCPRSRAWLKNPGESRLAAVASASESGSCFVLLTDEYATHESNRSLDDAAPKLLAIRGDTVDDIMAKVNAELALLRAH
 AETGSATDDDPAAAFAFTAHRLRFLRVGETVASHGATATLCLALLTPEKLEKELELAAGKGVPRSAKAGRNWMSPSGSAFAPTPTSDRVAFMY
 GEGRSPYYGVGLDLHRLWPALHERINDKTAALWENGDSWLMPRAVDADSQRVQTAFDADQIEMFRTGIFVSICLTDYARDVLGVQPKACFGLSL
 GEISMLFALSRNCGLSQDLTQRLTSPVWSTQLAVEFQALRKLWNVPADAPVESFWQGYLVASRAEIEKAGPDNRFVRLLIIVNDSSSALLIAGKP
 AECLRVLERLGGRLPPMPVKQGMIGHCEVAPYTPGIAHIEILEIPDSPVKMYTSTVNAELRGGSSSITEFVQKLYTRIADFPGVDKVSRDGHDFV
 VEVGPNNMRSAAVSDILGKAATPHVSVALDRPSESATQTLKSLALLTAHRVPLHNPTLFADLYHPTFLT AIDSAMQEPKPNRFLRSVEVNGYFC
 PDGISKQVAAASAKPSTHCMVRLHPAKAVVAAAAGAVVADSTPVVKAQKQTSLSLLVGDDAFLRCYDWDWPLYMGAMAEIGISSVDLVVAAAEAR
 MLASFGAARLPMQVELQIREIQRTSNFAFVNLMPPGDEAATVDALLRTGVSNVEASGYTGALSADLVRYRVTGLRRTSCGASVSATHRVVAKV
 SRTEVAEHFLRPAPAAVLEALVAAKQITPEQAALASRVAMADDVAVEADSGGHTDNRPIHVLLPLVVAQRNRWRHLVDTFVRVAGAGGGIACPRA
 ALLAFSLGAAFVVTGSVNQLAREAGTSDAVRLLLATATYSVAMAPGGVQVLKKQTMFAARAATMLAQLQAKFGSFDVPEPQLRKLERSVFKQS
 VADVWAAAREKFGVDATAASPQERMALCVRWYMSQSSRWATEATSARKADYQIWC/GPAIGSFNDFVRGTKLDATAGTGEFPRVVDINQHILLGA
 SHYRRVQQQQQDDDDVEYIV

(SEQ ID NO:4)

FIG. 12

MTSSKKTVPVWEMSKELLDDGKTVVFDYNELLEFAEGDVGVFGPEFDIIDKYRRRVRLPAREYLLVSRVTLMDAEVNNFRVGSRMVTEYDVPVNG
 ELSEGGDVPWA VLVESGQC DLM LISYMGIDFQCKGDRVYRLLNTSLTFGVAHEGETLVYDIRVTGFAKGAGGEISMFFFEYDCFVDGRLLIEMRD
 GCAGFTDAELAAAGKGVLTAKAELAAARAQIQKQDIAPFAPAPC'SHIKTSLDAREMRLLVDRQWARVFGSGMAGIDYKLCARKMLMIDRVTHLDP
 GGAHGLGLIGEKVLERDHWYFPCHFVRDEVMAAGSLVSDGCSQLLKVVYMLWGLHTTVGAFDRPVSGHANKVRCRGQISPHKGKLVYVMEIKE
 MGFDAKTGDPFALADVDIIDVNFEEGQAFAGVEDLHSYGQGDLRKKIVVDFKGIALSLQKRKEQQKESMTVTITTTTSRVIAPPSGCLKGDPTAPT
 SVTWHPMAEGNGGPGPTTSPSPSAYPPRAVCFSPFPNNPLDNDHTPGQMPLTWFNMFSEFCGKVSNC LGPEFARFDASKTSRSPAFDLALVTRVTS
 VADMEHGPFPYNDVNPQGTMVGEFDCPADAWFFGASSRDDHMPYSILMEIALQTSGLVTSVLKAPLTMDKDDILFRNLDA DAELVGDAMPDVR
 GKTIRNFTKCTGYSMLGKMGHRTFELSVDGAVFYKGSTSFGWVFEVFEQGLDNGKPLPWPYRENNVAVDTLSAPASASSAQGLQLQRRG
 SQAQFLDTIHLAGSGAGVHGQGYAHGEKAVNKQDWFSCHFVFDVPVMPGSLGIESMFQLVEAWCVKQGLAARHGIAHPVFAHAPGATSWKYRG
 QLT'PKNDRMDSEVHIKSVAAFSSWVDVVADGFLFVDGLRVYSADNLRVRIQTGAGHVEEQEVAAKATTKNSSIADVDVADLQALKQALLTLERPL
 QLDAGSEVPACAVSDLGDRGFMETYGVVAPLYSGAMAKGIASADLVIAMGQRKMLGSFGAGGLP'MHVVRAGIEKIQAAALPAGPYAVNLIHSPFD
 ANLEKGNVDLFLEKGVRVVEASAFMELTPQVVRYRATGLSRDARGGSVRTAHKTIKVSRTELAEMFIRPAPQAILDKLVASGEITPEQAALALEVP
 MADDIAVEADSGGHTDNRPIHVILPLILSLRNRLQRELKYPARHVRVVGAGGGIGCPQAAALGAFHMGAAFFVTGT'VNQLSRQAGTCDNVRRQLSR
 ATYSDITMAPAADMFEGQVELQVLKKGTMFPSPRAKKLLELFHKYDSFEAMPADELARVEKRIFSKSLAEVWAETKDFYITRLNNPEKIRKAENEDP
 KLKMSLCFRWYTLGLSSFWANNGIADRTMDYQIWCGPAIGA FNDFIADSYLDVAVSGEFPDVVQINLQILSGAAYLQRLLSVKLAPRIDVDTEDDLF
 TYRPDHAL

(SEQ ID NO:6)

[illegible]

(SEQ ID NO:68)

AAATTGGATATGGAAATTGGAAACCGAACTTGGTATTGATCTATCAAGAGAGTAGAATAATCTTTCTGAAGTTC AAGCTCAATTGAGTGTGGAAGCTAAAGATGATAGATGCTCTTT
GTAGAACCTCGTACTGTGTGGAAGTATTGATGCAATGAAGCCGAAATTTCTGGTGTGAGCCAGCTGCTCTGTTCAAGTTGCAAGTCTCTACTCAATAAGTTGCTCTGTTCAATAAGTTGCTCTGTT
AAGCATCTGCTCTGTGATTCTGGTTGTGGCAAGAGCGAAACAAGTTGATTGGAAAGTGTAGCATCC AAGACTGGTTATGAAACTGAGTTGATTGAATAGATATGGAAT
TGGAAACCGAACTTGGTATTGATCTATCAAGAGAGTAGAATAATCTTTCTGAAGTTC AAGCTCAAGTGAAGTGTGGAAGCTAAAGATGATAGATGCTCTTTAGTAGAACTCGTACTG
TTGGTGAAGTGAATTGATCAATGAAGCTGAAATTTCTGGTGTGCAACCAAGCTGCTCTGTTCAAGTTGCAGCTCTACTCAATAAGTTGCTCTGTTCAAGTATCTGCTCCTCTG
TTGATTCTGGTTGTAGCAAGCGCGAAACAAGTTGATTGGAAAGTATTGGCATCAAGACTGTTATGAGACTGATGTTGAATTGGATATGGAATTGGAAATTTGGAACCTGAAACTG
GTATTGATTCTATCAAGAGAGTAGAATAATCTTTCTGAAGTTC AAGCTCAATTGAATGTTGAAGCTAAAGATGATAGATGCTCTTAGTAGAACTGCTACTGTTGGTGAAGTGAATG
ATGCAATGAAGCCGAAATTTCTGGTGTCAACCAAGCTGCTCTGTTCAAGTTGCAGCTCTGCTCCAGTAGTTGCTCTGTTCAAGTATCTACTCTGTTGATTTCTGTTGTTGT
GGCAAAAGCGGAACAAGTTGATTGGAAAGTGTAGCATGCAAGACTGGTTATGAAACTGAGTTGATTGAATTTGGATATGGAATTTGGAACCTGAACTTGGTATTGATTCTATCA
AGAGAGTAGAATAATCTTTCTGAAGTTC AAGCTCAATTGAGTGTGGAAGCTAAAGATGATAGATGCTCTTAGTAGAACTGCTACTGTTGGTGAAGTGAATGATGCAATAAGAAAGC
GAAATTTCTGGTGTGCAACCAAGCTGCTCTGTTCAAGTTGCAGCTCTACTCAAGTAGTTGCTCTGTTAAAGTATCTACTCTGTTGATTTCTGTTGTTAGCAAGCGGAAAC
AAGTAGTATTGGAATATTGGCATCAAGACTGGTATTGAACTGATTGAATTAGATATGGAATTTGGAACCTGAACTTGGTATTGATTCTATCAAGAGAGTAGAATAAT
CTTTCTGAAGTTC AAGCTCAATTGAATGGAAGCTAAAGATGTTGGATGCTCTTAGTAGAACTGCTACTGTTGGTGAAGTGAATTGATCAATAAGAAAGCCGAAATTTGCTGGTGA
TCAACCTGCTCAGCTGATGTTCCAGTTCAAGCTAAGAGTGGTGTAGCCAACCTGCACCTTTTGGCAAGAGCGGAACAAGTAGTATTGGAAGTATTGGCATCCAAGACCTGGTT
ATGAAACTGAGCTGATTGAATTTGGATATGGAATTTGGAACCTGAACTTGGTATTGATTCAATCAAGAGAGTAGAATAATTTCTGCTCCGAAAGTTCAAGCAGAAATGAGTTGGAAGCA
AAAGATGTAGAGCGCTTAAGTAGAAACCCGTAAGTTGGGGAAGTAGCATGCAATAAGAGCTGAAATTTGCGCAGTGTCTACGGTTCAGCTTTGGATGATTC AACAAT
TATGGAGGACAGATGATGAAGATGAAGACTTTATTTATACGATCATGATACGGAAGCAATGTGAAGACTCTTAGTCTGAGTTTCTACCTGTAAGAGCACTCCCGCGC
CTGATAAACCTTTTGTGGATAACATTGCTGAAGGCCAATTGTTATTGGTATTGTGGATTTGGGAATGTGGAAC AAGCTTCAACTGAACCTTGAACCAAGCTATTGGAGAACCTGCTGGTGGCTA
CAITTCAGTGCAAGAGCTTGGTATCCCGTGGATTCGTTGGTAAATCATTTACTTAGGAAATACAGAAGAAAGTAGATCGAAAGATGGTTTCAAGCATTTGAATCTTCGTAT
GGGAAATAATTGGTGGCTTTGTTATCAACATTTTTCATGATAGCGACTATGGTAGCAACTTGGATGGCGTTAATGGCAGCGAAACAATTTGGAAGAGTCCCTCAACGACCAGAT
TAAGAATGGAAGAACCTCTTTTGGCTGTGCGCGTATGAATGGTGAACCTTGGTAGGAC AATGCTCAGTTCTATGATCAAGGAANTAGTGAATCATGCGCGTATCGCCGAAC
GTGTGGTGTCTATTGTGTTTGTGCAAAACTTTGGATTGGAATGGCCTAATGTTTGTCTGGTGTGATATTGCTGAAGGTATGAGTTATAGTTTGGCTGCGGAATTGATTGT
TGATGAGATTCTTGTGCAAACTTTCTCCATTCTGGGAATCTGGTTACAGATTAGCGGAGAAGATTCACAACCTGAAGCTCAACAATTTGGTCAAGCTGGAAGCTCATCTCCGAT
TAAGAAGAAAGGATGCTTCTAGTATCTGTGGTGTCTGCTGGTATTACTCCACTTTGATTCTGTAATTCGTGAAATTTGCTGAAGCTGAAGAGTGAAGTGTGATGTTGCTGGT
AGCTTTTGGTGAATGAAACCTTTGGCTTAATGTTAAATCCGGAAGATTATAGATAAGCTGTTTGGCATTTTGAAGAGAGATTTCAGCTGGCGTGTGATGAACCAA
CTCCAAGAGTTC ACAAATCTTTGATTATAAAGTGTCTGGTATTAGGAGGTTAGAGCATCTATTGCAATAATAGAAGCCCATGGAGCAAAAAGCTATATATTTGTTCTTGGCAT
GTATCTTCCGCTGAGAAAGTAAGGCTGCAGTGC AAAAAAGTTGA AAGGAGCATCTAGTTCTGATTACTGGTATTGTGCATGTCATCAGGCCGTTTGGAGGATAAATTTGGTTGA
GGAACAAACTTTTGGATTTC AAGCAGATATGGAACCA AAGTAAGTGAACCTGACTGAATAACTTGTCTGTCAGTGAACATGAANTTTGTCTGTCATTTGGTTATGTTTAGTTCT
TTTGGCTGGATATCATGGAATGTTGGTCAATCTGATTATGCAATGGCTAACGAATCACTTAACAAGATGGTGTTAGTGGTGCAGCTTATCTCAATTTCTCAATTTGTTGTTAAATCT
ATTTGTTTGGACCTTGGATGGATGGTAAGTGAATCCAGCTTTG AAAAAACAATTTCAATCAATGGGTGCCAGATTATTCCTCGTGAAGGTGGCGCGGAGACTGTTGCAAC
AATAGCTTATCTTCAATCTCTCAAGTTTATGTGGCACTGGGGTGTCTCCAGTTTCACTTTGTCA AATCTGTGCAAACTATTGTTCAAACTTATCCCTGAGTGAATAAT
CAITTTAAAGTCTCATAAATTCATGGTAAAAATGTTTGTGCTATGACTGTAGCAATTGGATATCTTGCTCACTTGGTTAAGAAATTTTATGCTGGTCACTATTTGTGGGAGT
TGAAGATGCTCAATTTGTCAGTGGTGTGTAATTGACCATCGCGGTGCAAGCTCAAGTGAATAATTAAACGGACAGAGTTGGATGATGGCAAGGTAAAAAGTTCAAGCTGTTC
TGACTGCTTCAACGATAATGGA AAAATGGTAACCTGCATAC A AAGCAGTGAATTGTTTGGGAAAAACAAGTAGACCTGCGTTTATTTTGAAGAAGATTTTTCATTGCAAGAATCT
AAATTCGCGAGTGTGATGATGGTAAAAACTTTGTTTTCATGGTCCATATTTCTGGGAATTACCAAGTGTGGAATGTATCTGATACCTTCACTCAACAACCTCAATGTA
CCAATATTGATTGACTGTCAACTGGTCAATTTCGGGATATCGAACCTCTTTTATGGCGGATGCTGCATTTCAAGCTATGCTTGTATGCTGTAGCAAACTTAATTA
GGAATATGCTATCTTACCAACAACATTTGGAAAGAGTAGATATCTATAAACCAATAGCACCTGGTGAAGAGTATTACACTACTTTTGAACGCTTTGGGTGAATACCTCCGGTCTCT
TCTCAAGTCTGATTTTATATGCACGATGAACAAGGAGAAAGTATTCTATCTGGAAGAGCTAGTGTGTTGTGAATGACAAGATGGAGTTTTT

[illegible]

[illegible]

FIG. 15

MEDQRIALVGLSAILPSGENVRESWEAIRDGLNCLSDLPADRVDVTAYNPTKGVKDKIYKRGHPIEYEFDSREFGLNMLQMEDSDANQTLTLKVKAEALDDANIPAFTEKKNIG
 CVLGI GGQOKASHEFFYSRLNVVVDKVLKMGLPDEDETAVAEKFKANFEWRLDSFGFLGNVTAGRCTNTFNMEGMCVVDAAACASSLIAIKVAIDELHGHDCDAMIAGATCTD
 NALGMYMAFSKTPVFTSDQSCLAYDEKTKGMLIGESAMFVLKRYADAVRDGDTVHAVIRSSSSDQKASGYTPTISGQEEAILRAYRRAGVSPNTITLVEGHGTGTPVGDKIET
 ALRNVFDKAY'GPGHKEEVAVGSKISOJGHLKAVAGCAGLVKLMALKHKTLTPOSINVENPPNLVDG'TVISDITLTYINTMNPWTGKPVPRRAGISSFGGANYHVALEEFEPQTK
 PYRLNVSQPMMLHANANSLQKLCEDQKLLKESREKCVNTKNTDYVAFSKFQDSFKLGSVPSQHARYGFAKSIEDTISLSAIVNRQKDITTSWALPKEGAFIRSTRALINDNKS
 VAAALFSGOGAQTTHMFENDVAMQWPQFRLCVNDMEKAQOEIVNDKSVKRISQVMFPRKPYARESPLDNKEISKTEYSQTTTVASSVGLFEFRDAGFAPAFVAGHSLGEFSALYAAGL
 IDREDLFLVCNRAMAMRDAPKKSADGAMAAVIGPNASSIKLSAPEVWVANNNSQTVITGANSVQAETSCLKTQGRPVVHLACDGAHSPHIMENAEOQKALSAVKFNKPTG
 SPSPKIFSNVTGGVFTDPKLTALSRHMTSSVQFLTQIKNMYAAGARVIEFGPKQVLSKL VNEIFFGDTSVLT VSNPNSAKDSIDIQRAAVQMAVAGVALTDFDKWELKDPTRMKEFP
 RKKTTTLTSAATVVSCKTLQERERIMNDGRTVSCVQRIENTNTGELEKLLKQLODKENEVVRVQALATQASADLQNTKAELQKAOATKSSNAASDAVVAKHKAILLAMLELETGK
 AVDYSSFSKGQV'ASPATV'RVVSAPVQAAAPVQVSASVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLNVEAKDVALSRTRTVGEVIDAMKAEIAGGQPAAP
 IAGGQPAAPVQVAAAPTQVVPVQASAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLSVEAKDVALSRTRTVGEVIDAMKAEIAGGQPAAP
 VQVAAAPTQVVPVQASAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLSVEAKDVALSRTRTVGEVIDAMKAEISGGQPAAPVQVAAAPTQI
 VAPVQVSAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLSVEAKDVALSRTRTVGEVIDAMKAEISGGQPAAPVQVAAAPTQIVAPVQVSAPV
 DSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLSVEAKDVALSRTRTVGEVIDAMKAEISGGQPAAPVQVAAAPTQIVAPVQASAPVDSGLLAKAEQ
 VVLEVLASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLNVEAKDVALSRTRTVGEVIDAMKAEISGGQPAAPVQVAAAPTQIVAPVQVSAPVDSGLLAKAEQVVLVLELASKT
 GYETELIEMLETELGHDSIKRVEILSEVQAQLNVEAKDVALSRTRTVGEVIDAMKAEIAGGQPAAPVQVAAAPTQIVAPVQVSAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLE
 MELETELGHDSIKRVEILSEVQAQLSVEAKDVALSRTRTVGEVIDAMKAEISGGQPAAPVQVAAAPTQIVAPVQVSAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHID
 SIKRVEILSEVQAQLNVEAKDVALSRTRTVGEVIDAMKAEIAGGQPAAPVQVAAAPTQIVAPVQVSAPVDSGLLAKAEQVVLVLELASKTGYETELIEMLETELGHDSIKRVEILSEVQAQLS
 AKDVALSRTRTVGEVIDAMKAEIAGS'AVTVATLDDSTIMEETDDEDEDFILYDHVYGSECDLSLSESSVKSIPRADKLLDNIAERP'VVDCTGKLTTELAKAIGERAVVATFSAQS
 LDLEWPNVFAARGVDIAEGMSYSLAAELIVDEISCANLSIRESGYTISGERFTTEAHKLVTKGPHAPIKKDAFLVSGGARGITPLCIREIAKAVKGGTYILMGRSALADEPLWANGKSGK
 DLDKAGLAFLEKKEFAAGRGSKPTPKVHKSLDKVLGIREVRASIANIEAHGAKAIVLSCDVSSAEKVKAAVQKVEKEHLVRITGIVHASGVRLDKLVENKTLDDFNAYVGTKYVTLGVN
 LLSAVNMNFVRHLVNFESSLAGYHGNV'GQSDYAMANESENKIGFRLGAAYSQLCVKSCIFGPDGGMVTPALCKQFQSMGVQIIPREGAETVARIVLSSNPQSVLVGNWGVPPVSP
 LSKSATIVQITFTPELNPLKSHQIHGKNVLPMTV'AIGYLAHLVKNFYAGHHLWGVEDAQLFSGVVIDHVAQAVKLTQSLDDDDGKVVKQAVLTASNNDNGKMPAYKAVIVLGKTS
 RPAFILKDFLSQESRSADELYDGKTLFHGFLFRGITKLLNVSDTSLTTQCTNIDLTA TERGQFADIEFPVNPFMADAFAQAMLVWVRNLRNLSASLPNNCERVDIYKPIAPGEKYYTTLQ
 ALGNTSGSVLKS'VYFMHDEQGEVFLSGRASVVVNDKMEF

(SEQ ID NO:69)

[illegible]

FIG. 16 (cont)

TATCCGCATTTATGTCATTAAACCCCTCAGGTTGTACGATACAGAGCCGCTGGTTGGCCAAAGCAAGAGATGGATCTGTGAAANNTTCAAAATTCGTATTATTGCCAAAATTTCAA
 GAACAGAGTTAGCGGAACCTGTTCTTTGAAACCAGCACCCAAAATAATTTTAGATGCATTGGTTGCGGATGGATCTATTAGTCAAGAAACAAGCCCAACTTGCATTACTTGTGCCA
 ATGGCTGATGATATTACTGTGGAAAGCTGATTCTGGTGGCATACTGACAATCGACCAATTCATGTTTGTACCTTTGATAATTCAGCAAGAATAAGAAATTTGTAACAATAC
 CCAAAACATTTAAAAGTTCGAATCGAGCAGCTGGTGGTATTGGATGCCGAAGGCAGCATTTGCTGTTGAGATGGTGCTGCATACATTGCAACTGGAACGGTAAATCA
 ACTTTCAAAGGAAGCAGGTACTTGTGACTATGTACGTAAAGTATTGAATAAAGCTACATAATTCGGATGTTACCATGGCTCCAGCCGCAGATATGTTCTGATCATGGTGTGAATT
 ACAAGTTTGAAGAAAGGTACTATGTTTCTTCACGTGCTAAAAAACTATACGATTTGTTCAAAAAAATAACAAATCGATTGAGGAATACCAGCAGATGAGGTGAAAAAAGCTTG
 AGCAAAAAAGTTTCAAAAAGTCGTTTGATGAAGTATGGATGAGACCAGAATTAATAATTAATCGTTTACATTCTCCGAAAAAATTTGAACGTGCTGAAAGAGATGCAAAA
 CTTAAAATGTCGTTATGTTTTCGTTGGTATTTGTCGAAGTCTCCAGATGGGCTAATACCGGTGAATCTGGAAAGAGTGCAGGATTATCAAAATTTGGTGGTCCAGCAATTGGG
 TCATATAATGATTTTGCGAAGGATCACCATGTTTGGATCCTGAGATTTTGGGTAGTTTTCCAAGTGTGTTTCAGATTAAATAAACATATTTTACGTGGTGCTTGTCTTCTATCAAA
 GACTCTCTCAGTTGAAATATCTGAATTTTAACTATGAGGAATTAGATACGTTAACATACACTCTGCATCGAATTTTAA

(SEQ ID NO:70)

FIG. 17

ATGGTGAAGCTTTCCGTTGGTGACAAACATTTGGCCACGATCAGCGCGTCGCCGTGGTGGCATGGCCGTCATGTACGCCGGCTGCCAGAACACGACGAGTTTGGCAGAGCCT
 CCAGGGTAAGAAACATGAACAGAGAGCATACGCAGAAACCGCTGGGGTCCCGAGTACCGGAGGAGAGCACTTTAAGCCGGAGCGCTCGAAGTACAGCGACACCTTCTGCAAC
 GAGCGTTACGGCTGATCGACGAGAACTGATCGAGCAGCATCTCTGAAGCTGCTAAGACGCGATCGATACAAAGGCGAGCATCGACCTTAACAAAGACCG
 GCATTGCTCCGGCTCTGTTCCCTATGGATAACCTCAGGGGACCTTCTCAACCTCTACAGTGGCCATTGAGAAGAAGATCGGGCCGAACGCCCTCAAGGATGTCA
 ACCTGTGGTGAAGCGCACGACCAAGGTAAGAGACATAAGAGCCATCTGATCCCGCAGCTTCTGCTGAGCACTTGACATGGTGGCCCTCCACTACTCGCTGAC
 GCTGCTGGCTCCGCTCTACGCTCCGCCAGGACACCTCTCAGCGGTGCCCGACACCATGCTCTGGCGCTCGTGGCTCCCGAGGCCCTTTTTCATCC
 TTTCCGGCTTTTCGACCTTCACGCCATGCCCTTTTCGGGTGACGTGGGCCCTCTTCAAGAACAGAGCCAGGGCTCACTCCGGGCGAGGGCGGTGCTATCATGGTCTGA
 AGCCCTCAACGATGCCATTCCGCGACGGCGACCGCATCTACGGCAGCTCTGGGCGCGAGCTTTCAGCGGGTTGGCGCTCCCGCTCCCGCTCCCGCCACATGCCGTCCGAG
 TTCGACTGATGGAAGGCCCTCAGCGGTTTCAGCGCTCCCTCTCATCCAGTACGTGGAGTGGCACGCCACTGGCACCCCGAGGGCGACAAGGTGAGATCGAGCGC
 CATGACGAAGTGTCTCGGAGCATCTGCCCTGGCTTCGCCACGAGGTAACTTCGGCACACCTCTGAGCTTGTGGCTTTGGGGCATGTGCAAGGTCTCTCTCGAT
 GCAGTACGGTGAAGATTCCTCTACGCTGGCTGAGAACCCCGACAACTTATGCAGATCTTGTGTTACCGAGACTATTCCTGGCCGAACACCAACCGCGATCTTAAGC
 GTGCGGCTCAGCGCTTTGGCTTTGGCGGTACTAACGCCACCGCGGTTCGAGAGTACCGCAGCGACTTCAGGCCACAAAGACCTTGGAGAACGAGAGCAAGTCCAC
 GAGATCTTTCTCCCTTAAGATTGCCATTGTGGCATGGAGTCGAGTTGGCACTCTCAAGGCCCTCAGGAGTTCCAGCTGGCATCTAACCGCGGCCACGGCGGTGC
 GACTTCCGGGAACCGCTGGGCTTTCTCGGTGAGGACAAGGAGTTTCTCAGGCTGGGCTCCAGAGCTCCAGAGCTCCAGCGCTCAACAAAGGAGTGGAGACTGACT
 TAAAGCCCTTCGCTCCCATGATCCAGGAGGACATCTCCGCCCTCCAGCTCTCGCGGTGTCGATCAACCGCGCCCTCAACGCCAGCGGGCTTAAAGCCCAACGCA
 AGTTCGCTCTCGTGGGCTCGGCAACGATCTTGTAGCTTACCGCCACCGGCTCGCTGCCCTGAAGGAGCGCTTCAGACCGCGTCAAGGAGGACATCCGCCCTGTG
 GAGAAGCTCATGAACACCGCGGACCTCACGCTCTACACCTGTACATGGAACCTCTGTGCGACCGCTGATGCTGAGCGGCTTACACCGGCTTACACGGGCTTAG
 CTTACGATCACGGAGGCGAGACTCGGTTTACGCTTGGCTGACCTCGGCGGTGTTCTCGCCACCGGTGAGGTGATGCCGTGGTTGCTGGCTGGCTGATCTCTCGGG
 CTCGGCGAGAACCTGTTGTCAGTCCGCGCTCCAAAGTGTCCACCAAGAGGCCCTTGTAACTTGTAGTCAAGCGCGAGCGGTACTTTCGCCGGCGACGGGTGGG
 GTGCCCTGTTCTCAAGCGCTTTCGGACTGCATGACTCCACCGAGAAAGATCTACGCCACCGGTGACAGCAATTTGTCTGGCGGACGAGGTGGGCGCGACTATAAGCAGGCC
 CTGAAGAGCCTCGATCGCGGAAGGACATCGAGCTCGGAGCTCTCCGCTCCAGCGGCAAGCACCACTCCGGCCGCTACCTCGGAGGAGAGCTTAACGAGCTCG
 CGGAGATCTTCAACGAGGCAATTCAGCGGTGGCCATCGGCAAGCGTCAAGGCCAAGCTCGGCGACGTCCGCTACGCCCTCCGGTGTGCGAGCTCATCAAGAGCGGCCCTCTGC
 CTCTAACACCGCTACCTCCCAAGCTCCCACTGGAACAAGCCGACCAAGGAGTCCGAGTGGTGAAGAGCTTCTTGTCTGGAGCACTCGCGCTGCTCAAGAACGT
 GGACGAGAACCGCCACGGGTGTGAGCGGCTGTGCGAGAACGGCTCTGTGTACGGCATGTGTCATGAGCGAGTGCACATGAGGAGTGCACACCTGTGTCTCTC
 GATAAGAACGAGCCCAAGGTGCTCGGTATCTACGGGATTCCTCGTGAAGATATTTGTGTTCCAGCTGAACAAAGTACCTGGAGAACTTCTTCAAGGAGACTGGCACTGTCTCGGC
 TGGCAGAAAGTGAAGCCCTACCATTTGATCGACTCGAAGCTTTTCCGAGATGCTGAACCTTCCCGAGGACAAAGAACAAAGTTTGGCTCGCTCGGTACGAGCC
 CCAACAGCTCCAGCGCGAATTGAGCTCGCGTTAAGGGCATCCCTCGTGTGAAGGCCAAGCGGCACTGGTGTCTCCCTCCCGAGCAATTTTGGTGTGCAACCGGCTC
 AAGTCGGACAACATTGCCTTATGTACGCGAGGCGCGCTCGCTTACGCCGGCTCGGCTACGATCCACCGCATCTGGCCATGTTCACGAGCTCGTGAACACCGCAC
 GACTGACTGTGGACCAAGGTGACTGTGGTACTTCGCCGAGCTCTCCGTGGCCGAGAAGAGAGGTCTTTTGGCGACTTCGACAAAGAACAGATCGAGATGTTCGCG
 CTCGGTATTTTCTGCAGCATGTGCTTACCGCATGGCGAGGACTCTCGGCTTAAGCCGAAGGCCGCTTTCGGCTCTCCCTCGGCGAGATCAGCATGCTTTTGTCTTCT
 CGAAGAAGAACCAAGCTCTCAAGGAGCTTACTCGCCGCTCAAGGAGGCCAAGGTGTGGGCGTGCAGTGGCCGTGAGTTCGCCGCTCCGCTTGTGAACAT
 CCCGCGGACAAAGTCTCATGATGTTCTGCGAGGTTACTTCTGTTACCGTACCGCAACCTGCTGTGGAGAACACCAATTGCGAGAACAAAGTTGCTCCGCTCTTATCTGCTAA
 CGACTCCAGTCTGCTCATTTGCCGTTAAGCCGATGAGTGGCAAGAGTTCATCGAAGAGTCCACCTTAAGTCCCGCTCCCGCTACCCAGGGCATGATTGGCCACT
 GCGGAGGCCATTCCTACCTCGACGAGTACAGCACATCACGAGATGTTGAGATTCGCGAAGCTGAGTCAAGCTTTCACGAGCTTCGAGAACCGCGAGCTGTGTC
 TCGTGAAGAGCTCTGTAGCAAGCTGTGCGGAGATCTACAGCATCGGAGTCTCCCAAACTTGTCAACAAAGTCAAGGAGTCAAGAGCAATTTTCAATTCATCA
 GCTGGGAGCAACAACACTACGTTCCGGTGTCAAGACTATCTCGGTCCGAGATCGTAGAGTTCGATCGACCGCTCAGAACGAGACTGCCTGGGCGCAGCTCATGAAGA
 TGGTCCGAGCTGATCTCCACCGGCTCCCGGCTCGAGTCAAGAGTTCATCCGAGTTCGATCCCGAGGCCAAGCCCAACCGCTTATCTCCGCAAC
 ATCGAGCTCAACGGCTTTTTCGACCGCACGAACATCATCTGATAAGCAAGCTTTTCCCTCGGGACCCGAAGCTCGCCGAGATCTCTCAACAACCGCAACATGCCGAAGGATAA
 CTGTTACGTCCCATTGAGCGGCTCAAGAGTGAATCAAGGCGAGCGCTAACCTCAGGTGTCCGGTGTGAGGCGGTGTTACCGAGGATCTCTCTGGAGGACA
 ACCTCTTGAAGAGCTCTGGAGATCACTAAGCTTTCAGCGGTGTCAACGCTCACCGAGCCATCTCGCGGATTCGGGCTTCTCAAGACGTACGAGGTGACTACCCGC
 TCTACACCGGCGCTATGGCCAAGGGTATCGCTCCGCCGAGCTCTGTCATTGCGTGGCGGCGCTCGCTCTCAGGTGGTCTC

FIG. 17 (cont)

AGGATGCCATTAAAGCAGATCAAGGCTGAGCTTGGAACGGTCCCTTTGCCGTCAACCTCATCCACTCGCCTTTCGACCCCTCGCTTGAGAAAGGCAACGTTGACCTTTTCTCTCA
 AGTACAAACGTCCGCTTTGTCGAGGTGAGCGGTTTCATGAGCCTCACCCCAAGGTCTGCTACCGCTGCCGGCTTGCCAAAGGCCGTGACGGCTCGGTCAAGATTTCAG
 AACCGCATCATCGCCAAGATTTCGCGCACGGAGCTGGCGAGCTCTTCTCAAGCCCGCTCCGAAAGAACATCC'CTGATGCCCTCGTTGCCGACGGCTCGATTCTCCAGGAGCA
 GGCTCAGCTCGGCTCTGTCCTATGGCCGATGACATCACCGTTGAGGCCGACTCCGGTGGCCACACCGACACCGCCCCATTTCATGTGCTCTCTCCCTCATCATCCAGCA
 GCGCAACCGCATTTTGCAAGCAGTAC'CCGAAGCACCTCAAGGTCCGCAATCGGTGGCTTAAGCGGCTTTTGCCGCCCTTGAGATGGGTGGGCT
 ACATCGCCACGGGCAACCGTTAAAC'AGCTCTCGAAGGAGGCCGCACTGCGACTACGTGCGCAAGGTGCTCAACAGGCCACCTACTCCGACGTCAAGATGGCTCCCGCTGCC
 GACATGTTTCGACCACGGTGTGAGCTCCAGGTTCTCAAGAAAGGGCAACCATGTTCCGTGGCGCCAAAGAAAGCTCTACGACCTCTTTAAGAAAGTACAAGTCGATCGAGGAGCT
 CCCTGCCGACGAGGTCAAGAAAGCTGGAGCAGAGGTTTAAAGAAAGTCTGACGAGGTCTGGGACGAGACTAAGAACTACTACATTAAACCGCTCCACTCCCTCGAGAAAG
 ATCGAGCGCGCGGAGCGTGACGCCAAGCTGAAGATGTGCTCTGCTTTGTTGTTACCTGAGCAAGTCGTCCTGGGCCAACACCGGCGAGTCGGGCCGTGTGTCAGGACTA
 CCAGATCTGGTGGGCCCCGCCATCGGCTCGTACAACGACTTCGCCAAGGGCTCGCCCTGACCTGAGATCCTTGCTGCTGTTCCCGTCCGATCAACAGCA
 TATTCTGCGCGGCTTGCTTCTACCAAGCGTCTTTCGCAAGTCAAGTACCTTAAC'TTCAACTACGAGGAGCTCGATACGCTCACCTACAGCGCTAGCAACTTTATCTAA

(SEQ ID NO: 121)

FIG. 18

MVKLSVGDNICHQQRVA VVGMAV MYAGCQNHIEFWQSLQGNMNSKSISQNR LGSFYREHFKPERSKYSDTFCNERYGCIDENVQSEHELLLKLAKDAIADTKGSIDLNKTGIVS
GCLSPMDNLQGDLLNLYQCHIEKKIGPNALKDVNLWSKRTTNGKDDKKAYFDPASFVAEQDMDGPHLYSLDAAACASALYVLRLAQDHLLSGAADTMLCGASCLPEPFFHLSGFSTF
HAMPLSGDVSAPLHIKTSOGLTPEGGAIMVLRKLNDAIRDGRIYGTLLGAELNAGCGPLSPHMPSEFDCMEKALQSVHRLPSSIQYVECHATGTPQGDVKEIDAMTKCFGEHLPR
FGSTKGNFGHTLVAAFGAGMCKVLLSMQYGEIPPTGLEPNIMHDLV VTE TIPWNTNGDLKRACLSAGFGGTNAHAYFEYRSDLQANKTLNENESKSHIESSFKIAIVGMESEF
GTLKGLQEFERAIYNGGHGACDLPENRWRFLEGEDEKFLQACGLQKLPKRGCIKEVETDFKRLRLPMIQEDILRPLQALLAVSHIDRALNASGVKPNKGVAVLVGLGTDLELYRHRARVA
LKERLOTAVKEDIP LLEKLMNYVNDRGTSYSTSYIGNLVATRVSSLW/GFTGPSFTTTEGENSVYRCLDLGRWFLANGEYDAVVVAGVDLCGSAENLVFKSRKSVSTQNEPFANFES
NADGYFAGDGC G GALVLRKLSDCDSTDEKIVATVDSIAVGDDEVGPTIKQALKNASAAKDIELAE LSASSGKHHSGRITCEDELNELGEIFNEGHRVAIGSVKANVGDVGYASGAASLI
KTALCLYNRYLPKLPNWNKPTKDVESKSFVCEHSRAW/LKNVDENRHA VVSGVCENGSCYGVMSDVQGHIIEENLVSLDKNEPKVLYGIYGDSDVDILYQLNKYLEKFLQETGT
AAAAQVKSP TIDIDSNVFAEMLNPQDKNKKFAVALVTPNKLQRHELA VKGIPRCV KAKRDWCSPSGSIFACNPLKSDNIAFMYGEGRSPYAGLYDHLRWPMMLHELNNRRTTE
LWDQGSWYLP R SSV AEKEKVFEGDFDKNQIEMFRLGIFYSMCFDMMATELLGLPKAAFGLSGHSIMLF AFSKKNTKLSKELTRRLKEAKVWASQLAVEFAAIRDLWNIPADKSID
EFWQGYFVYANRTLVENTIGENKFVRLIIVNDSQSCLIAKPEDCEQKVIEKLHLKLPAPVVTQGMIGHICPEAIPYLDQISHIHEMLEIPKPVNPKLFTSENRELYSMKDSVSKLVAEIV
QHVA DFPNIVNKVKETCKTDIFIELGNNYRS GAVKTLGPEIVSAIDRQNETAWGQMKMVASLISHRVPGVELKKLYHPELKFDPQAKPNRFRINIELNGFFDRTNIVDKQLSPA
DPKLAIEIVNRRNMPKDNVYVPIERVKTMIKAE PANLQVSVGSKPVVTERISSDDNLFELKSEITKSF DGVNACTE/AMLGDSGFLKTYEVDYPLYTGAMAKGIASADLVIAAGSKILA
SFGAGALQVVEDAIKQKAE L GNGPFAVNLJHSPDP SLEKGNVDLFLKYNVRFEVS AFMSLTQV VYRAAGLAKARDGSVKIQNRRIAKISRTELAE LFKPAPKNILDALVAD
GSISQEQALQALLVPMADIDITVEADSGGHTDNRPJHVLPLIHOORNRICKQYPKHLKVRIGAAGGIGCPKAAFAAFEMGAAYIATGTVNQLSKEAGTCDYVRKV/LNKATYS DVTMAP
AADMFDHGVELQVLLKKGTMPPSRAKKLYDLFKKYKSEELPADEVKKLEQKVPKKSFEVWDETKNYINRNLHSPEKIERAERDAKLKMSLCFRWYLSKSSRWANTGESGRVQDY
QIWCCGPAIGSYNDFAKGSPCLDPEILGSFPSVVVQINKHILRGACFYQRLSQLKYLNFNYEELDTLTY/SASNFI

(SEQ ID NO:71)

(SEQ ID NO:72)

ATGGTTGGTTTACAAATGAAAAAGAACCCAGTATGGGAGATGAGTAAAGGAAGAAATGTTGTATTGTGACTATGATGAATGTTGGAAATTTGCTGAAG
CTGATATTAATGAAGTCTTGGACCTAAAGTTTGATATTTATCGATAGTATAGTTCAGCTGAGTAAATCTTCTAGTACAGAGTATCTTGTAGTGAATCTG
CTGAAGTTGGAAATTTACAGATTGGACTAGAATGGTACTGAATATGATGTCCAGTAAATGGTGAACTTTCAGAAAGTGGTGATGTTCCATGGGCTGTTCTGTGTGAATCTG
GACAAAGIGATCTTAAGTTAATATCTTAATAGGGATATGATTTTCTAAGTAAAGGTGATCGTGCTATCGAATTAATAAATACACGTGTGACGTTTACGGTGTGCTCAATGAGG
TGAACAACACTAGTATACGATATTCGTGTAACCTGATTTGCAAAAGGATATGCAACGGTGAATCTCCATGTTTTTTTGTGAATGATGTTGATGTAAGTGAAGTAAATTAATTCGA
ATAGAGAGATGGTGTGCGGGATTTTTTCTGATGAAGAACTTGCAGAGTAAAGAGATTATTAACACTGTCTGGAACCTTCATAAAAAGAAATCTATTTGCTCCAAAATCCA
TTAAGACCTTTTGCTCTAAATCCAGCTACAACAAAGATTTCTGGAATAATGATATGGAATAATTTGTGAGCTCAATGGGAATAATGATTTGGGTATGGTCAATCAAGGT
ATTGACTACAAGTATGTCACGGAAATGCTTATGATTGCTATTAATAACACATATGTTGGTGCATATGGTCTGGATTATGGTGTGGCAAAAATCTTGAA
CGTGACTATTGGTATTTCAACCATTTTGAAAGGATCAAGTTAGCTGGCTCACTTTAGTAGTGTTGCACTACAGCTACATAAACTTTATATGTTATGTTGGTGTAC
ATGATGGTGTCCAGATTTTCAATTTCCGTCCAGTTCTGGACACCAATAAAGATTGTTGCCGTGGACAAATATGATCCCACTGTGGTAACTTGTATTGTTATGGAATAATA
GAGAAATGGGATTCATGATCAATCTGGACACCAATATGCTATTTGCTGATTTGATTTGATTAATGATTAATGAACTTTGGTCAATCATGATATGGCTGATATGGAATGATT
ATGACGCTGGTAATTTGCTCAAAAGAAATTTGGTGTGATTTTAAAGGAATTTGTTGCAATGGAAGTACCGTGAATCAATAATCATATGATCTCTACCAAAAATCAACTA
TTATACAACTCCAAATGCTTCTGGTGATCCATGGCACCAATCAACAGATGATGATCCAAATGGCAGGATTAATGGGCAAGTAAATGGGCACTCTTCAATTAGTCAATCTGT
ATTATCCACCACTGCTGTTGTCTCAACCAATCTTGGTATCTTTPAGATAACGATCATACACTGTGTAATAATGCTTTAACTGGTGTAAATATGTCGAGTTATGTGTG
TAAAGTATCAAAATGCTTTGGACCAAGAAATTAAGAGATTTGATAACTCTAAACATCCAGAACTCTGCTTGTATCTTGC ACTTGTACAGGTTGTGAGTATCAGATAT
GGAATTTAAACCTCAATTAATATGATGATATCCCAATGAAGGTACAATGATGAGTGAAATTTGATGCCCTGACATGCTGGTGTTTTTCAGGATCATGTAACCAATGGTCA
TATGCCGTATCTATTGTTATGGAATAATGCTCTTCAACTCTGGTGTATTAACCTTGAAAGCACTTGTTGACTATGGATAAAGATGATATCTTTTCCGCAAAATTTGGAT
GCCACTGCTGAATGTTCCGAAGTGATTTGATTTGAGAGTAAACACTATCAAAACTTTACTCAATGACCGGTTACAGTATGCTCGGAATAATGGGAATTCATAGATTCAC
ATTGTAATGATCTGTGATGATGATGTTTCTCAAGAGGATCAACATCTTTGGTGGTCTACCCCTGAAGTATTCGAGTCAAGGTTGGTGTGATAAATGGTAAAAAGGTACAA
ACCATGGTATTGGAAACAAAATCATCTAATGATGATACTATGACTTGGCTCACTGCTGGCAAGGATAAGTTATTTCAAAAGATTGGATTAAGGATGCACAAGTCAAA
GAGAAATACAAATGATGATTTAGATACTATGATATTTCCAATATCTGAAATGATACAAAGGTTATGCTCATGGAAAGAAAGAAATTAATCCAAACGACTGGTCTC
TTTTTCTGTCAITTTCTGGTTGATCCTGTGATCGCTGTTCAATAGTATTGAAGTATGTTTCACTATTGAAGCATTTTCAATTGATCAAGGAATCGCTTCAAAACATGGTA
TTTGGAATCCAACTTTGCTCATITCCAATGGGAACAACTTCTTGGAATACAGAGGTAATTAAGTAATAAAGGTAAGGATGAGTAAGAAATTAATATCAAGAAATATGTTCT
AAAAATGCTGATGTTGATGTTGATGTTTATGTTGGTGAATTCACATAAGATATACTGCAAGTATGCTGCGGTGTAATAATGTTGACCGGGAACCAAAAGCT
GCCACTAAATAGTAGTCTGCTCCAAGACATGTGGCAACCAATTCAGGAGTGCCTTCGAATACAGCAGTGTTGAAATCAGTTTGGAAATCTTTGAAAGAAAGAAATGHT
AAMCTTGAGAAACCAATGATCTTGGAACCTTCCAATATTTGTAACCAATCTGGTGAGTAAATATGGCAAGCATCCGTATTCCACCATGCACTCAATGATTTGGTGG
TGACGTGATTTTATGGAAACATAACAATGTTGTGCAACCTTACTGGACCATGGCTAAAGHATTGCATGCTGATTATGCTGATTTGGTAAATTTGCAGCTGGTGAAGAAATTT
GGGTCTTTGGCGCTGGAGCTTACCAATGCATCTGTTCTGTCTTGAAATAATCCAGCGCATCTCAGAAAGTCCATACGCTGTCAACTTGATTTATAGTCCAT
CGACTCAAACTTGAAAAGGGAATATGATCTATTTTGGAAAAGGGTGTTCATGTTGTGGAAGCATCTCATTCCTGCTCGACCACTCAAGTAGTTCGTTACCGTGCATG
TTGATTTATCTCGGCTCAAGACGGATCTGTAATGATCAAAAATAGAAATCATCGCTPAAAGTTTCAAGGACCAATGGCTGAAATTTTTCAGACCTGCACCAACAAATGTTGTA
TGCAAAAGCTTTGCTTAGTGAGAAATCACTAAAGAACAGCTCTATGCTCTTGGAAGTACCAATGCTGATGATGCTGTGAAGCTGATAGCGTGCGACATACTGATGA
ATAGACCAATTCATGTAATCCTACCTTTGATTATCAATCTACGAATAAGAAATTCATAAAGATGGTTTTCTGCTGCTTTGAGAGTTGCGTTGGTGGTGGGAAATGTG
TTTGTGCCAAGTGTCTGATTTGCTGATCAATAATGGAGCTGCACTGTTCTTACTGCGAGCGTCAACCAAGTTAGCAACAATCTGTTACCTGTGATATCTGTTAGAAAGCAAT
TATCTGAAGCTTCATAGATATTACCATGGCACAGCGCTGATATGTTTATGATCACTGAGTTCAGGTTCAAGTATTAAAAAGGAACATATGTTTCTCATCTGCTGCAAGA
AAATGTATGAATTTATCTGTATGTACAACTATTGTATGATGATGCCAAAAGCGAACTCTAAAGACTAGAAAGCAAGAAATTTTTCAAAAATCGCTTCGGAAAGTTCGGGAAAGA
ACTAAAGATTTTATATCAATCTGTGTAATAAATCTGAGAAAGATTGAACATGCTGAGAAAGAAAGATCCAAAGTTGAAGATGTCATTATGCTTTAGATGGTATTGGGTTTAAGT
TCAATTTTGGGCAACAATGGAATTAAGAAAGATCAATGACATTTGGTGGTGGCAGCAATGTTTCATACAATGATTTTGTAAAAAGGAACATTATTTGGATCTGTCGA
TGACAGGTTTCATATCCATGTTGTCTCAAAATTAACATGAAATCTACCGGCTGTGTTTCTTCAACGAGTTCGTGCAATCAAGCATCAAGTATTTGATTTGATGTC
ATGAGAAGATGTTTACCTATCGTCCAGAATCAACCCCTATAG (SEQ ID NO:72)

ATGGTTGGCTGCGAGATGAAGAAGGCTGTGTGGGAGATGTGGAAGGAGGAGAGTGTCTCGGCAAGAAAGTCTCTTGTACTACGACGAGCTCTCTGAATTCGCGGAGG
GCTGATCGCAAGGTGTTTCGGCCCAAGTTTGACATATCGAAGAGTACAGCCGCTGTGTGCTCCGCGGAGTACCTCTCTGTCACCGGTGACCGCTCATGTGAT
GGCGAGAGTCTGGCAACTTCCGCTCGCGCATGGTCAACCGAGTACGAGTCCCGGTGAACGGCGAGCTTTCCACAGGCGGCGAGTTCCTCGGCGGTCTCGCTCGAGTG
GGGCGAGTGGAGACTCTGTTATCTGTACATGGCAATGACATTTTACGTGAAGGGTGACCGGTTTACCTCCGCTTCTAACACGACCCCTACGTTTACGGTGTGCGCCAGA
CGAGATGCGGACGGTGTGCGTGTGCTTCTACGGAGGAGGAGTCTGCGCGGCAAGGGAGTCTATCAAGACCCGCTGTGTGAGTCTCAACAAGTGTATGTGCGCAAG
TCTGATCAAGCTTTTGGCTCAACCCCGGCTTCAACAGAGATGTTCACGAGAACATGTGAGAAGCTTTTCGAGCGCCAGTGGGAGAACGTCCTCGGCTCGGCGCTCA
GGGGATCGACTACAACTGTGCGCGCCAGAGATGCTCATATGACCGCGCATACAGAAATCCAGCAACGGCGGTGCTACGGCTCGGCTCTCTGTTCGCGAGAGAAGATTG
TTTGAGCGCGACCAATTGGTATCTCCCTTGGCACTTCTGCAAGGACAGGTGATGGCGGGCTCCCTGTGTAGCGAGCGGTGCTCGACGTGCTCAAGCTTTACATGCTTTGGCTCG
TTTGAGCGCGACCTGTGCTCCGATTTCTAGTTCTGCTCGGCGAGCCCAACAAAGCCAGCTCGGCTGGCCAGATCAACGCCCATGTGGCAAGCTGTGTACGTGTACGTGATG
AGATTCTCGGAGATGGGTTTCAAGAGATCTCACGGCCACGCCCTACGCGATCGCTGACGTTGACATTTGTAACTAGAGCTCGGCCAGTCTCTTGACATGGCCGACATC
GACTCTGATCGGCCGTGGCAACCTCTCAAGAAAGATGTGTGATATTCAGGCAATTTGCGCTCCAGATGAGGAGCCACCTGTCAAGAGCTCAACATCAATGATTCGTCCTCCCA
CTGCTACGATTATCAAGCGCGCCCAACTCGGCTCGCGGATCTCTCGGCCCTCGAGGTCAACGTGACCGGCTGCTGACGCGCCCGCCCTCTCAG
CCCGTCGGATTACCTCTCTGTGCGGTTTGTTTAAAGCCTTCTCTGGCAACCCCTCGACAACGATCATACGCCGGCAAGATCGCCGTGACCTGGTTTAAATCATGTCCGGAGTT
TATGTGCGCAAGTCTGCTGGCCCTGAGTTTAAAGCTGTTCGAAACATCAACGCTGACAGGCGCAAGATGATTGGCGGAGTCTGACCTGCGCGGTGACGCTGTGTTCTCTGCA
CAAGGATATGAGGTTCAAGCCCACTCAACATCAATCGATCAACCGCTGACAGGCAAGTGAATGGCGGAGTCTGACCTGCGCGGTGACGCTGTGTTCTTACGGGCTCTCTCA
ACGAGGGCCACATATCGCTACAGATCGTCAAGAGATCGCCCTTACAGCAAGCGGTGTCTCACTCTGCTCAAGGCCCGCTCACTATGGACAAGGACGACATCTCTTTCT
CAAGCACTCAACGCCACCGCGAGATGCTGCTTTCGAGTGTTCGAGCTGTGCGGTGAAGCAATTCACAGGTCAACCGGTACAGGCTGTGCGCAAGATGGGCAATC
CAACGGTTCATTTAGCTCTCGGTGATGACGCTGTGTTTCAAGAGGCTCGACCGGCTGCGACCGGCTGAGTGTGTTGATTGCTGCGCAGGTCCGATACAGCG
AAGAAGGHTCAACGGGTATCTGAGCAAGAGCTGTGACGATGTGCTTGCATGCCGCGGCAAGGATCAAGGCTTCTTCTGAAGATCTGGCTGAAGAGG
CTAGGTCAGCGTCTCAACACAGGTGAGTTCTCGACAGTGCACATTATTCGAAACACAGTACAAAGGGTACCGCGACGTGAGAGAAAGGTCAAACCC
CAAGCACTGCTTCTCGCCACTTTCTGTGCAACCGGTGATCGCCGCTCTCGGTATGAGTGCATGTCTCAAGTCAAGGCTTTTTCGATGACAGGAGTATCGG
TCCCAAGCAATGTCATCTCGACACTTTCGACACTCGAACCGGCTGAGAGTACCGGCGAGCTCAACAAAGGGCAAGCGATGACAGGAGATTCACA
TCAAGGATATTGTCAAGAACCGGACCGTACTGTGATCTCATCTGCGCAATGTTTCTCTGTGGACTGTCTGCGCTTACAGCGCGATGACCTCCGCTCAAGATCTGCTT
CCGGCACTAAGGCTGTCCCAAGAGCTCGCGCGCTCGGCACTGTGGCCACTCCGATCCCGGCGTCCCTCAACACTCTCTGGTGGAGATCTGCTTGAATGCCCTTA
ATTAGAGAGCTCTCAACTCGAAGAGCCCTCTACTTTGAGACTTCAACCAATCTGAAGCTGTGGCAAGTAAACGCGGATGGCAATCCGCTCATCTCCGTCGACCC
ATTAAAGATCTCGGTGAGCGCTCGTTTATGAGACTTAAACGCTGCTGCTCCCTCTACACCGCGGATGGCAAGGCAATCGCTTTCGGGGAACCTCTGTCATCGTCCCGGT
AAGCGCAAGATCTCGGCACTTCGGCGGTGGGCTCCCGATCTGTGCGACCTCTGGTGAGAGATTCAGCGCGCTTCGAGCGGCTCGCGGTACCGGTGCAACCT
CTGCTACTCGCTTCTCGACTCGAACTCGAAGAGGTTAAGCTGACCTCTTTCGGAGAAGGGCTCGACGTGTGAGGCTCTGCGGCTTACGGCCTCAACGCCAGGTCGT
TCGCTACC GGCTCGGCTCTCGGTGCTAAGGACGGTTCGTGTGATTAAAGAACCGCATCATCGGTAAAGTACGCGGACGGAGCTTGGCAGATGTTCTTTCGCGCTCG
CGCCCGAGAACTCTCTGATGAATGCTATCGCCAGCGGAGTACUAAAGAGCAAGGCTCTCTGCTGTGAGCTGTATGGCGACATGTCTGCTGAGCGGACTCGG
GGCGCCACACCGATAACCGTCCCATCCAGCTATCCCTCCGCTGATTAACTCACTCCCAAGGATGGGTCTTCTCGCCGCTCTCCGCTCGCGGCTCGGCTCGG
CTGTGTGTGGCATCGTTGCCCCCTCGGCGCTGTGCGGCTTCAACATGGGCGGGCTTCTGTATACCGGCTCCGTTTAAACGAGTGAAGCAAGCTCGCGGACGTGGCAG
ATTGTGCGGCAAGCACTTACGAGGCAAGCTACCGACTCAAGTGGCTCCCGCTGACATGTTCGACAGGCGGTGGAGTCCAGGTCTCTAAAGAAAGGTACAGATGT
TCCCTCGCGCCCAAGCTACAGACTCTTTTGATGTACAAAGCTTTGACAGCAATGCGCAAGTCCGAGCTCGAAGAAAGGCAATTTTCAGAAGAGGCTCG
CCGAGGCTGTGGAGGAGATAAGGACTTTTACATCAACCGCTCAACAACTCGGAGAGATCGAGCACCGCGAGAAAGGACCCAGGCTCAAGATGTGCCCTTGTGCTTTCG
TCTGTATCTCGGCCCTTCTGAGCTTTTGGGCCAACAAACGGCATCAAGGAGCGCAGCATGTGATTACAGATTGTGTGCGCGCCGCTATTTGGCAGTCTGTGAAGG
GCACCTACTCTGACCCCGCTGCGGTCTGCTACCTCTGCTGGTGTGAGTCAACATGAGTACTCTTCGAGCGGTCTCGGCCATTAAAGCAGGAC
(SEQ ID NO: 122)

FIG. 21

MYGLQMKKKPVWEMSKEEQSSGKNNVFDYDELLEFAEGDIGKVFPGPKFDIDDKYSRRVRLPAREYLLVTRVTLMDAEVGNFRYGSRMVTEYDVPVNGELSQGGDVPWAVLVESGQ
CDLMLISYMGIDFQCKGDRVYRLNLTTLTFYGV AHFGFTLVYDIRV TGF AKGMHGFI SMFFFYDCYVNGRLLJFMRDGCAGFFTDFELAAGKGVKTV AFLHKKRSIVPKSIKPFAL
NP AVHKTMFSENDMEKLCERQWENVLGSGLQGIDY KLCARKMLMIDRITKIQHNGGAYGLGLLVGEKILERDHWYFPCHFVKDQVMAGSLVSDGCSQLLKLMLWGLHDVVPD
FQRPVPQPNKVRCRGQISPHRGKLVYVMEIREMGFNESTGQPYALADVDIIDVNYELGQSFDMADIDSYGRGNLSKKIVVDFKGIALQMEGYTKSSNIIDSSPKSTIIQPPNCLRGDP
LAPSQVTHPMAGVNGAPAPSPSPDYPPRAVCFKPPGYNPLDNHTPGKMPLTWENMSEFMC GKVSNC LGPEFKREDNSKTSRPAFDLALVTRVVSVDMEFKPHLNIDVNP SKG
TMIGFDCPADAWFFQGCNDGHMPYSIVMEIALQISGVLTSVLKAPLTMDKDIDILFRLNLDATAEMVRSVDVDCRGKTIKNFTQCTGY SMLGKMGHRTFTELSVDVVVYKGSISFG
WFTPEVESQVGLDNGKKVQPYWYLEQKSSNVVTYDVASTAGDKLFSKIGSKDAQVQRNTOCEFLDTMHIIPNTGKYNKGVAHGEKKVNPNDWFESCHFEDVPVMPGSLGIESMF
QLIEAFSIDQGLASKHGIVNPTFAHNSGKTSWKYRGQLNNKGRMDSEIHKIDIVKNADGTVDLIADGFLVDSLRYVSADDLRVKIVPGTKAAPKSVAAAPRHVATIPGVPSNTSSV
EISLESLLKELLNLEKPLYLETNSHVKQFGDVNNGQASVDPCTINDLGFERSMEFTYNNVAPLYTGAMAKGIASADLVIAAGKRKILGSFGAGGLPMHLVRASVEKIQALPEGPYAV
NLHSPEDSNLEKGNVDLEKGVHVVEASFTALTQOVVRYRACGLSRADGSLVLIKNRIIGKVSRTELAEMFRRPAPQNLLDKLIASGEITKEQASLALEVPMADDVAVEADSGGHT
DNRPHVILPLIINLRNRIHKECGPPAALRVRVGAGGIGCPSAAVAAFNMGAALITGSYNQVSKQSGTCDIVRRKQLSEASYSDITMAPAADMFQGVLEQLVKKGTMTFTSRACKLYE
LFCMVNSFDDMPKSELQRLFKRIFQKSLAFVWEFTKDFYNRLNNPEKIFHAFFKKDPKLMKMSLCFRWYLLGLSSFWANNIGKFRSMDYQIWC GPAIGSYNDFVKGTYLDPVAVAGSYPC
VVQNMQLRGACFLQVRRAIKHDPRLDIDVDEDDVFTYRPESTL

(SEQ ID NO:73)

FIG. 22

Uso de Codón de *Schizochytrium*

AA*	Codón	Fracción	AA*	Codón	Fracción	AA*	Codón	Fracción	AA*	Codón	Fracción
Ala	GCC	0,84	Terminal	TAA	0,34	Leu	CTT	0,18	Ser	TCG	0,33
Ala	GCA	0,03	Terminal	TGA	0,33	Leu	TTG	0,02	Ser	TCC	0,31
Ala	GCT	0,18	Terminal	TAG	0,33	Leu	CTG	0,12	Ser	AGT	0,03
Ala	GCG	0,16	Gin	CAA	0,08	Leu	CTC	0,69	Ser	TCA	0
Arg	CGG	0,01	Gin	CAG	0,92	Leu	TTA	0	Ser	TCT	0,09
Arg	AGA	0	Glu	GAA	0,09	Leu	CTA	0	Thr	ACG	0,3
Arg	CGC	0,8	Glu	GAG	0,91	Lys	AAA	0,04	Thr	ACC	0,54
Arg	CGA	0,01	Gly	GGA	0,1	Lys	AAG	0,96	Thr	ACA	0,02
Arg	AGG	0	Gly	GGT	0,2	Met	ATG	1	Thr	ACT	0,14
Arg	CGT	0,17	Gly	GGG	0	Phe	TTT	0,45	Trp	TGG	1
Asn	AAC	0,94	Gly	GGC	0,7	Phe	TTC	0,55	Tyr	TAC	0,94
Asn	AAT	0,06	His	CAC	0,83	Pro	CCT	0,21	Tyr	TAT	0,06
Asp	GAT	0,24	His	CAT	0,17	Pro	CCG	0,34	Val	GTC	0,62
Asp	GAC	0,76	He	ATC	0,7	Pro	CCC	0,43	Val	GTA	0
Cys	TGC	0,95	He	ATA	0	Pro	CCA	0,02	Val	GTT	0,14
Cys	TGT	0,05	He	ATT	0,3	Ser	AGC	0,24	Val	GTG	0,24

* AA=Aminoácido