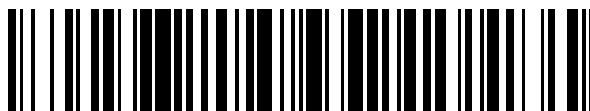


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 688 953**

51 Int. Cl.:

C12N 1/12 (2006.01)

C12N 15/79 (2006.01)

C12P 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.04.2002 E 10178886 (7)**

97 Fecha y número de publicación de la concesión europea: **01.08.2018 EP 2325319**

54 Título: **Producto y proceso para transformación de microorganismos de Thraustochytriales**

30 Prioridad:

16.04.2001 US 284116 P

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

07.11.2018

73 Titular/es:

DSM IP ASSETS B.V. (100.0%)

Het Overloon 1

6411 TE Heerlen, NL

72 Inventor/es:

ROESSLER, PAUL G.;

MATTHEWS, T. DAVE;

RAMSEIER, TOM M. y

METZ, JAMES G.

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 688 953 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

Producto y proceso para transformación de microorganismos de Thraustochytriales

5 Campo de la invención

La presente invención generalmente se refiere a un método para producción de una proteína, que comprende: cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína; y recuperar la proteína. La invención también se refiere a un método para producción de una proteína, que comprende: cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína, para producir la proteína; y extraer el microorganismo que comprende la proteína.

15 Antecedentes de la invención

El desarrollo han dado como resultado la revisión de la taxonomía de los Traustoquítridos. Los teóricos de la taxonomía sitúan a los Traustoquítridos con las algas o los protistas similares a algas. Sin embargo, debido a la incertidumbre taxonómica, en el contexto de la presente invención lo mejor podría ser tener en cuenta las cepas que se describen en la presente invención como Traustoquítridos (Orden: Thraustochytriales; Familia: *Thraustochytrium*, *Schizochytrium Labyrinthoides*, o *Japonochytrium*). A continuación se resumen los cambios taxonómicos.

Las cepas de ciertos microorganismos unicelulares que se desvelan y se reivindican en el presente documento son miembros del orden Thraustochytriales. Los Thraustochytriales son eucariotas marinos con una historia taxonómica problemática. Moss (1986), Bahnweb y Jackle (1986) y Chamberlain y Moss (1988) han revisado los problemas con la colocación taxonómica de los Traustoquítridos.

Por fines de conveniencia, los taxónomos colocaron al principio a los Traustoquítridos con otros eucariotas zoospóricos incoloros en los Ficomisetos (hongos similares a algas). El nombre Ficomisetos, sin embargo, se eliminó con el tiempo del estado taxonómico, y los Traustoquítridos se mantuvieron en los Oomicetos (los hongos zoospóricos biflagelados). Inicialmente se supuso inicialmente que los Oomicetos estaban relacionados con las algas heterocontas, y con el tiempo una amplia gama de estudios ultraestructurales y bioquímicos, resumidos por Barr (1983) apoyaron esta suposición. Los Oomicetos fueron de hecho aceptados por Leedale (1974) y otros ficólogos como parte de las algas heterocontas. Sin embargo, por razones de conveniencia que resultaban de su naturaleza heterotrófica, los Oomicetos y Traustoquítridos han sido estudiados en gran parte por micólogos (científicos que estudian hongos) en vez de por ficólogos (científicos que estudian algas).

Desde otra perspectiva taxonómica, los biólogos evolucionistas han desarrollado dos escuelas de pensamiento generales en cuanto a cómo evolucionaron los eucariotas. Una teoría propone un origen exógeno de los orgánulos unidos a membrana a través de una serie de endosimbiosis (Margulis 1970); por ejemplo, las mitocondrias se obtuvieron a partir de endosimbiontes bacterianos, los cloroplastos de cianofitas y los flagelos de espiroquetas. La otra teoría sugiere una evolución gradual de los orgánulos unidos a membrana a partir de los sistemas no unidos a membrana de la especie primitiva procariota por medio de un proceso autógeno (Cavalier-Smith 1975). Ambos grupos de biólogos evolucionistas, sin embargo, han retirado a los Oomicetos y a los Traustoquítridos de los hongos y los colocan o bien con las algas cromofitas en el reino Chromophyta (Cavalier-Smith 1981) (este reino se ha ampliado más recientemente para incluir otros protistas y en la actualidad los miembros de este reino se denominan Stramenopiles) o bien con todas las algas en el reino Protoctista (Margulis y Sagan (1985).

Con el desarrollo de la microscopía electrónica, los estudios de la ultraestructura de las zoosporas de dos géneros de Traustoquítridos, *Thraustochytrium* y *Schizochytrium*, (Perkins 1976; Kazama 1980; Barr 1981) han proporcionado pruebas válidas de que los Thraustochytriaceae están solo relacionados de manera distante con los Oomicetos. Además, los datos genéticos que representan un análisis de correspondencia (una forma de estadística multivariante) de secuencias de ARN ribosómico 5 S indican que los Thraustochytriales son claramente un único grupo de eucariotas, completamente separados de los hongos, y lo más estrechamente relacionados con las algas rojas y marrones, y con miembros de los Oomicetos (Mannella *et al.*, 1987). La mayoría de los taxónomos están de acuerdo con retirar a los Traustoquítridos de los Oomicetos (Bartnicki-Garcia 1988).

En resumen, usando el sistema taxonómico de Cavalier-Smith (1981, 1983), los Traustoquítridos se clasifican con las algas cromofitas en el reino Chromophyta, (Stramenopiles). Esto los sitúa en un reino completamente diferente de los hongos, que se colocan todos en el reino Eufungi. La colocación taxonómica de los traustoquítridos se resume por tanto a continuación:

Reino: Chromophyta (Stramenopiles)
 Filo: Heterokonta
 Orden: Thraustochytriales
 Familia: Thraustochytriaceae

Género: *Thraustochytrium*, *Schizochytrium*, *Labyrinthuloides*, o *Japonochytrium*

Algunos taxónomos pioneros separaron unos pocos miembros originales del género *Thraustochytrium* (aquellos con una fase de vida ameboide) en un género separado denominado *Ulkenia*. Sin embargo, en la actualidad se sabe que la mayoría, si no todos, los Traustoquitrídos (incluyendo *Thraustochytrium* y *Schizoclaytrium*) presentan fases ameboides y, como tal, algunos no consideran que *Ulkenia* sea un género válido. Tal como se utiliza en el presente documento, el género *Thraustochytrium* incluirá *Ulkenia*.

A pesar de la incertidumbre de la colocación taxonómica dentro de clasificaciones superiores de Filo y Reino, los Traustoquitrídos siguen siendo una agrupación distintiva y característica cuyos miembros siguen pudiéndose clasificar dentro del orden Thraustochytriales.

Los *Schizochytrium* y otros microorganismos de Thraustochytriales presentan un valor comercial actual y potencial sustancial debido a su capacidad para producir grandes cantidades de compuestos lipídicos, incluyendo ácidos grasos altamente insaturados (HUFA) y diversos carotenoides (por ejemplo, astaxantina). Los ácidos grasos altamente insaturados omega-3 son de interés comercial significativo porque se han reconocido recientemente como compuestos dietéticos importantes para prevenir arteriosclerosis y cardiopatía coronaria, para aliviar afecciones inflamatorias y para retrasar el crecimiento de células tumorales. Estos efectos beneficiosos son un resultado tanto de los HUFA omega-3 que provocan inhibición competitiva de compuestos producidos a partir de ácidos grasos omega-6, como de los compuestos beneficiosos producidos directamente a partir de los propios HUFA omega-3 (Simopoulos *et al.*, 1986). Los ácidos grasos omega-6 son los HUFA predominantes que se encuentran en plantas y animales. Por tanto, el desarrollo adicional de microorganismos de Thraustochytriales como organismos de producción comercial se beneficiará significativamente de la capacidad para producir cambios genéticos específicos en los organismos por medio de tecnología de ADN recombinante, incluyendo la potenciación de la producción de HUFA y carotenoides y altamente valiosos por tales organismos. Además, la capacidad para lograr una mejor comprensión de la bioquímica y biología molecular de este grupo escasamente caracterizado de organismos proporcionará una información valiosa que se puede usar para guiar los futuros esfuerzos de desarrollo de cepas. Antes de la presente invención, sin embargo, los métodos y construcciones recombinantes adecuados para transformar Traustoquitrídos, incluyendo miembros de los géneros *Schizochytrium* y *Thraustochytrium*, no estaban disponibles. De manera importante, antes de la presente invención no se disponía del desarrollo de marcadores seleccionables que son particularmente útiles para transformar microorganismos de Thraustochytriales y de la identificación de secuencias de promotor específicas de Thraustochytriales.

Los investigadores anteriores han descrito métodos y reactivos de transformación para uso en diversos microorganismos, incluyendo en microalgas que no son miembros del orden Thraustochytriales. El documento de Patente de Estados Unidos N.º 6.027.900 concedida a Allnut *et al.*, desvela fusiones genéticas para uso en modificación por ingeniería genética de algas eucariotas, y en particular, *Phaeodactylum tricorutum*, usando un promotor para un gen de captación de luz de algas fotosintéticas y el gen *Sh ble* de *Streptoalloteichus hindustanus* como marcador seleccionable. Las células se hacen crecer en altas concentraciones de sal (por ejemplo, 10-35 g/l) y Zeocin™ para selección de transformantes. Las células de microalgas adecuadas para su transformación usando un método de este tipo son microalgas fotosintéticas que se pueden hacer crecer en condiciones de alta concentración de sal. El documento de Patente de Estados Unidos N.º 5.661.017 concedida a Dunahay *et al.*, desvela un método para transformar algas que contienen clorofila C (por ejemplo, Diatomeas) usando una construcción recombinante que comprende un marcador seleccionable unido de forma operativa a una secuencia de control reguladora adecuada para la expresión del marcador en las algas que contienen clorofila C. Se desvela que el marcador seleccionable es cualquier marcador adecuado, incluyendo los marcadores aislados de fuentes bacterianas y fúngicas, y es preferentemente neomicina fosfotransferasa. La secuencia de control reguladora puede incluir cualquier secuencia reguladora obtenida a partir de un alga que contiene clorofila C y, preferentemente, de *Cyclotella cryptica* (por ejemplo, una secuencia reguladora de acetil-CoA carboxilasa de *C. cryptica*). Ti Zhou-Chou *et al.*, 1995 (Journal of Biotechnology, 38 (2): 137-139) describe la expresión, purificación y caracterización de una proteína recombinante en *Dictyostelium discoideum*. Sin embargo, los métodos de ese tipo no son fácilmente transferibles para la transformación de microorganismos de Thraustochytriales, porque, antes de la presente invención, la transformación de microorganismos tales como Thraustochytriales (por ejemplo, microalgas) estaba lejos de ser rutinaria. Los marcadores y sistemas de transformación que han alcanzado un gran desarrollo para bacterias y levaduras no son necesariamente adaptables fácilmente a otros microorganismos. De hecho, el documento de Patente de Estados Unidos N.º 5.661.017 indica que "ha habido poco éxito en el desarrollo de sistemas de transformación para microalgas eucariotas" (col. 1, líneas 49-51), lo que se debe en parte a la dificultad de introducción de ADN extraño en tales microorganismos, y se debe en parte a una falta de marcadores y vectores adecuados para su uso en una transformación de ese tipo. El sistema que se describe en el documento de Patente de Estados Unidos N.º 5.661.017 se desarrolló de forma específica para las algas que contienen clorofila C porque los inventores creían que eran más susceptibles de transformación genética, en particular en comparación con otras algas. De forma análoga, el documento de Patente de Estados Unidos N.º 6.027.900, que enseña un método de transformación que es específico para microalgas fotosintéticas, habla de la creencia de que la mayoría de las algas son resistentes a cualquier tipo de manipulación genética (col. 1, líneas 39-47). Los sistemas adaptados para bacterias, levaduras, células de insecto y animales no se han adaptado fácilmente a microalgas. Por lo tanto, antes

de la presente invención, existía una necesidad en la técnica de sistemas y métodos de transformación eficaces que fueran específicos para microalgas.

Adicionalmente, aunque en la actualidad el orden Thraustochytriales se agrupa con las algas cromofitas en los Stramenopiles, algunos sostienen todavía en la técnica que estos microorganismos son bastante diferentes de la mayoría de las microalgas, y algunos de esos expertos en la materia tienen la opinión de que los miembros de Thraustochytriales pueden no estar apropiadamente clasificados como microalgas en absoluto. Los microorganismos que se considera que son microalgas han evolucionado por lo menos cuatro veces separadas durante la evolución, lo que conduce a que los microorganismos de tipo "microalga" sean ubicados en diferentes reinos (por ejemplo algas rojas, algas verdes y algas doradas (Chromophyta) están todas en reinos separados). Como resultado, no se espera que los sistemas de transformación que se ha demostrado que son útiles en otras microalgas sean útiles para Thraustochytriales. Por lo tanto, a pesar del valor comercial de microorganismos de Thraustochytriales, la capacidad para usar el potencial completo de tales microorganismos mediante ingeniería genética no se ha reconocido hasta el momento. Antes de la presente invención, los presentes inventores no eran conscientes de ningún promotor, marcador seleccionable o vector útil para la transformación de microorganismos de Thraustochytriales, ni había ningún conocimiento con respecto a qué sistemas de selección se podrían usar en o adaptar para Thraustochytriales.

En resumen, existe una necesidad en la técnica de desarrollar métodos para transformar microorganismos de Thraustochytriales, proporcionando de ese modo un medio para crear cepas con un aumento del valor comercial. Además, existe una necesidad en la técnica de desarrollar métodos para mutación o inactivación de genes específicos mediante recombinación homóloga o no homóloga en microorganismos de Thraustochytriales, proporcionando un nuevo modo para alterar el metabolismo celular y para identificar las funciones de genes específicos en Thraustochytriales.

Sumario de la Invención

La invención proporciona un método para producción de una proteína, que comprende: cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína; y recuperar la proteína. La invención también proporciona un método para producción de una proteína, que comprende: cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína, para producir la proteína; y extraer el microorganismo que comprende la proteína. La invención se define mediante las reivindicaciones. La materia objeto fuera del alcance de las reivindicaciones se proporciona solamente para información.

Breve descripción de las figuras

- FIG. 1 ilustra la construcción del plásmido recombinante pTUBZEO-11.
- FIG. 2 ilustra la construcción del plásmido recombinante pTUBZEO11-2.
- FIG. 3A ilustra el plásmido o recombinante pMON50200.
- FIG. 3B ilustra el plásmido o recombinante pMON50201.
- FIG. 3C ilustra el plásmido o recombinante pMON50202.
- FIG. 3D ilustra el plásmido o recombinante pMON50203.

Descripción detallada de la invención

Los métodos de la presente invención usan métodos y materiales relacionados para transformar por vía genética microorganismos del orden Thraustochytriales. Todas las cepas de microorganismos unicelulares que se desvelan en el presente documento para uso como un transformante de las construcciones recombinantes de la presente invención, que generalmente también se pueden denominar Traustoquítridos, son miembros del orden Thraustochytriales. De acuerdo con la presente invención, las expresiones "Traustoquítrido", "microorganismo de Thraustochytriales" y "microorganismo del orden Thraustochytriales" se pueden usar indistintamente. Los presentes inventores no son conscientes de ningún informe anterior que describa un sistema de transformación para *Schizochytrium* o cualquier otro microorganismo de Thraustochytriales. Los sistemas de transformación que se describen en el presente documento se pueden usar para introducir genes extraños en microorganismos del orden Thraustochytriales, proporcionando de ese modo un medio para crear cepas con un aumento del valor comercial. Además, la presente descripción permite la mutación o inactivación de genes específicos mediante recombinación homóloga o no homóloga, proporcionando un nuevo modo para alterar el metabolismo celular y para identificar las funciones de genes específicos en microorganismos de Thraustochytriales.

De forma más específica, los presentes inventores han demostrado la transformación genética de un microorganismo de Thraustochytriales del género, *Schizochytrium* (Orden: Thraustochytriales; Familia: Thraustochytriaceae; Género: *Schizochytrium*), mediante el uso de dos tipos de vectores de transformación. Estos vectores se pueden introducir en células mediante métodos convencionales, seguido por identificación y aislamiento

de células recombinantes basándose en su capacidad para crecer en presencia de compuestos selectivos. Los presentes inventores han demostrado la eficacia de estos vectores introduciéndolos mediante bombardeo de partículas, pero también se pueden usar otros medios para introducir los vectores (por ejemplo, electroporación) y se conocen en la técnica.

Para un vector de transformación, ejemplificado en el presente documento mediante el vector recombinante denominado pTUBZEO11-2, se creó un gen quimérico en el que el gen *ble* (que codifica para una "proteína de unión a bleomicina") de *Streptoalloteichus hindustanus* se colocó en el sentido cadena abajo de un promotor de un gen de tubulina de *Schizochytrium*. Un terminador de SV40 se colocó en el sentido cadena abajo del gen *ble* en esta construcción. Este vector permite la expresión del gen *ble* en *Schizochytrium*, confiriendo de ese modo resistencia a Zeocin™ y compuestos relacionados, que son tóxicos para células de tipo silvestre cuando se incluyen en el medio de crecimiento a niveles apropiados. La fuente del gen *ble* y el terminador de SV40 en esta construcción era un vector disponible en el mercado, denominado pSV40/Zeo, que se adquirió en Invitrogen Corporation (Carlsbad, CA) (Manual Técnico 180202, versión B, "ZeoCassette Vectors"; Invitrogen Corporation, 1600 Faraday Ave., Carlsbad, CA 92008). El promotor del gen de tubulina se aisló mediante la reacción en cadena de la polimerasa; uno de los cebadores usados para la reacción se basaba en los datos de secuencia obtenidos a través de un proyecto de secuenciación de ADNc de *Schizochytrium* de forma aleatoria. El mapa de pTUBZEO11-2 se muestra en la Fig. 2, y la secuencia de nucleótidos de pTUBZEO11-2 se representa mediante la SEQ ID NO: 9. La transformación de *Schizochytrium* con este vector se confirmó mediante el uso de la reacción en cadena de la polimerasa y análisis de transferencia de tipo Southern para detectar la presencia de secuencias de vector integradas en el genoma de *Schizochytrium*.

Los investigadores anteriores han usado el gen *ble* como un marcador seleccionable para transformación genética de una Diversidad de organismos, incluyendo bacterias, microalgas distintas de Traustochytridos, hongos, protozoos, plantas y células animales (Véase, por ejemplo, el documento de Patente de Estados Unidos N.º 6.027.900; Lumbreras *et al.*, 1998, Plant J. 14: 441-447; Rohe *et al.*, 1996, Curr. Genet. 29: 587-590; Messina *et al.*, 1995, Gene 165: 213-217; Guerrero *et al.*, 1992, Appl. Microbiol. Biotechnol. 36: 759-762; Perez *et al.*, 1989, Plant Mol. Biol. 13: 365-373; Gatigno *et al.*, 1990 Gene 91: 35-41). El gen *ble* codifica una "proteína de unión a bleomicina" que confiere resistencia a varios antibióticos, incluyendo bleomicina, fleomicina y Zeocin™ (Drocourt *et al.*, 1990, Nucleic Acids Res. 18: 4009). Este gen está disponible en el mercado en Invitrogen Corporation, que era la fuente del gen que los presentes inventores utilizaron para crear el vector de transformación de *Schizochytrium*, pTUBZBO11-2. Sin embargo, se cree que los presentes inventores son los primeros en producir un vector de transformación en el que el gen *ble* está unido a un promotor de Traustochytridos de una manera que permite la expresión del gen en Traustochytridos.

Un segundo conjunto de vectores de transformación se creó mediante mutagénesis dirigida al sitio *in vitro* de un gen de acetolactato sintasa (*als*) que los presentes inventores aislaron a partir de una biblioteca genómica de *Schizochytrium*. Estas mutaciones cambian la secuencia de aminoácidos de la enzima codificada (ALS) de tal modo que es mucho menos sensible a sulfometurón metilo y otros compuestos de sulfonilurea, así como inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo, a los que los microorganismos del orden Thraustochytriales son sensibles. Los compuestos de sulfonilurea tales como sulfometurón metilo (SMM) a menudo tóxicos son para las células porque son capaces de unirse a e inactivar la enzima acetolactato sintasa (ALS) de una diversidad de organismos. La ALS cataliza la primera etapa de la biosíntesis de los aminoácidos valina, leucina e isoleucina. Se ha mostrado que las imidazolinonas, triazolpirimidinas y otros compuestos se unen a e inactivan la ALS de ciertos organismos. Las formas mutantes de genes que codifican la acetolactato sintasa (también conocida como ácido acetohidroxi sintasa) de otros organismos se han usado previamente como marcadores seleccionables para la transformación de levaduras y plantas (Annu. Rev. Plant Physiol Plant Mol. Biol. 40: 441-470, 1989). Sin embargo, no existen informes anteriores a la presente invención que describan la secuencia o las propiedades del gen *als* de *Schizochytrium* o cualquier otro miembro de Thraustochytriales, o el uso de genes *als* de Thraustochytriales mutantes que confieran resistencia a compuestos de sulfonilurea imidazolinona u oxibenzoato de pirimidinilo. De hecho, según el conocimiento de los presentes inventores, no se ha publicado ningún informe con respecto a la sensibilidad de microorganismos de Thraustochytriales a estos agentes selectivos, incluyendo sulfometurón metilo y, por tanto, antes de la presente invención no se sabía si un marcador seleccionable de este tipo podría ser incluso viable para uso en un sistema de transformación de Traustochytridos. Se debe observar que en diversos organismos se producen genes con una homología sustancial con respecto a genes *als* conocidos, pero no codifican Enzimas que sean capaces de catalizar la síntesis de acetolactato (Biochimica et Biophysica Acta 1385: 401-419, 1998). Por lo tanto, no podría haber sido evidente que un homólogo de *als* clonado de hecho codifique ALS. Con el fin de determinar definitivamente si el gen de *Schizochytrium* clonado era un gen *als* verdadero, los presentes inventores demostraron, a través de experimentos de transformación, una correlación positiva de la resistencia a sulfometurón metil con la expresión del gen *als* de *Schizochytrium* supuesto mutado.

Los presentes inventores han producido tres vectores de transformación diferentes que contienen genes de *als* mutantes: un gen *als* mutante codifica una enzima con una valina en la posición 595 en lugar de a prolina (plásmido pMON50201, o ALSmut1-7), otro codifica una glutamina en la posición 192 en lugar de a prolina (plásmido pMON50202, o ALSmut2-2), y una tercera forma contiene ambas de estas mutaciones (plásmido pMON50203, o ALSmut3-5). En estos vectores, la expresión de los genes *als* mutantes está bajo el control del promotor y el

terminador del gen *als* nativo. Los mapas de estos vectores, junto con un vector que contiene el gen *als* de *Schizochytrium* de tipo silvestre (plásmido pMON50200, o AE-5), se muestran en las Figs 3A-3D. La transformación de *Schizochytrium* con estos vectores que codifican ALS mutante se confirmó mediante el uso de la reacción en cadena de la polimerasa y análisis de transferencia de tipo Southern para detectar la presencia de secuencias de vector integradas en el genoma de *Schizochytrium*. Tal como se describe en detalle a continuación, ahora que los presentes inventores han identificado la secuencia completa del gen *als*, también se pueden hacer otras mutaciones, que se especifican a continuación. Por lo tanto, los genes *als* mutantes que se describen pretenden ser a modo de ejemplo, y no incluyen todas las posibles mutaciones.

Los sistemas de transformación de la presente invención se han usado para introducir genes extraños en células de Thraustochytriales por medio de cotransformación. En estos casos, los genes extraños se colocaron entre diversos promotores de *Schizochytrium* y un terminador apropiados (por ejemplo, SV40 o una región terminadora genética de *Schizochytrium*). Por ejemplo, los presentes inventores han producido e introducido un gen sintético que codifica una ácido graso ω -3 desaturasa del nemátodo *Caenorhabditis elegans*, representado en el presente documento por la SEQ ID NO: 29, para aumentar los niveles de ácido docosahexaenoico en *Schizochytrium*. La SEQ ID NO: 30 representa la secuencia de aminoácidos de la desaturasa codificada por la SEQ ID NO: 29. En células de Thraustochytriales también se pueden introducir casetes de expresión que contienen genes extraños mediante inclusión directa dentro del vector de transformación que contiene el marcador seleccionable.

Además, los presentes inventores también han demostrado con los vectores que codifican ALS mutante que se produce recombinación homóloga en *Schizochytrium*, lo que indica la viabilidad del uso de medios recombinantes para inactivar o mutar genes de *Schizochytrium* nativos específicos.

Con respecto a las secuencias de promotor de Thraustochytriales que se describen en el presente documento, una búsqueda en base de datos de secuencias (GenBank) para todas las secuencias de nucleótidos y proteínas informadas para miembros del orden Thraustochytriales indica que, en el momento de la presente invención, no se había informado de ninguna secuencia promotora de *Schizochytrium* o cualquier otro miembro de Thraustochytriales. El único gen que se ha informado de cualquier especie de *Schizochytrium* es para el ARN ribosómico 5S de *S. aggregatum* (números de registro en GenBank X06104 y M13616). Se han informado secuencias de ARN ribosómico 5S y 18S para los miembros de Thraustochytriales, especie *Ulkenia*, y géneros *Labyrinthuloides* y *Thraustochytrium*, pero esto no presenta relación con la presente invención. Se ha descrito una región codificante parcial de un gen de "ARN polimerasa de tipo T3/T7 supuesta" de *Thraustochytrium aureum* (Nucleic Acids Research 15: 648-654, 1996), pero no se informó de una secuencia promotora para este gen.

Los métodos de la presente invención se pueden usar para introducir cualquier gen u otras secuencias de nucleótidos que son de interés en un microorganismo del orden Thraustochytriales. Tales secuencias de nucleótidos incluyen, pero no se limitan a, ácidos nucleicos que codifican proteínas (por ejemplo, enzimas) asociadas con la síntesis de ácidos grasos (por ejemplo, los ácidos grasos: ácido docosahexaenoico (DHA), ácido docosapentaenoico (DPA), ácido eicosapentaenoico (EPA) y ácido araquidónico (ARA)). Tales proteínas comprenden, pero no se limitan a: ácido graso sintasas, ácido graso desaturasas y ácido graso elongasas, así como proteínas asociadas con un complejo de policétido sintasa y/o proteínas asociadas con la incorporación de tales ácidos grasos en fosfolípidos o en moléculas de triacilglicerol. Por ejemplo, la invención se ha usado para introducir genes que codifican diversas ácido graso ω -3 desaturasas en *Schizochytrium* en un intento para aumentar el nivel de ácido docosahexaenoico (DHA) en las células a través de desaturación de ω -3 de ácido docosapentaenoico (DPA). Como otro ejemplo, también se ha demostrado la expresión de una ácido graso polienoico isomerasa supuesta del alga roja, *Ptilota*, en *Schizochytrium*. Los genes que codifican un complejo de policétido sintasa de *Schizochytrium* (es decir, un sistema de policétido sintasa) se han depositado como Números de Registro en GenBank, AF378329 (ORFA), AF378328 (ORFB) y AF378329 (ORFC).

Los métodos de la presente invención también son útiles para introducir, en microorganismos de Thraustochytriales, genes y otras secuencias de nucleótidos que codifican proteínas asociadas con la ruta biosintética de isoprenoides. Tales proteínas comprenden, pero no se limitan a, HMG-CoA sintasa y HMG-CoA reductasa. Otras proteínas adecuadas incluyen proteínas asociadas con la síntesis de moléculas derivadas de subunidades de isoprenoides que incluyen, pero no se limitan a, diversos compuestos esteroides y diversos compuestos carotenoides. Las proteínas asociadas con la síntesis de diversos compuestos carotenoides incluyen, pero no se limitan a, escualeno sintasa, fitoeno sintasa, fitoeno desaturasa y diversas carotenoide ciclasas, hidroxilasas y cetolasas.

Los métodos de la presente invención también son útiles para introducir, en Thraustochytriales, una o más secuencias de ácidos nucleicos que codifican proteínas asociadas con la síntesis de compuestos antioxidantes que incluyen, pero no se limitan a, vitamina E y ácido lipoico.

Además, los métodos de la presente invención se pueden usar para introducir cualquier gen u otros vectores de secuencias de nucleótidos en microorganismos de Thraustochytriales con el fin de inactivar o producir delección de genes (es decir, "desactivación" o "alteración genética dirigida"). La inactivación o delección de genes se usa generalmente con el fin de potenciar el valor comercial del microorganismo. Por ejemplo, puede ser deseable eliminar genes que codifican enzimas (o ácidos nucleicos que regulan la expresión de tales genes) de las rutas de

síntesis de ácidos grasos saturados y poliinsaturados. Por ejemplo, puede ser deseable inhibir o desactivar genes que codifican proteínas que están implicadas en la degradación de otros compuestos valiosos producidos por el microorganismo de Thraustochytriales o que de otra manera disminuyen el valor del compuesto deseado. Por ejemplo, los genes que codifican lipasas, enzimas de oxidación de ácidos grasos y proteínas que presentan olores o sabores desagradables pueden ser dianas de desactivación deseables. Además, puede ser deseable desactivar genes que codifican proteínas que están asociadas con la síntesis de compuestos cuya síntesis está en competición con otras moléculas de interés. Por ejemplo, tales genes incluyen, pero no se limitan a, genes que codifican proteínas implicadas en la biosíntesis de hidratos de carbono, genes que codifican proteínas implicadas en la síntesis de diversos productos de rutas de isoprenoides (por ejemplo, esteroides o compuestos carotenoides específicos), y genes que codifican proteínas implicadas en la síntesis de componentes de la pared celular. A modo de ejemplo, se han introducido genes en células de *Schizochytrium* mediante el uso de la presente invención en un intento de inactivar genes que son homólogos a los genes de la policétido sintasa de *Shewanella* con el fin de evaluar su papel en la producción de ácidos grasos altamente insaturados (HUFA). Tal como se muestra a modo de ejemplo con el Ejemplo 6, la presente descripción también se pueden usar para inactivar, producir delección o mutar genes nativos que están implicados en la producción de ácidos grasos, carotenoides, esteroides, vitaminas u otros compuestos con el fin de mejorar la rentabilidad o aceptabilidad de productos que están relacionados con estos compuestos. Se indica que en algunas realizaciones, tal como se ha discutido anteriormente, puede ser deseable potenciar la producción de una proteína dada, mientras que en otras formas de realización, puede ser deseable a aumentar la producción de una proteína dada.

Los métodos que se describen en el presente documento también son útiles para determinar el proceso de recombinación genética en *Schizochytrium*.

Otros genes y moléculas de ácido nucleico útiles para su introducción en Thraustochytriales resultarán evidentes para los expertos en la materia.

A continuación se describen diversas realizaciones de los métodos de la presente invención inicialmente con respecto a un gen *als* y/o proteína ALS de Thraustochytriales. Sin embargo, se debe entender que las definiciones generales de términos y métodos se pretenden aplicar a la discusión de otros genes, ácidos nucleicos y proteínas que se describen a continuación.

Los métodos de la presente invención se basan en parte en la identificación, aislamiento y producción de secuencias de ácidos nucleicos que codifican marcadores seleccionables que son adecuados para uso en construcciones recombinantes para la transformación de microorganismos Traustoquítridos. Los marcadores seleccionables de ese tipo permiten la selección de microorganismos que se han transformado de forma satisfactoria con las construcciones recombinantes que se describen en el presente documento. Un marcador seleccionable útil para la transformación de Thraustochytriales adecuado para uso de acuerdo con los métodos de la presente invención es una acetolactato sintasa de Thraustochytriales (es decir, ALS). Preferentemente, la acetolactato sintasa se ha modificado, mutado o seleccionado de otro modo, para que sea resistente a la inhibición por compuestos de sulfonilurea, inhibidores de la clase de imidazolinona y/u oxibenzoatos de pirimidinilo (es decir, una ALS de ese tipo es un homólogo de una acetolactato sintasa que se produce de manera natural).

Una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención es una proteína que presenta actividad biológica de acetolactato sintasa, incluyendo proteínas de longitud completa, proteínas de fusión, o cualquier homólogo de una acetolactato sintasa que se produce de manera natural. Un homólogo de una acetolactato sintasa incluye proteínas que difieren de una acetolactato sintasa que se produce de manera natural porque al menos uno o unos pocos, pero sin limitarse a uno o unos pocos, aminoácidos han sufrido delección (por ejemplo, una versión truncada de la proteína, tal como un péptido o fragmento), insertado, invertido, sustituido y/o derivatizado (por ejemplo, mediante glicosilación, fosforilación, acetilación, miristoilación, prenilación, palmitación, amidación y/o adición de glicosilfosfatidil inositol). A continuación se describen con detalle los homólogos preferentes de una acetolactato sintasa que se produce de manera natural.

Una proteína aislada, tal como una acetolactato sintasa aislada, de acuerdo con los métodos de la presente invención, es una proteína que se ha retirado de su medio natural (es decir, que se ha sometido a manipulación humana) y puede incluir proteínas purificadas, proteínas parcialmente purificadas, proteínas producidas de manera recombinante y proteínas producidas de manera sintética, por ejemplo. Como tal, "aislado" no refleja el grado hasta el que la proteína se ha purificado. Una "acetolactato sintasa de Thraustochytriales" se refiere a una acetolactato sintasa (incluyendo un homólogo de una acetolactato sintasa que se produce de manera natural) de un microorganismo de Thraustochytriales o que se ha producido de otra manera a partir del conocimiento de la estructura (por ejemplo, secuencia) de una acetolactato sintasa que se produce de manera natural de un microorganismo de Thraustochytriales. En otras palabras, una acetolactato sintasa de Thraustochytriales incluye cualquier acetolactato sintasa que tiene la estructura y la función de una acetolactato sintasa que se produce de manera natural a partir de un microorganismo de Thraustochytriales o que es un homólogo biológicamente activo (es decir, presenta actividad biológica) de una acetolactato sintasa que se produce de manera natural a partir de un microorganismo de Thraustochytriales tal como se describe en detalle en el presente documento. Como tal, una

acetolactato sintasa de Thraustochytriales puede incluir proteínas purificadas, parcialmente purificadas, recombinantes, mutadas/modificadas y sintéticas.

5 En general, la actividad biológica o acción biológica de una proteína se refiere a cualquier función o funciones presentadas o realizadas por la proteína que se atribuye a la forma que se produce que se produce de manera natural de la proteína tal como se mide u observa *in vivo* (es decir, en el entorno fisiológico natural de la proteína) o *in vitro* (es decir, en condiciones de laboratorio). Por ejemplo, una actividad biológica de una acetolactato sintasa incluye actividad enzimática de acetolactato sintasa. Las modificaciones de una proteína, tal como en un homólogo o mimético (que se discute a continuación), pueden dar como resultado proteínas que presentan la misma actividad biológica que la proteína que se produce de manera natural, o en proteínas que presentan un aumento o disminución de la actividad biológica en comparación con la proteína que se produce de manera natural. Las modificaciones que dan como resultado una disminución en la expresión de la proteína o una disminución en la actividad de la proteína se pueden denominar inactivación (completa o parcial), regulación por disminución o disminución de la acción de una proteína. De forma análoga, las modificaciones que dan como resultado un aumento en la expresión de la proteína o un aumento en la actividad de la proteína se pueden denominar amplificación, sobreproducción, activación, potenciación, regulación por incremento o aumento de la acción de una proteína.

20 Con respecto a la acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención, es preferente que las modificaciones presentes en homólogos de acetolactato sintasa, en comparación con una acetolactato sintasa que se produce de manera natural, no cambien sustancialmente, o al menos no disminuyan sustancialmente, la actividad biológica básica de la sintasa en comparación con la proteína que se produce de manera natural. Sin embargo, tales homólogos pueden presentar diferencias en características distintas de la actividad funcional o enzimática de la proteína en comparación con la forma que se produce de manera natural, tal como una disminución de la sensibilidad a la inhibición mediante ciertos compuestos en comparación con la proteína que se produce de manera natural. Preferentemente, un homólogo de una acetolactato sintasa que se produce de manera natural presenta reducción de la sensibilidad (es decir, disminuida, reducida) con respecto a compuestos que se unen a e inactivan acetolactato sintasas que se produce de manera natural en comparación con la acetolactato sintasa que se produce de manera natural a partir de la que se obtuvo el homólogo. Por ejemplo, los compuestos de sulfonilurea, tales como sulfometurón metilo (SMM), a menudo son tóxicos para las células porque se pueden unir a e inactivar la acetolactato sintasa (ALS). También se ha mostrado que las imidazolinonas, triazolopirimidinas y otros compuestos similares (denominados generalmente inhibidores de la clase de imidazolinona en el presente documento) se unen a e inactivan ALS. Por tanto, un homólogo de una acetolactato sintasa que se produce de manera natural presenta preferentemente una reducción de la sensibilidad a compuestos de sulfonilurea, así como a inhibidores de la clase de imidazolinona (por ejemplo, al presentar sitios de unión alterados para tales inhibidores o sitios de unión con afinidad reducida por el inhibidor) y a oxibenzoatos de pirimidinilo, mientras que mantiene la actividad enzimática de acetolactato sintasa.

40 Como se usa en el presente documento, una proteína que presenta "actividad biológica de acetolactato sintasa" o que se denomina "acetolactato sintasa" se refiere a una proteína que cataliza la primera etapa en la biosíntesis de los aminoácidos valina, leucina e isoleucina. De forma más específica, una acetolactato sintasa aislada adecuada para uso de acuerdo con los métodos de la presente invención, incluyendo proteínas de longitud completa, proteínas truncadas, proteínas de fusión y homólogos, se puede identificar de una manera clara mediante la capacidad de las proteínas para catalizar la síntesis de acetolactato a partir de piruvato. La actividad biológica de la acetolactato sintasa se puede evaluar por alguien con experiencia en la materia mediante cualquier ensayo *in vitro* o *in vivo* adecuado para actividad enzimática.

50 En una realización, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 65 % con respecto a una secuencia de aminoácidos seleccionada entre el grupo de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, sobre al menos aproximadamente 600 aminoácidos de cualquiera de tales secuencias, en la que la proteína es una acetolactato sintasa (es decir, tiene actividad biológica de acetolactato sintasa). Más preferentemente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 70 %, y más preferentemente, idéntica en al menos aproximadamente un 75 %, e incluso más preferentemente idéntica en al menos aproximadamente un 80 %, e incluso más preferentemente idéntica en al menos aproximadamente un 85 %, e incluso más preferentemente idéntica en al menos aproximadamente un 90 % e incluso más preferentemente idéntica en al menos aproximadamente un 95 %, e incluso más preferentemente idéntica en al menos aproximadamente un 96 %, e incluso más preferentemente idéntica en al menos aproximadamente un 97 %, e incluso más preferentemente idéntica en al menos aproximadamente un 98 %, e incluso más preferentemente idéntica en al menos aproximadamente un 99 % a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, sobre al menos aproximadamente 600 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa.

65 En otra realización, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 75 % con respecto

a una secuencia de aminoácidos seleccionada entre el grupo de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, sobre al menos 95 aminoácidos de cualquiera de tales secuencias, en la que la proteína es una acetolactato sintasa (es decir, tiene actividad biológica de acetolactato sintasa). Más preferentemente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 80 %, e incluso más preferentemente idéntica en al menos aproximadamente un 85 %, e incluso más preferentemente idéntica en al menos aproximadamente un 90 % e incluso más preferentemente idéntica en al menos aproximadamente un 95 %, e incluso más preferentemente idéntica en al menos aproximadamente un 96 %, e incluso más preferentemente idéntica en al menos aproximadamente un 97 %, e incluso más preferentemente idéntica en al menos aproximadamente un 98 %, e incluso más preferentemente idéntica en al menos aproximadamente un 99 % a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, sobre al menos 95 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa. Incluso más preferentemente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos que tiene cualquiera de los porcentajes de identidad que se han mencionado anteriormente con respecto a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24 sobre al menos 100 aminoácidos, y más preferentemente 125, y más preferentemente 150, y más preferentemente 175, y más preferentemente 200, y más preferentemente 225, y más preferentemente 250, y más preferentemente 275, y más preferentemente 300, y más preferentemente 325, y más preferentemente 350, y más preferentemente 375, y más preferentemente 400, y más preferentemente 425, y más preferentemente 450, y más preferentemente 475, y más preferentemente 500, y más preferentemente 525, y más preferentemente 550, y más preferentemente 575 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa.

Como se usa en el presente documento, a menos que se indique de otro modo, la referencia a un porcentaje (%) de identidad se refiere a una evaluación de homología que se realiza usando: (1) una búsqueda de homología BLAST 2.0 Basic BLAST usando blastp para búsquedas de aminoácido y blastn para búsquedas de aminoácido con parámetros por defecto convencionales, en el que la secuencia de consulta se filtra para regiones de baja complejidad por defecto (se describe en Altschul, S.F., Madden, T.L., Schääffer, A.A., Zhang, J., Zhang, Z., Miller, W. Y Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". Nucleic Acids Res. 25: 3389-3402; (2) un alineamiento BLAST 2 (usando los parámetros que se describen a continuación); (3) y/o PSI-BLAST con los parámetros por defecto convencionales (BLAST Iterado Específico de Posición). Se indica que debido a algunas diferencias en los parámetros convencionales entre BLAST 2.0 Basic BLAST y BLAST 2, se podría reconocer que dos secuencias específicas presentan homología significativa usando el programa BLAST 2, mientras que una búsqueda realizada en BLAST 2.0 Basic BLAST usando una de las secuencias como secuencia de consulta puede no identificar la segunda secuencia en las coincidencias máximas. Además, PSI-BLAST proporciona una versión automatizada, fácil de usar de una búsqueda de "perfil", que es un modo sensible para buscar homólogos de secuencia. El primer programa realiza una búsqueda en la base de datos BLAST con huecos. El programa PSI-BLAST usa la información a partir de cualquier alineamiento significativo devuelto para construir una matriz de puntuación específica de posición, que reemplaza a la secuencia de consulta para la siguiente ronda de búsqueda en la base de datos. Por lo tanto, se debe entender que el porcentaje de identidad se puede determinar usando uno cualquiera de estos programas.

Dos secuencias específicas se pueden alinear entre sí utilizando un alineamiento de secuencias BLAST 2 tal como se describe en Tatusova y Madden, (1999), "Blast 2 sequences - a new tool for comparing protein and nucleotide sequences", FEMS Microbiol Lett. 174: 247-250. El alineamiento de secuencias BLAST 2 se realiza en blastp o blastn usando el algoritmo de BLAST 2.0 para realizar una búsqueda de BLAST con huecos (BLAST 2.0) entre las dos secuencias que permite la introducción de huecos (delecciones e inserciones) en El alineamiento resultante. Para fines de claridad en el presente documento, el alineamiento de secuencias con BLAST 2 se realiza usando los parámetros por defecto convencionales tal como sigue a continuación.

Para blastn, usando una matriz 0 de BLOSUM62:

Recompensa para coincidencia = 1

Penalización por falta de coincidencia = -2

Penalizaciones por apertura de huecos (5) y extensión de huecos (2)

Caída de alineamiento de hueco (*gap x_dropoff*) (50) esperado (10) tamaño de palabra (11) filtro (activado)

Para blastp, usando matriz 0 de BLOSUM62:

Penalizaciones por apertura de huecos (11) y extensión de huecos (1)

Caída de alineamiento de hueco (*gap x_dropoff*) (50) esperado (10) tamaño de palabra (3) filtro (activado).

Una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención también puede incluir proteína que tenga una secuencia de aminoácidos que comprende al menos 30 restos de aminoácido contiguos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, (es decir, 30 restos de aminoácido contiguos que tienen una identidad de un 100 % con 30 aminoácidos contiguos de cualquiera de

SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24). En una realización preferente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención incluye proteínas que tienen secuencias de aminoácidos que comprenden al menos 50, y más preferentemente al menos 75, y más preferentemente al menos 100, y más preferentemente al menos 115, y más preferentemente al menos 130, y más preferentemente al menos 150, y más preferentemente al menos 200, y más preferentemente, al menos 250, y más preferentemente, al menos 300, y más preferentemente, al menos 350, y más preferentemente, al menos 400, y más preferentemente, al menos 450, y más preferentemente, al menos 500, y más preferentemente, al menos 550, y más preferentemente, al menos 600, y más preferentemente, al menos 650, restos de aminoácidos contiguos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24. Una proteína de ese tipo tiene actividad biológica de acetolactato sintasa.

Como se usa en el presente documento, el término "contiguo" o "consecutivo", con respecto a las secuencias de ácidos nucleicos o aminoácidos que se describen en el presente documento, se refiere a que están conectados en una secuencia ininterrumpida. Por ejemplo, para una primera secuencia, comprender 30 aminoácidos contiguos (o consecutivos) de una segunda secuencia, se refiere a que la primera secuencia incluye una secuencia ininterrumpida de 30 restos de aminoácido que es idéntica en un 100 % con respecto a una secuencia ininterrumpida de 30 restos de aminoácido en la segunda secuencia. De forma análoga, que una primera secuencia presente una "identidad de un 100 %" con una segunda secuencia se refiere a que la primera secuencia coincide exactamente con la segunda secuencia sin huecos entre nucleótidos o aminoácidos.

En otra realización, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención, incluyendo un homólogo de acetolactato sintasa, incluye una proteína que tiene una secuencia de aminoácidos que es lo suficientemente similar con respecto a una secuencia de aminoácidos de acetolactato sintasa que se produce de manera natural que una secuencia de ácidos nucleicos que codifica el homólogo es capaz hibridarse en condiciones de rigurosidad moderada, elevada, o muy elevada (que se describen a continuación) a (es decir, con) una molécula de ácido nucleico que codifica la acetolactato sintasa que se produce de manera natural (es decir, al complemento de la hebra de ácido nucleico que codifica la secuencia de aminoácidos de la acetolactato sintasa que se produce de manera natural). Preferentemente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención está codificada por una secuencia de aminoácidos que se hibrida en condiciones de rigurosidad moderada, elevada o muy elevada al complemento de una secuencia de ácidos nucleicos que codifica una proteína que comprende una secuencia de aminoácidos representada por SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24. Incluso más preferentemente, una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención está codificada por una secuencia de ácidos nucleicos que se hibrida en condiciones de rigurosidad moderada, elevada o muy elevada al complemento de los nucleótidos 1260-3314 de la SEQ ID NO: 15, nucleótidos 1260-3314 de la SEQ ID NO: 18, nucleótidos 1260-3314 de la SEQ ID NO: 21, o nucleótidos 1260-3314 de la SEQ ID NO: 23. Las condiciones de hibridación de ese tipo se describen con detalle a continuación. Un complemento de secuencia de ácidos nucleicos de secuencia de ácidos nucleicos que codifica una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención se refiere a la secuencia de ácidos nucleicos de la hebra de ácido nucleico que es complementaria a la hebra que codifica la acetolactato sintasa. Se observará que un ADN bicatenario que codifica una secuencia de aminoácidos dada comprende un ADN monocatenario y su hebra complementaria que tiene una secuencia que es un complemento al ADN monocatenario. Como tal, las moléculas de ácido nucleico adecuadas para uso de acuerdo con los métodos de la presente invención pueden ser bicatenarias o monocatenarias, e incluyen las moléculas de ácido nucleico que forman híbridos estables en condiciones de hibridación rigurosas con una secuencia de ácidos nucleicos que codifican las secuencias de aminoácidos SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, y/o con el complemento de la secuencia de ácidos nucleicos que codifica cualquiera de tales secuencias de aminoácidos. Los expertos en la materia conocen métodos para deducir una secuencia complementaria. Se debería indicar que dado que las tecnologías de secuenciación de aminoácidos y de secuenciación de ácidos nucleicos no están completamente libres de errores, las secuencias que se presentan en el presente documento, en el mejor de los casos, representan secuencias evidentes de una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención.

Los homólogos de acetolactato sintasa pueden ser el resultado de variación alélica natural o mutación natural. También se pueden producir homólogos de acetolactato sintasa adecuados para uso de acuerdo con los métodos de la presente invención usando técnicas conocidas en la materia que incluyen, pero no se limitan a, modificaciones directas en la proteína o modificaciones en el gen que codifica para la proteína usando, por ejemplo, técnicas de ADN clásicas o recombinantes para efectuar mutagénesis aleatoria o dirigida. Una variante alélica que se produce de manera natural de un ácido nucleico que codifica una acetolactato sintasa es un gen que se encuentra esencialmente en el mismo o los mismos locus en el genoma que el gen que codifica una secuencia de aminoácidos de SEQ ID NO: 15, pero que, debido a variaciones naturales causadas por, por ejemplo, mutación o recombinación, presenta una secuencia similar pero no idéntica. Las variantes alélicas naturales generalmente codifican proteínas que presentan actividad similar a la de la proteína codificada por el gen con el que se están comparando. Una clase de variantes alélicas puede codificar la misma proteína pero puede presentar diferentes secuencias de ácido nucleico debido a la degeneración del código genético. Las variantes alélicas también pueden comprender alteraciones en las regiones no traducidas en las posiciones 5' o 3' del gen (por ejemplo, en regiones de control reguladoras). Los expertos en la materia conocen bien las variantes alélicas.

Las proteínas de acetolactato sintasa adecuadas para uso de acuerdo con los métodos de la presente invención también incluyen productos de expresión de fusiones genéticas (por ejemplo, usadas para sobreexpresar formas activas, solubles de la proteína recombinante), de genes sometidos a mutagénesis (tales como genes que presentan modificaciones de codones para potenciar la transcripción y traducción genéticas), y de genes truncados (tales como genes que presentan dominios de unión a membrana eliminados para generar formas solubles de una proteína de membrana, o genes que presentan secuencias señal eliminadas que se toleran escasamente en un hospedador recombinante en particular).

El tamaño mínimo de una proteína acetolactato sintasa y/u homólogo adecuado para uso de acuerdo con los métodos de la presente invención es un tamaño suficiente para tener actividad biológica de acetolactato sintasa. Preferentemente, una proteína acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención tiene una longitud de al menos al menos 30 aminoácidos, y más preferentemente, una longitud de al menos aproximadamente 50, y más preferentemente al menos 75, y más preferentemente al menos 100, y más preferentemente al menos 115, y más preferentemente al menos 130, y más preferentemente al menos 150, y más preferentemente al menos 200, y más preferentemente, al menos 250, y más preferentemente, al menos 300, y más preferentemente, al menos 350, y más preferentemente, al menos 400, y más preferentemente, al menos 450, y más preferentemente, al menos 500, y más preferentemente, al menos 550, y más preferentemente, al menos 600, y más preferentemente, al menos 650, y más preferentemente, al menos 684 aminoácidos. No existe ningún límite, distinto de un límite práctico, con respecto al tamaño máximo de una proteína de ese tipo porque la proteína puede incluir una parte de una proteína de acetolactato sintasa o una acetolactato sintasa de longitud completa, además de una secuencia adicional (por ejemplo, una secuencia de proteína de fusión), si se desea.

En el presente documento se describe una proteína de fusión que incluye un dominio que contiene acetolactato sintasa (es decir, una secuencia de aminoácidos para una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención) unida a uno o más segmentos de fusión. Los segmentos de fusión adecuados para uso de acuerdo con los métodos de la presente invención incluyen, pero no se limitan, segmentos que pueden: potenciar la estabilidad de una proteína; proporcionar otra actividad biológica deseable; y/o ayudar con la purificación de una acetolactato sintasa (por ejemplo, mediante cromatografía de afinidad). Un segmento de fusión adecuado puede ser un dominio de cualquier tamaño que presente la función deseada (por ejemplo, confiere un aumento de estabilidad, solubilidad, acción o actividad biológica; y/o simplifica la purificación de una proteína). Los segmentos de fusión se pueden unir a extremos amino y/o carboxilo terminales del dominio que contiene acetolactato sintasa de la proteína y pueden ser susceptibles de escisión con el fin de permitir la recuperación sencilla de una acetolactato sintasa. Las proteínas de fusión se producen preferentemente cultivando una célula recombinante transfectada con una molécula de ácido nucleico de fusión que codifica una proteína que incluye el segmento de fusión unido al extremo o bien carboxilo y/o bien amino terminales de un dominio que contiene acetolactato sintasa.

Un mimético de una acetolactato sintasa también es adecuado para uso de acuerdo con los métodos de la presente invención. Como se usa en el presente documento el término "mimético" se usa para hacer referencia a cualquier compuesto peptídico o no peptídico que es capaz de imitar la acción biológica de un péptido que se produce de manera natural, a menudo porque el mimético presenta una estructura básica que imita la estructura básica del péptido que se produce de manera natural y/o presenta las propiedades biológicas destacadas del péptido que se produce de manera natural. Los miméticos pueden incluir, pero no se limitan a: péptidos que presentan modificaciones sustanciales con respecto al prototipo de manera que no presentan similitud de cadenas laterales con el péptido que se produce de manera natural (tales modificaciones, por ejemplo, pueden disminuir su susceptibilidad a la degradación); anticuerpos anti-idiotípicos y/o catalíticos, o fragmentos de los mismos; partes no proteicas de una proteína aislada (por ejemplo, estructuras de hidratos de carbono); o moléculas orgánicas sintéticas o naturales, incluyendo ácidos nucleicos y fármacos identificados a través de química combinatoria, por ejemplo.

Tales miméticos pueden ser diseñados, seleccionados y/o identificados de otro modo usando una diversidad de métodos conocidos en la técnica. Diversos métodos de diseño de fármacos, útiles para diseñar miméticos u otros compuestos terapéuticos útiles en los métodos de la presente invención se desvelan en Maulik *et al.*, 1997, *Molecular Biotechnology: Therapeutic*. Un mimético de acetolactato sintasa se puede obtener, por ejemplo, a partir de estrategias de diversidad molecular (una combinación de estrategias relacionadas que permiten la construcción rápida de bibliotecas de moléculas grandes, químicamente diversas), bibliotecas de compuestos naturales o sintéticos, en particular a partir de bibliotecas químicas o combinatorias (es decir, bibliotecas de compuestos que se diferencian en secuencia o tamaño pero que presentan elementos estructurales similares) o mediante diseño de fármacos racional, dirigido o aleatorio. Véase por ejemplo, Maulik *et al.*, mencionado anteriormente.

En una estrategia de diversidad molecular, se sintetizan bibliotecas de compuestos grandes, por ejemplo, a partir de péptidos, oligonucleótidos, carbohidratos y/o moléculas orgánicas sintéticas, usando enfoques biológicos, enzimáticos y/o químicos. Los parámetros críticos en el desarrollo de una estrategia de diversidad molecular incluyen diversidad de subunidades, tamaño molecular y diversidad de bibliotecas. El objetivo general de la identificación sistemática de tales bibliotecas es usar la aplicación secuencial de la selección combinatoria para obtener ligandos de alta afinidad para una diana deseada, y a continuación optimizar las moléculas líder mediante

estrategias de diseño aleatorias o bien dirigidas. En Maulik, *et al.*, se describen métodos de diversidad molecular con detalle, igual que la referencia mencionada anteriormente.

5 Maulik *et al.*, también desvelan, por ejemplo, métodos de diseño dirigido, en los que el usuario dirige el proceso de creación de nuevas moléculas a partir de una biblioteca de fragmentos de fragmentos seleccionados de forma apropiada; diseño aleatorio, en el que el usuario usa un algoritmo genético o de otro tipo para mutar de forma aleatoria fragmentos y sus combinaciones mientras que se aplica simultáneamente un criterio de selección para evaluar la idoneidad de los ligandos candidatos; y un enfoque basado en rejilla en el que el usuario calcula la energía de interacción entre estructuras de receptores tridimensionales y sondas de fragmentos pequeños, seguido por unión entre sí de sitios de sondas favorables.

15 Las acetolactato sintasas adecuadas para uso de acuerdo con los métodos de la presente invención se pueden obtener a partir de cualquier microorganismo de Thraustochytriales, y en particular, a partir de cualquier microorganismo de *Schizochytrium*. En una realización, una acetolactato sintasa preferente adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de aminoácidos seleccionada entre el grupo de la SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24. La proteína que tiene una secuencia de aminoácidos representada por la SEQ ID NO: 15 es una acetolactato sintasa que se produce de manera natural (es decir, de tipo silvestre) de un microorganismo de Thraustochytriales, y de forma específica, es una acetolactato sintasa de *Schizochytrium*. Las secuencias de aminoácidos representadas por las SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24 son secuencias que se han modificado, de manera que las enzimas resultantes presentan una reducción de la sensibilidad a compuestos de sulfonilurea, así como a inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo, en comparación con la proteína que se produce de manera natural representada por la secuencia de aminoácidos SEQ ID NO: 15. Se indica que las proteínas representadas por las SEQ ID NO: 19, SEQ ID NO: 22 y SEQ ID NO: 24 tienen actividad biológica de acetolactato sintasa. Las acetolactato sintasas con reducción de la sensibilidad a compuestos de sulfonilurea, así como a inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo son acetolactato sintasas preferentes adecuadas para uso de acuerdo con los métodos de la presente invención, porque las secuencias de ácidos nucleicos que codifican tales sintasas se pueden usar en vectores recombinantes adecuados para uso de acuerdo con los métodos de la presente invención como marcadores seleccionables.

30 Por lo tanto, en una realización se presenta una acetolactato sintasa modificada adecuada para uso de acuerdo con los métodos de la presente invención, incluyendo cualquier homólogo de cualquiera de las SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 y SEQ ID NO: 24, en la que el homólogo tiene actividad biológica de acetolactato sintasa, y en particular, en la que el homólogo tiene una reducción de la sensibilidad a compuestos de sulfonilurea, así como a inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo, en comparación con la proteína que se produce de manera natural representada por la secuencia de aminoácidos SEQ ID NO: 15. En un aspecto, tales homólogos de acetolactato sintasa incluyen proteínas que tienen una secuencia de aminoácidos que difiere de la SEQ ID NO: 15 en una sustitución, inserción o delección de aminoácido en una o más de las siguientes posiciones: 116G, 117A, 192P, 200A, 251K, 358M, 383D, 592V, 595W, o 599F. Estas posiciones corresponden a sitios de mutación de ALS conocidos en una acetolactato sintasa de levadura (es decir, 116G, 117A, 192P, 200A, 251K, 354M, 379D, 583V, 586W, y 590F, respectivamente) (Véase Mazur y Falco, 1989, Annu. Rev. Plant Physiol. Plant Mol. Biol. 40: 441-470, que se incorpora en el presente documento por referencia en su totalidad). Los expertos en la materia conocerán otros posibles sitios de mutación basándose en mutaciones de aminoácidos satisfactorias en ALS de otros organismos. La aplicación de tales sitios a los sitios correspondientes en la ALS de Thraustochytriales puede producir una proteína adecuada para uso de acuerdo con los métodos de la presente invención.

50 Como se ha analizado anteriormente, los métodos de la presente invención se basan en parte en el descubrimiento y producción de construcciones recombinantes para la transformación de microorganismos de Traustochytridos. Por lo tanto, en una realización una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención comprende una secuencia de ácidos nucleicos que codifica una de acetolactato sintasa de Traustochytridos, y secuencia de ácidos nucleicos totalmente complementaria a la misma. Una molécula de ácido nucleico adecuada para uso de acuerdo con los métodos de la presente invención que codifica una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención incluye una molécula de ácido nucleico que codifica cualquiera de las proteínas de acetolactato sintasa, incluyendo homólogos, que sea discutido anteriormente. Más particularmente, en una realización una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención comprende la secuencia de ácidos nucleicos decodifica una proteína que tiene una secuencia de aminoácidos que es idéntica en al menos aproximadamente un 65 % con respecto a una secuencia de aminoácidos seleccionada entre el grupo de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, sobre al menos aproximadamente 600 aminoácidos de cualquiera de las secuencias de ese tipo; en las que la proteína es una acetolactato sintasa (es decir, tiene actividad biológica de acetolactato sintasa). Más preferentemente, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de ácidos nucleicos que codifica la secuencia de aminoácidos que es idéntica en al menos aproximadamente un 70 %, y más preferentemente, idéntica en al menos aproximadamente un 75 %, e incluso más preferentemente idéntica en al menos aproximadamente un 80 %, e incluso más preferentemente idéntica en al menos aproximadamente un 85 %, e incluso más preferentemente idéntica en al menos aproximadamente un 90 % e incluso más preferentemente idéntica en al menos aproximadamente un 95 %, e incluso más preferentemente idéntica en al menos aproximadamente un 95 %.

e incluso más preferentemente idéntica en al menos aproximadamente un 96 %, e incluso más preferentemente idéntica en al menos aproximadamente un 97 %, e incluso más preferentemente idéntica en al menos aproximadamente un 98 %, e incluso más preferentemente idéntica en al menos aproximadamente un 99 % a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, sobre al menos aproximadamente 600 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa.

En otra realización, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de ácidos nucleicos que codifica la secuencia de aminoácidos que es idéntica en al menos aproximadamente un 75 % con respecto a una secuencia de aminoácidos seleccionada entre el grupo de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22, SEQ ID NO: 24, sobre al menos 95 aminoácidos de cualquiera de tales secuencias, en la que la proteína es una acetolactato sintasa (es decir, tiene actividad biológica de acetolactato sintasa). Más preferentemente, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de ácidos nucleicos que codifica la secuencia de aminoácidos que es idéntica en al menos aproximadamente un 80 %, e incluso más preferentemente idéntica en al menos aproximadamente un 85 %, e incluso más preferentemente idéntica en al menos aproximadamente un 90 % e incluso más preferentemente idéntica en al menos aproximadamente un 95 %, e incluso más preferentemente idéntica en al menos aproximadamente un 96 %, e incluso más preferentemente idéntica en al menos aproximadamente un 97 %, e incluso más preferentemente idéntica en al menos aproximadamente un 98 %, e incluso más preferentemente idéntica en al menos aproximadamente un 99 % a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, sobre al menos 95 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa.

En otra realización más, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención tiene una secuencia de ácidos nucleicos que codifica una secuencia de aminoácidos que tiene cualquiera de los porcentajes de identidad que se han mencionado anteriormente con respecto a cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24 sobre al menos 100 aminoácidos, y más preferentemente 125, y más preferentemente 150, y más preferentemente 175, y más preferentemente 200, y más preferentemente 225, y más preferentemente 250, y más preferentemente 275, y más preferentemente 300, y más preferentemente 325, y más preferentemente 350, y más preferentemente 375, y más preferentemente 400, y más preferentemente 425, y más preferentemente 450, y más preferentemente 475, y más preferentemente 500, y más preferentemente 525, y más preferentemente 550, y más preferentemente 575 aminoácidos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, en la que la proteína tiene actividad biológica de acetolactato sintasa. El porcentaje de identidad se determina usando parámetros por defecto de BLAST 2.0 Basic BLAST, como se ha descrito anteriormente.

En una realización, las moléculas de ácido nucleico adecuadas para uso de acuerdo con los métodos de la presente invención que codifican una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención incluyen moléculas de ácido nucleico aisladas que se hibridan en condiciones de rigurosidad moderada, e incluso más preferentemente en condiciones de alta rigurosidad, e incluso más preferentemente en condiciones de rigurosidad muy elevada con el complemento de una secuencia de ácidos nucleicos que codifica una acetolactato sintasa que se produce de manera natural. Preferentemente, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención que codifica una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención comprende una secuencia de ácidos nucleicos que se hibrida en condiciones de rigurosidad moderada o elevada al complemento de una secuencia de ácidos nucleicos que codifica una proteína que comprende una secuencia de aminoácidos representada por SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24. En una realización, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención comprende una secuencia de ácidos nucleicos que se hibrida en condiciones de rigurosidad moderada, elevada o muy elevada al complemento de una secuencia de ácidos nucleicos representada por los nucleótidos 1260-3314 de la SEQ ID NO: 14, nucleótidos 1260-3314 de la SEQ ID NO: 18, nucleótidos 1260-3314 de la SEQ ID NO: 21, o nucleótidos 1260-3314 de la SEQ ID NO: 23.

Como se usa en el presente documento, las condiciones de hibridación se refieren a las condiciones de hibridación convencionales en las que se usan moléculas de ácido nucleico para identificar moléculas de ácido nucleico similares. Tales condiciones convencionales se desvelan, por ejemplo, en Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Labs Press, 1989. Sambrook *et al.*, igual que la referencia mencionada anteriormente, se incorpora como referencia en el presente documento en su totalidad (véanse específicamente las páginas 9.31-9.62). Además, se desvelan fórmulas para calcular las condiciones de hibridación y lavado apropiadas para conseguir una hibridación que permita grados variables de faltas de coincidencias de nucleótidos, por ejemplo, en Meinkoth *et al.*, 1984, Anal. Biochem. 138, 267-284; Meinkoth *et al.*, igual que la referencia mencionada anteriormente, que se incorpora como referencia en el presente documento en su totalidad.

Más particularmente, las condiciones de hibridación y lavado de rigurosidad moderada, tal como se les hace referencia en el presente documento, se refieren a condiciones que permiten el aislamiento de moléculas de ácido nucleico que presentan una identidad de secuencia de ácido nucleico de al menos aproximadamente un 70 % de

identidad con la secuencia de ácido nucleico con la molécula de ácido nucleico que se está usando para estudiar con sonda en la reacción de hibridación (es decir, condiciones que permiten una falta de coincidencia de nucleótidos de aproximadamente un 30 % o inferior). Las condiciones de hibridación y lavado de alta rigurosidad, tal como se les hace referencia en el presente documento, se refieren a condiciones que permiten el aislamiento de moléculas de ácido nucleico que tienen una identidad de secuencia de ácidos nucleicos de al menos aproximadamente un 80 % con la molécula de ácido nucleico que se está usando para estudiar con sonda en la reacción de hibridación (es decir, condiciones que permiten una falta de coincidencia de nucleótidos de aproximadamente un 20 % o inferior).

Las condiciones de hibridación y lavado de rigurosidad muy alta, tal como se les hace referencia en el presente documento, se refieren a condiciones que permiten el aislamiento de moléculas de ácido nucleico que presentan una identidad de secuencia de ácido nucleico de al menos aproximadamente un 90 % con la molécula de ácido nucleico que se está usando para estudiar con sonda en la reacción de hibridación (es decir, condiciones que permiten una falta de coincidencia de nucleótidos de aproximadamente un 10 % o inferior). Tal como se ha analizado anteriormente, un experto en la materia puede usar las fórmulas en Meinkoth *et al.*, igual que la referencia mencionada anteriormente, para calcular las condiciones de hibridación y lavado apropiadas para conseguir estos niveles particulares de falta de coincidencia de nucleótidos. Tales condiciones variarán dependiendo de si se están formando híbridos de ADN:ARN o ADN:ADN. Las temperaturas de fusión calculadas para híbridos de ADN:ADN son 10 °C menores que para híbridos de ADN:ARN. En realizaciones en particular, las condiciones de hibridación rigurosa para híbridos de ADN:ADN incluyen hibridación a una fuerza iónica de 6X SSC (Na⁺ 0,9 M) a una temperatura entre aproximadamente 20 °C y aproximadamente 35 °C (rigurosidad más baja), más preferentemente, entre aproximadamente 28 °C y aproximadamente 40 °C (más rigurosa), e incluso más preferentemente, entre aproximadamente 35 °C y aproximadamente 45 °C (incluso más rigurosa), con condiciones de lavado apropiadas.

En realizaciones en particular, las condiciones y liquidación rigurosas para híbridos de ADN:ARN incluyen una fuerza iónica de 6X SSC (Na⁺ 0,9 M) a una temperatura entre 30 °C y aproximadamente 45 °C, más preferentemente entre aproximadamente 38 °C y aproximadamente 50 °C, e incluso más preferentemente, entre aproximadamente 45 °C y aproximadamente 55 °C, con condiciones de lavado igualmente rigurosas. Estos valores se basan en cálculos de una temperatura de fusión para moléculas mayores de aproximadamente 100 nucleótidos, formamida al 0 % y un contenido de G + C de aproximadamente un 40 %. Como alternativa, la T_m se puede calcular de forma empírica tal como se establece en Sambrook *et al.*, mencionado anteriormente, páginas 9.31 a 9.62. En general, las condiciones de lavado deberían ser tan rigurosas como fuera posible, y deberían ser apropiadas para las condiciones de hibridación elegidas. Por ejemplo, las condiciones de hibridación pueden incluir una combinación de condiciones de sal y temperatura que están aproximadamente 20-25 °C por debajo de la T_m calculada de un híbrido en particular, y las condiciones de lavado incluyen normalmente una combinación de condiciones de sal y temperatura que están aproximadamente 12-20 °C por debajo de la T_m del híbrido en particular. Un ejemplo de condiciones de hibridación adecuadas para su uso con híbridos de ADN:ADN incluye una hibridación de 2-24 horas en 6X SSC (formamida al 50 %) a aproximadamente 42 °C, seguido de etapas de lavado que incluyen uno o más lavados a temperatura ambiente en SSC a aproximadamente 2X, seguido de lavados adicionales a temperaturas más elevadas y fuerza iónica inferior (por ejemplo, al menos un lavado a aproximadamente 37 °C en aproximadamente 0,1X-0,5X SSC, seguido por al menos un lavado a aproximadamente 68 °C en aproximadamente 0,1X-0,5X SSC).

En otra realización, las moléculas de ácido nucleico adecuadas para uso de acuerdo con los métodos de la presente invención que codifican una acetolactato sintasa adecuada para uso de acuerdo con los métodos de la presente invención incluyen moléculas de ácido nucleico aisladas que comprenden una secuencia de ácidos nucleicos que codifica una proteína que tiene una secuencia de aminoácidos que comprende al menos 30 restos de aminoácidos contiguos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24, (es decir, 30 restos de aminoácidos contiguos que tiene una identidad de un 100 % con 30 aminoácidos contiguos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24). En una realización preferente, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención comprende una secuencia de ácidos nucleicos que codifica una proteína que tiene una secuencia de aminoácidos que comprende al menos 50, y más preferentemente al menos 75, y más preferentemente al menos 100, y más preferentemente al menos 115, y más preferentemente al menos 130, y más preferentemente al menos 150, y más preferentemente al menos 200, y más preferentemente, al menos 250, y más preferentemente, al menos 300, y más preferentemente, al menos 350, y más preferentemente, al menos 400, y más preferentemente, al menos 450, y más preferentemente, al menos 500, y más preferentemente, al menos 550, y más preferentemente, al menos 600, y más preferentemente, al menos 650, restos de aminoácidos contiguos de cualquiera de SEQ ID NO: 15, SEQ ID NO: 19, SEQ ID NO: 22 o SEQ ID NO: 24. Una proteína de ese tipo tiene actividad biológica de acetolactato sintasa. En una realización, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención que codifica una acetolactato sintasa comprende una secuencia de ácidos nucleicos que tiene al menos 60 nucleótidos contiguos, y más preferentemente al menos 150, y más preferentemente al menos 225, y más preferentemente al menos 300, y más preferentemente al menos 345, y más preferentemente al menos 390, y más preferentemente al menos 450, y más preferentemente al menos 525, y más preferentemente al menos 600, y más preferentemente al menos 750, y más preferentemente al menos 900, y más preferentemente al menos 1050, y más preferentemente al menos 1200, y más preferentemente al menos 1350, y más preferentemente al menos 1500, y más preferentemente al menos 1650, y más preferentemente al menos 1800, e incluso más preferentemente al menos 1950, nucleótidos contiguos

de los nucleótidos 1260-3314 de la SEQ ID NO: 15, nucleótidos 1260-3314 de la SEQ ID NO: 18, nucleótidos 1260-3314 de la SEQ ID NO: 21, o nucleótidos 1260-3314 de la SEQ ID NO: 23.

Las moléculas de ácido nucleico particularmente preferentes adecuadas para uso de acuerdo con los métodos de la presente invención incluyen los nucleótidos 1260-3314 de la SEQ ID NO: 14 (codifica la SEQ ID NO: 15), nucleótidos 1260-3314 de la SEQ ID NO: 18 (codifica la SEQ ID NO: 19), nucleótidos 1260-3314 de la SEQ ID NO: 21 (codifica la SEQ ID NO: 22), o nucleótidos 1260-3314 de la SEQ ID NO: 23 (codifica la SEQ ID NO: 24), SEQ ID NO: 14, SEQ ID NO: 18, SEQ ID NO: 21 o SEQ ID NO: 23.

De acuerdo con los métodos de la presente invención, una molécula de ácido nucleico aislada es una molécula de ácido nucleico que se ha retirado de su medio natural (es decir, que se ha sometido a manipulación humana), siendo su medio natural el genoma o cromosoma en el que la molécula de ácido nucleico se encuentra en la naturaleza. Como tal, "aislado" no refleja necesariamente el grado hasta el que la molécula de ácido nucleico se ha purificado, sino que indica que la molécula no incluye un genoma completo o un cromosoma completo en el que la molécula de ácido nucleico se encuentra en la naturaleza. Una molécula de ácido nucleico aislada puede incluir un gen, tal como un gen de acetolactato sintasa que se describe en el presente documento. Una molécula de ácido nucleico aislada que incluye un gen no es un fragmento de un cromosoma que incluya tal gen, sino que más bien incluye la región codificante y las regiones reguladoras asociadas con el gen, sin genes adicionales que se encuentran de manera natural en el mismo cromosoma. Una molécula de ácido nucleico aislada también puede incluir una secuencia de ácidos nucleicos especificada flanqueada por (es decir, en el extremo en la posición 5' y/o 3' terminales de la secuencia) ácidos nucleicos adicionales que no flanquean normalmente la secuencia de ácidos nucleicos especificada en la naturaleza (es decir, son secuencias heterólogas). La molécula de ácido nucleico aislada puede incluir ADN, ARN (por ejemplo, ARNm), o derivados de ADN o ARN (por ejemplo, ADNc). Aunque la expresión "molécula de ácido nucleico" se refiere principalmente a la molécula de ácido nucleico física y la expresión "secuencia de ácidos nucleicos" se refiere principalmente a la secuencia de nucleótidos en la molécula de ácido nucleico, las dos frases se pueden usar de manera intercambiable, especialmente con respecto a una molécula de ácido nucleico, o una secuencia de ácidos nucleicos, que puede codificar para una proteína.

Preferentemente, una molécula de ácido nucleico aislada adecuada para uso de acuerdo con los métodos de la presente invención se produce usando tecnología de ADN recombinante (por ejemplo, amplificación por reacción en cadena de la polimerasa (PCR), clonación) o síntesis química. Las moléculas de ácido nucleico aisladas incluyen moléculas de ácido nucleico naturales y homólogos de las mismas, que comprenden de manera no limitante variantes alélicas naturales y moléculas de ácido nucleico modificadas en las que se han insertado, sometido a delección, sustituido y/o invertido nucleótidos de una manera tal que tales modificaciones proporcionan el efecto deseado con respecto a la actividad biológica de la proteína. Anteriormente se ha expuesto con detalle variantes alélicas y homólogos de proteína (por ejemplo, proteínas codificadas por homólogos de ácido nucleico).

Un homólogo de molécula de ácido nucleico se puede producir usando una serie de métodos conocidos por los expertos en la materia (véase, por ejemplo, Sambrook *et al.*, igual que la referencia mencionada anteriormente.) Por ejemplo, las moléculas de ácido nucleico se pueden modificar usando una diversidad de técnicas que incluyen, pero no se limitan, técnicas de mutagénesis clásicas y técnicas de ADN recombinante, tales como mutagénesis dirigida al sitio, tratamiento químico de una molécula de ácido nucleico para inducir mutaciones, escisión con enzimas de restricción de un fragmento de ácido nucleico, ligación de fragmentos de ácido nucleico, amplificación por PCR y/o mutagénesis de regiones seleccionadas de una secuencia de ácidos nucleicos, síntesis de mezclas de oligonucleótidos y ligación de grupos de mezcla para "construir" una mezcla de moléculas de ácido nucleico y combinaciones de las mismas. Los homólogos de moléculas de ácido nucleico se pueden seleccionar a partir de una mezcla de ácidos nucleicos modificados mediante identificación sistemática de la función de la proteína codificada por el ácido nucleico y/o mediante hibridación con un gen de tipo silvestre.

De forma análoga, el tamaño mínimo de una molécula de ácido nucleico adecuada para uso de acuerdo con los métodos de la presente invención es un tamaño suficiente para codificar una proteína que presenta la actividad biológica deseada, o suficiente para formar una sonda o cebador de oligonucleótido que puede formar un híbrido estable con la secuencia complementaria de una molécula de ácido nucleico que codifica la proteína natural (por ejemplo, en condiciones de rigurosidad moderada, elevada o muy elevada). Como tal, el tamaño de la molécula de ácido nucleico que codifica una proteína de este tipo puede depender de la composición de ácido nucleico y el porcentaje de homología o identidad entre la molécula de ácido nucleico y la secuencia complementaria así como de las condiciones de hibridación *per se* (por ejemplo, temperatura, concentración de sal y concentración de formamida). El tamaño mínimo de una molécula de ácido nucleico que se usa como cebador de oligonucleótido o como sonda es normalmente de al menos aproximadamente 12 a aproximadamente 15 nucleótidos de longitud si las moléculas de ácido nucleico son ricas en GC y de al menos aproximadamente 15 a aproximadamente 18 bases de longitud si son ricas en AT. No existe ningún límite, distinto de un límite práctico, en el tamaño máximo de una molécula de ácido nucleico adecuada para uso de acuerdo con los métodos de la presente invención, porque la molécula de ácido nucleico puede incluir una parte de una secuencia que codifica una proteína (por ejemplo, una secuencia que codifica acetolactato sintasa) o una secuencia de ácidos nucleicos que codifica una proteína de longitud completa.

Una realización de los métodos de la presente invención incluye un vector recombinante que se va a usar para la transformación de un microorganismo de Thraustochytriales. De acuerdo con los métodos de la presente invención, un vector recombinante es una molécula de ácido nucleico diseñada por ingeniería genética (es decir, producida artificialmente) que se usa como herramienta para manipular una secuencia de ácido nucleico de elección y para introducir una secuencia de ácidos nucleicos de este tipo en una célula hospedadora. Por lo tanto el vector recombinante es apropiado para su uso en clonación, secuenciación y/o de otro modo manipulación de la secuencia de ácidos nucleicos de elección, tal como mediante expresión y/o administración de la secuencia de ácidos nucleicos de elección a una célula hospedadora para formar una célula recombinante. Un vector de este tipo contiene normalmente secuencias de ácidos nucleicos heterólogas, es decir, secuencias de ácidos nucleicos que no se encuentran de manera natural adyacentes a una secuencia de ácidos nucleicos que se va a administrar, aunque el vector también puede contener secuencias de ácidos nucleicos reguladoras (por ejemplo, promotores, regiones no traducidas) que se encuentran de manera natural adyacentes a moléculas de ácido nucleico de la presente invención (tratado en detalle a continuación). El vector puede ser ARN o ADN, ya sea procariota o eucariota, y normalmente es un plásmido. El vector se puede mantener como un elemento extracromosómico (por ejemplo, un plásmido) o se puede integrar en el cromosoma del microorganismo recombinante. Todo el vector puede permanecer en su sitio dentro de una célula hospedadora o, en ciertas condiciones, el ADN de plásmido se puede someter a delección, dejando atrás la molécula de ácido nucleico de la presente invención. La molécula de ácido nucleico integrada puede estar bajo el control de un promotor cromosómico, bajo el control de un promotor de plásmido o nativo, o bajo una combinación de varios controles de promotores. En el cromosoma se pueden integrar copias únicas o múltiples de la molécula de ácido nucleico. Un vector recombinante adecuado para uso de acuerdo con los métodos de la presente invención contiene al menos un marcador seleccionable para microorganismos de Thraustochytriales adecuado para uso de acuerdo con los métodos de la presente invención, tal como una secuencia de ácidos nucleicos que codifica una acetolactato sintasa de Thraustochytriales (homólogo o proteína natural) o una secuencia de ácidos nucleicos que codifica el gen *b1e* (que se describe a continuación). Como se usa en el presente documento, la expresión "molécula de ácido nucleico recombinante" se usa principalmente para hacer referencia a un vector recombinante en el que se ha ligado la secuencia de ácidos nucleicos que se va a clonar, manipular, transformar en la célula hospedadora (es decir, la inserción).

Normalmente, un vector recombinante, y por lo tanto una molécula de ácido nucleico recombinante, incluye al menos una molécula de ácido nucleico adecuada para uso de acuerdo con los métodos de la presente invención unida de forma operativa a una o más secuencias de control de la transcripción. Como se usa en el presente documento, la expresión "molécula recombinante" o "molécula de ácido nucleico recombinante" se refiere principalmente a una molécula de ácido nucleico o secuencia de ácidos nucleicos unida de forma operativa a una secuencia de control de la transcripción, pero puede se puede usar de manera intercambiable con la expresión "molécula de ácido nucleico", cuando tal molécula de ácido nucleico es una molécula recombinante como se discute en el presente documento. De acuerdo con los métodos de la presente invención, la expresión "unido de forma operativa" se refiere a la unión de una molécula de ácido nucleico a una secuencia de control de transcripción de una manera tal que la molécula se puede expresar cuando se transfecta (es decir, se transforma, transduce, transfecta, conjuga o conduce) a una célula hospedadora. Las secuencias de control de la transcripción son secuencias que controlan el inicio, elongación o terminación de la transcripción. Las secuencias de control de la transcripción particularmente importantes son las que controlan el inicio de la transcripción, tales como secuencias promotoras, potenciadoras, de operador o represoras. Las secuencias de control de la transcripción adecuadas incluyen cualquier secuencia de control de la transcripción que puede funcionar en un microorganismo del orden Thraustochytriales. Se cree que los presentes inventores son los primeros en aislar e identificar al menos tres de tales promotores, que se describen con detalle en cualquier parte en el presente documento.

Los promotores precedentes incluyen, pero no se limitan a, un promotor de acetolactato sintasa de Thraustochytriales (representado en el presente documento por los nucleótidos 1-1259 de la SEQ ID NO: 14), un promotor de α -tubulina de Thraustochytriales (representado en el presente documento por los nucleótidos 441-894 de la SEQ ID NO: 9, un promotor de un sistema de policétido sintasa de Thraustochytriales (PKS) (contenido dentro de la SEQ ID NO: 34), y un promotor de ácido graso desaturasa de Thraustochytriales (contenido dentro de la SEQ ID NO: 31; se indica que el promotor de la ácido graso desaturasa se denomina promotor de desaturasa en la Solicitud Provisional U.S. con N.º de Serie 60/284.116, mencionada anteriormente). La clonación y secuenciación del promotor de α -tubulina, el promotor de acetolactato sintasa, y el promotor de ácido graso desaturasa se describen en la sección de Ejemplos. En una realización preferente, el promotor de α -tubulina comprende la secuencia de promotor de α -tubulina de Thraustochytriales que se produce de manera natural (nucleótidos 441-894 de la SEQ ID NO: 9), o una secuencia de ácidos nucleicos que es idéntica en al menos un 95 % a los nucleótidos 441-894 de la SEQ ID NO: 9, en la que el promotor tiene al menos actividad de transcripción de promotor de α -tubulina basal. De forma análoga, un promotor de acetolactato sintasa preferente comprende una secuencia de ácidos nucleicos del promotor de acetolactato sintasa de Thraustochytriales que se produce de manera natural (representado dentro de los nucleótidos 1-1259 de la SEQ ID NO: 14), o una secuencia de ácidos nucleicos que es idéntica en al menos un 75 %, y más preferentemente un 80 %, y más preferentemente un 85 %, y más preferentemente un 90 %, y más preferentemente un 95 % a los nucleótidos 1-1259 de la SEQ ID NO: 14, en la que el promotor tiene una actividad de transcripción de promotor de acetolactato sintasa al menos basal. Un promotor de PKS preferente comprende una secuencia de ácidos nucleicos de un promotor de PKS de Thraustochytriales que se produce de manera natural (Representado dentro de la SEQ ID NO: 34), o una secuencia de nucleótidos que es idéntica en al menos un 75 %, y

más preferentemente un 80 %, y más preferentemente un 85 %, y más preferentemente un 90 %, y más preferentemente un 95 % a la SEQ ID NO: 34, en la que el promotor tiene actividad de transcripción de promotor de PKS al menos basal. Por último, un promotor de ácido graso desaturasa preferente comprende una secuencia de ácidos nucleicos del promotor de ácido graso desaturasa de *Thraustochytriales* que se produce de manera natural (Representado dentro de la SEQ ID NO: 31), o está contenido dentro de una secuencia de ácidos nucleicos que es idéntica en al menos un 75 %, y más preferentemente un 80 %, y más preferentemente un 85 %, y más preferentemente un 90 %, y más preferentemente un 95 % a la SEQ ID NO: 31, en la que el promotor tiene actividad de transcripción de promotor de ácido graso desaturasa al menos basal. En el presente documento se han descrito anteriormente métodos para determinar el porcentaje de identidad para las secuencias de acetolactato sintasa, y se incluyen en el presente documento.

En una realización, un vector recombinante adecuado para uso de acuerdo con los métodos de la presente invención es un vector de expresión. Como se usa en el presente documento, la expresión "vector de expresión" se usa para hacer referencia a un vector y es adecuado para producción de un producto codificado (por ejemplo, una proteína de interés). En esta realización, una secuencia de ácidos nucleicos que codifica el producto a producirse se inserta en el vector recombinante para producir una molécula de ácido nucleico recombinante. La secuencia de ácidos nucleicos que codifica la proteína a producir se inserta en el vector de un modo tal que une de forma operativa a la secuencia de ácidos nucleicos a las secuencias reguladoras en el vector (por ejemplo, un promotor de *Thraustochytriales* de la presente invención) que emite la transcripción y traducción de la secuencia de ácidos nucleicos dentro del microorganismo recombinante. Los marcadores seleccionables que se describen en el presente documento permiten la selección de un microorganismo recombinante en el que una molécula de ácido nucleico recombinante adecuada para uso de acuerdo con los métodos de la presente invención se ha introducido de forma satisfactoria.

Como se describe en el presente documento, un vector recombinante de la presente descripción es un vector de direccionamiento. Como se usa en el presente documento, la expresión "vector de direccionamiento" se usa para hacer referencia a un vector que se usa para suministrar una molécula de ácido nucleico en particular en una célula recombinante, en el que la molécula de ácido nucleico se usa para hacer delección o inactivar un gen endógeno dentro de la célula hospedadora (es decir, se usa para tecnología de alteración o desactivación genética dirigida). Un vector de este tipo también se puede conocer en la técnica como vector de "desactivación genética". Como se describe en el presente documento, una parte del vector, pero más normalmente, la molécula de ácido nucleico insertada en el vector (es decir, la inserción), presenta una secuencia de ácidos nucleicos que es homóloga a una secuencia de ácido nucleico de un gen diana en la célula hospedadora (es decir, un gen que se selecciona como diana para su delección o inactivación). La secuencia de ácidos nucleicos de la inserción del vector se diseñó para que se unan al gen diana de manera que el gen diana y la inserción experimentan recombinación homóloga, mediante lo cual el gen diana endógeno sufre delección, inactivación o atenuación (es decir, en al menos una parte del gen diana endógeno que se está mutando o sometiendo a delección).

En una realización, un vector recombinante preferente adecuado para uso de acuerdo con los métodos de la presente invención es un vector recombinante que es adecuado para uso en un microorganismo de *Thraustochytriales*, y que incluye una secuencia de ácidos nucleicos que codifica una molécula de acetolactato Sintasa adecuada para uso de acuerdo con los métodos de la presente invención. Preferentemente, la acetolactato sintasa se modifica en comparación con la forma que se produce de manera natural (SEQ ID NO: 15), de modo que la sintasa confiere a un microorganismo de *Thraustochytriales* que se ha transfectado con un vector recombinante, una reducción de la sensibilidad a compuestos de sulfonilurea, inhibidores de la clase de imidazolinona y/u oxibenzoatos de pirimidinilo. Preferentemente, un vector recombinante de este tipo comprende una secuencia de ácidos nucleicos que codifica una proteína de acetolactato sintasa que comprende una secuencia de aminoácidos que difiere de la SEQ ID NO: 15 en una sustitución, inserción o delección de aminoácido en una o más de las siguientes posiciones: 116G, 117A, 192P, 200A, 251K, 358M, 383D, 592V, 595W, o 599F. En una realización, tales proteínas de acetolactato sintasa presentan una secuencia de aminoácidos que incluye, pero no se limita a: SEQ ID NO: 19, SEQ ID NO: 22 y SEQ ID NO: 24. Preferentemente, un vector recombinante de este tipo comprende una secuencia de ácidos nucleicos seleccionada entre: nucleótidos 1260-3314 de la SEQ ID NO: 18, nucleótidos 1260-3314 de la SEQ ID NO: 21, y nucleótidos 1260-3314 de la SEQ ID NO: 23. En una realización particularmente preferente, los vectores recombinantes que contienen secuencias de ácidos nucleicos que codifican ALS que confieren la resistencia deseada incluyen SEQ ID NO: 18, SEQ ID NO: 21 y SEQ ID NO: 23.

En una forma de realización, un vector recombinante que es adecuado para uso de acuerdo con los métodos de la presente invención porque confiere a un microorganismo de *Thraustochytriales* S.A. transfectar no con el vector recombinante una reducción de la sensibilidad a compuestos de sulfonilurea, inhibidores de la clase de imidazolinona y/u oxibenzoatos de pirimidinilo, comprende la SEQ ID NO: 15, que es la secuencia de acetolactato sintasa *Schizochytrium* que se produce de manera natural. En esta realización, el vector recombinante se diseña para sobreexpresar la sintasa que se produce de manera natural, mediante lo cual tal sobreexpresión presenta el efecto de conferir resistencia a los compuestos especificados en el microorganismo. En esta realización, un experto en la materia observará que el uso de tecnología de ADN recombinante puede mejorar el control de la expresión de moléculas de ácido nucleico transformadas manipulando, por ejemplo, el número de copias de las moléculas de ácido nucleico dentro de la célula hospedadora, la eficacia con la que esas moléculas de ácido nucleico se

transcriben, la eficacia con la que las transcripciones resultantes se traducen, y la eficacia de las modificaciones posteriores a la traducción. Adicionalmente, la secuencia promotora se podría modificar por ingeniería genética para mejorar el nivel de expresión en comparación con el promotor nativo. Las técnicas recombinantes útiles para controlar la expresión de moléculas de ácido nucleico comprenden incluyen, pero no se limitan a, integración de las moléculas de ácido nucleico en uno o más cromosomas celulares, adición de secuencias de estabilidad de vector a plásmidos, sustituciones o modificaciones de señales de control de la transcripción (por ejemplo, promotores, operadores, potenciadores), sustituciones o modificaciones de señales de control de la traducción (por ejemplo, sitios de unión a ribosomas, secuencias de Shine-Dalgarno), modificación de moléculas de ácido nucleico para que correspondan al uso de codones de la célula hospedadora, y delección de secuencias que desestabilizan transcripciones.

En una realización, un vector recombinante adecuado para uso de acuerdo con los métodos de la presente invención en la transformación de microorganismos de Thraustochytriales contiene el gen *Sh ble* de *Streptoalloteichus hindustanus* como marcador seleccionable (que codifica una "proteína de unión a bleomicina"), en combinación con un promotor de Thraustochytriales como se ha descrito anteriormente en el presente documento. Un vector recombinante preferente que comprende el gen *ble* y un promotor de Thraustochytriales incluye, por ejemplo, la secuencia de vector representada por la SEQ ID NO: 8 o 9. La secuencia de aminoácidos de la proteína de unión a bleomicina de *Streptoalloteichus hindustanus* se representa en el presente documento como la SEQ ID NO: 10.

Las moléculas de ácido nucleico recombinante adecuadas para uso de acuerdo con los métodos de la presente invención, que pueden ser cualquiera de ADN o ARN, también deben contener secuencias reguladoras adicionales, tales como secuencias reguladoras de la traducción, orígenes de replicación y otras secuencias reguladoras que son compatibles con la célula recombinante. En una realización, una molécula recombinante de la presente invención, incluyendo las que se integran en el cromosoma de la célula hospedadora, también contiene señales secretoras (es decir, secuencias de ácido nucleico de segmento señal) para permitir que una proteína expresada se secrete de la célula que produce la proteína. Los segmentos de señal adecuados incluyen un segmento de señal que está asociado de manera natural con la proteína que se va a expresar o cualquier segmento de señal heterólogo que puede dirigir la secreción de la proteína de acuerdo con los métodos de la presente invención. En otra realización, una molécula recombinante adecuada para uso de acuerdo con los métodos de la presente invención comprende una secuencia líder para permitir que una proteína expresada se suministre a y se inserte en la membrana de una célula hospedadora. Las secuencias líder adecuadas incluyen una secuencia líder que está asociada de manera natural con la proteína, o cualquier secuencia líder heteróloga que puede dirigir el suministro y la inserción de la proteína a la membrana de una célula.

Habiendo descrito diversas herramientas que son útiles para transformar microorganismos del orden Thraustochytriales, en una realización, adecuada para uso de acuerdo con los métodos de la presente invención se presenta un método para la transformación de células de un microorganismo del orden Thraustochytriales. El método incluye una primera etapa de: (a) introducir en células de un microorganismo del orden Thraustochytriales una molécula de ácido nucleico recombinante tal como se ha descrito anteriormente en el presente documento. La molécula de ácido nucleico recombinante comprende una secuencia de ácidos nucleicos que codifica una acetolactato sintasa que confiere a dichas células una reducción de la sensibilidad a compuestos seleccionados de entre el grupo constituido por: compuestos de sulfonilurea, inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo. Tales acetolactato sintasas se han descrito en detalle anteriormente e incluyen las acetolactato sintasas que presentan una secuencia de aminoácidos representada por las SEQ ID NO: 19, SEQ ID NO: 22 y SEQ ID NO: 24, así como homólogos de cualquiera de las secuencias de ese tipo o SEQ ID NO: 15 como se ha discutido anteriormente. El método incluye una segunda etapa que consiste en: (b) seleccionar células que se han transformado satisfactoriamente con la molécula de ácido nucleico recombinante cultivando las células de (a) en un medio que contiene al menos un compuesto que es inhibidor para células no transformadas, seleccionándose el compuesto entre el grupo que consiste en: un compuesto de sulfonilurea, un inhibidor de la clase de imidazolinona y oxibenzoatos de pirimidinilo. Las células que crecen en presencia de tales compuestos se han transformado satisfactoriamente.

La molécula de ácido nucleico recombinante usada en el presente método comprende cualquiera de los vectores recombinantes que se han descrito anteriormente en el presente documento, y normalmente incluye al menos una secuencia de ácidos nucleicos que codifica una proteína que va a ser producida por la célula recombinante (es decir, que comprende un vector de expresión recombinante), o una secuencia de ácidos nucleicos útil para la delección o inactivación dirigida de un gen endógeno en la célula recombinante (es decir, que comprende un vector de direccionamiento recombinante).

En una realización, la molécula de ácido nucleico recombinante adecuada para uso de acuerdo con los métodos de la presente invención comprende adicionalmente una secuencia de ácidos nucleicos que codifica una proteína a producir por la célula, en la que la secuencia de ácidos nucleicos que codifica para la proteína está unida de forma operativa a una secuencia de control de la transcripción. Los expertos en la materia conocerán proteínas que puede ser deseables para su producción en Thraustochytriales, y todas pretenden quedar incluidas por los métodos de la presente invención. Las proteínas particularmente preferentes incluyen, pero no se limitan a, proteínas asociadas

con la síntesis de un ácido graso seleccionado entre el grupo que consiste en ácido docosahexaenoico (DHA), ácido docosapentaenoico (DPA), ácido eicosapentaenoico (EPA) y ácido araquidónico (ARA). Tales proteínas incluyen, por ejemplo, una ácido graso sintasa, una ácido graso desaturasa, una ácido graso elongasa, una proteína asociada con un complejo de policétido sintasa y una proteína asociada con la incorporación de ácidos grasos a fosfolípidos o a moléculas de triacilglicerol. En un aspecto, la proteína es un ácido graso ω -3 desaturasa codificada por una secuencia de ácidos nucleicos de SEQ ID NO: 29. La SEQ ID NO: 30 representa la secuencia de aminoácidos de la desaturasa. En otro aspecto, la proteína es un ácido graso polienoico isomerasa. En una realización, las proteínas que se pueden producir en microorganismos de Thraustochytriales usando el presente método incluyen proteínas asociadas con las rutas biosintéticas de isoprenoides. Tales proteínas comprenden incluyen, pero no se limitan a HMG-CoA sintasa, HMG-CoA reductasa, escualeno sintasa, fitoeno sintasa, fitoeno desaturasa, una carotenoide ciclasa, una carotenoide hidroxilasa, una carotenoide cetolasa. En otra realización más, las proteínas que se pueden producir en microorganismos de Thraustochytriales usando el presente método incluyen, pero no se limitan a, vitamina E y ácido lipoico.

Como se describe en el presente documento, la molécula de ácido nucleico recombinante útil en el método de la presente descripción incluye una secuencia de ácidos nucleicos que se hibrida con una secuencia de ácidos nucleicos diana en el microorganismo de manera que un gen que comprende la secuencia de ácidos nucleicos diana se muta o inactiva mediante recombinación homóloga con la segunda secuencia de ácidos nucleicos. Una secuencia de ácidos nucleicos de este tipo puede ser homóloga con respecto a genes que codifican enzimas (o ácidos nucleicos que regulan la expresión de tales genes) de las rutas de síntesis de ácidos grasos saturados y poliinsaturados, genes que codifican proteínas que están implicadas en la degradación de otros compuestos valiosos producidos por el microorganismo de Thraustochytriales que de otra manera disminuyen el valor del compuesto deseado, o genes que codifican proteínas que están asociadas con la síntesis de compuestos cuya síntesis está en competición con otras moléculas de interés. Por ejemplo, las secuencias de ácidos nucleicos diana incluyen, pero no se limitan a, secuencias que codifican lipasas, enzimas de oxidación de ácidos grasos, proteínas implicadas en la síntesis de hidratos de carbono, proteínas implicadas en la síntesis de productos de rutas de isoprenoides, proteínas implicadas en la síntesis de componentes de la pared celular, proteínas implicadas en las rutas de síntesis de ácidos grasos saturados, proteínas implicadas en las rutas de síntesis de ácidos grasos poliinsaturados, proteínas asociadas con un complejo de policétido sintasa y proteínas asociadas con la incorporación de ácidos grasos a fosfolípidos o moléculas de triacilglicerol.

En una realización de los métodos de la presente invención, el método para la transformación de microorganismos de Thraustochytriales incluye una etapa que consiste en introducir en la célula al menos una molécula de ácido nucleico recombinante adicional que comprende una secuencia de ácidos nucleicos que codifica una proteína que se va a expresar, estando la secuencia de ácido nucleico unida de forma operativa a una secuencia de control de la transcripción. Como alternativa, la molécula de ácido nucleico recombinante adicional puede incluir una segunda secuencia de ácidos nucleicos que se hibrida con una secuencia de ácidos nucleicos diana en el microorganismo de manera que un gen que comprende la secuencia de ácido nucleico diana se muta o inactiva mediante recombinación homóloga con la segunda secuencia de ácidos nucleicos. De esta manera, se pueden introducir múltiples proteínas en las células, se pueden inactivar múltiples genes, o son posibles combinaciones de los dos. La molécula de ácido nucleico recombinante adicional se puede introducir en el microorganismo de Thraustochytriales de forma simultánea con la primera molécula de ácido nucleico recombinante (es decir, cotransformación), o como una transformación posterior (por ejemplo, con los fines de rasgos de "apilamiento").

En una realización, el método de la presente invención incluye adicionalmente la etapa de introducir en la célula al menos una molécula de ácido nucleico recombinante adicional que comprende una secuencia de ácidos nucleicos que comprende una proteína de unión a bleomicina. En esta realización, la molécula de ácido nucleico recombinante adicional se introduce preferentemente en una etapa posterior, en vez de como una cotransformación. Preferentemente, la molécula de ácido nucleico recombinante que comprende una secuencia de ácidos nucleicos que codifica una proteína de unión a bleomicina comprende además una secuencia de ácidos nucleicos que codifica una segunda proteína que se va a expresar por la célula, en la que la secuencia de ácido nucleico que codifica la segunda proteína está unida de forma operativa a una secuencia de control de la transcripción. Como alternativa, o además, la molécula de ácido nucleico recombinante que comprende una secuencia de ácidos nucleicos que codifica una proteína de unión a bleomicina comprende además una segunda secuencia de ácidos nucleicos que se hibrida con una secuencia de ácidos nucleicos diana en el microorganismo de manera que un gen que comprende la secuencia de ácidos nucleicos diana se muta o inactiva mediante recombinación homóloga con la segunda secuencia de ácidos nucleicos. Como se describe en el presente documento, una molécula de ácido nucleico recombinante de este tipo comprende una secuencia de ácidos nucleicos de SEQ ID NO: 9.

Las células hospedadoras adecuadas para su transformación para uso de acuerdo con los métodos de la presente invención incluyen, pero no se limitan a, cualquier microorganismo del orden Thraustochytriales. Las células hospedadoras pueden ser células no transformadas o células que ya se han transfectado con al menos una molécula de ácido nucleico. Las células hospedadoras preferentes para uso de acuerdo con los métodos de la presente invención incluyen microorganismos de un género que incluye, pero no se limita a: *Thraustochytrium*, *Labyrinthuloides*, *Japonochytrium*, y *Schizochytrium*. Las especies preferentes 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 de *Ulkenia* anteriores tales como *U. visurgensis*, *U. amoebaida*, *U. sarkariana*, *U. profunda*, *U. radiata*, *U. minuta* y *Ulkenia sp.* BP-5601), e incluyendo *Thraustochytrium striatum*, *Thraustochytrium aureum*, *Thraustochytrium roseum*; y cualquier especie de *Japonochytrium*. Las cepas particularmente preferentes 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 y Belsky) (ATCC 28209); *Schizochytrium limacinum* (Honda y 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).

De acuerdo con los métodos de la presente invención, el término "transformación" se usa para hacer referencia a cualquier método mediante el cual una molécula de ácido nucleico exógena (es decir, una molécula de ácido nucleico recombinante) se puede insertar en células microbianas, tales como células microbianas de Thraustochytriales. En los sistemas microbianos, el término "transformación" se usa para describir un cambio heredado debido a la adquisición de ácidos nucleicos exógenos por el microorganismo y es esencialmente sinónimo al término "transfección". Las técnicas de transformación adecuada se incluyen, pero no se limitan a, bombardeo de partículas, electroporación, microinyección, lipofección, adsorción, infección y fusión de protoplastos.

En una realización, una proteína que se va a producir usando un método de la presente invención se produce cultivando una célula que expresa la proteína (es decir, un microorganismo de Thraustochytriales recombinante) en condiciones eficaces para producir la proteína. En algunos casos, la proteína se puede recuperar, y en otros, el microorganismo se puede recoger como un todo o como un lisado y se puede usar como una "biomasa". Como se describe en el presente documento, un gen diana se somete a delección o se inactiva cultivando una célula que se ha transformado con una molécula recombinante que comprende un vector de direccionamiento de la presente invención en condiciones eficaces para permitir la recombinación dentro de la célula, dando como resultado la delección o inactivación de un gen diana. Una célula preferente para cultivar es una célula recombinante que se describe en el presente documento. Las condiciones de cultivo eficaces incluyen, pero no se limitan a, medios eficaces, biorreactor, temperatura, pH y condiciones de oxígeno que permiten la producción y/o la recombinación de proteína. Un medio eficaz se refiere a cualquier medio en el que se cultiva normalmente una célula de Thraustochytriales. Tal medio comprende normalmente un medio acuoso que presenta fuentes de carbono, nitrógeno y fosfato asimilables, y sales, minerales, metales y otros nutrientes apropiados, tales como vitaminas. Los ejemplos de medios y condiciones de cultivo adecuados se discuten con detalle en la sección de Ejemplos. Las condiciones de cultivo adecuadas para microorganismos de Thraustochytriales también se describen en el documento de Patente de Estados Unidos N.º 5.340.742, concedido el 23 de agosto de 1994, a Barclay. Las células adecuadas para uso de acuerdo con los métodos de la presente invención se pueden cultivar en biorreactores de fermentación convencionales, matraces de agitación, tubos de ensayo, placas de microtitulación y placas de Petri convencionales. El cultivo se puede realizar a una temperatura, un pH y un contenido de oxígeno apropiados para una célula recombinante. Tales condiciones de cultivo están dentro de la experiencia de alguien con una experiencia habitual en la materia.

Dependiendo del vector y sistema de hospedador usados para la producción, las proteínas resultantes de los métodos de la presente invención pueden o bien permanecer dentro de la célula recombinante; se pueden secretar al medio de fermentación; se pueden secretar al espacio entre dos membranas celulares; o se pueden retener en la superficie externa de una membrana celular. La expresión "recuperación de la proteína" se refiere a recoger el medio de fermentación completo que contiene la proteína y no implica necesariamente etapas adicionales de separación o purificación. Las proteínas producidas los métodos de la presente invención se pueden purificar usando una diversidad de técnicas de purificación de proteínas convencionales, tales como, pero que no se limitan a, cromatografía de afinidad, cromatografía de intercambio iónico, filtración, electroforesis, cromatografía de interacción hidrófoba, cromatografía de filtración en gel, cromatografía en fase inversa, cromatografía de concanavalina A, cromatografía de solubilización diferencial. Las proteínas producidas con los métodos de la presente invención se recuperan preferentemente en forma "sustancialmente pura". Como se usa en el presente documento, "sustancialmente pura" se refiere a una pureza que permite el uso eficaz de la proteína como producto comercial.

En otra realización más, un microorganismo recombinante del orden Thraustochytriales adecuado para uso de acuerdo con los métodos de la presente invención, incluye un microorganismo que se ha transformado un microorganismo que se ha transformado con una molécula de ácido nucleico recombinante con una molécula de ácido nucleico recombinante que comprende una secuencia de ácidos nucleicos que codifica una acetolactato sintasa que se describe en el presente documento. Preferentemente, la acetolactato sintasa confiere al microorganismo una reducción de la sensibilidad a compuestos seleccionados entre el grupo que consiste en: compuestos de sulfonilurea, inhibidores de la clase de imidazolinona y oxibenzoatos de pirimidinilo. Las secuencias y moléculas de ácido nucleico recombinantes para su uso en la transformación de un microorganismo de este tipo se han descrito con detalle anteriormente. Un microorganismo de este tipo se puede transformar adicionalmente con otras moléculas de ácido nucleico recombinante, incluyendo moléculas de ácido nucleico recombinante que comprenden un marcador seleccionable de gen *ble* y secuencias de control de transcripción de Thraustochytriales. Como se ha descrito anteriormente en el presente documento. Los microorganismos de Thraustochytriales recombinantes adecuados para uso de acuerdo con los métodos de la presente invención se describen en la sección

de Ejemplos. De acuerdo con los métodos de la presente invención, un microorganismo de Thraustochytriales recombinante adecuado para uso de acuerdo con los métodos de la presente invención se modifica mediante ingeniería genética para expresar una proteína de interés (los ejemplos de tales proteínas se han discutido anteriormente) usando los vectores recombinantes que se describen en el presente documento.

Como se usa en el presente documento, un microorganismo recombinante presenta un genoma que está modificado (es decir, mutado o cambiado) con respecto a su forma normal (es decir, de tipo silvestre o que se produce de manera natural) usando tecnología recombinante. En el presente documento se describe microorganismo recombinante en el que se han insertado, sometido a delección o modificado moléculas de ácido nucleico (es decir, mutado; por ejemplo, mediante inserción, delección, sustitución y/o inversión de nucleótidos), de una manera tal que las modificaciones de ese tipo proporcionan el efecto deseado dentro del microorganismo. Como se usa en el presente documento, las modificaciones genéticas que dan como resultado una disminución de la expresión genética, la función del gen, o la función del producto genético (es decir, la proteína codificada por el gen) se pueden denominar inactivación (completa o parcial), delección, interrupción, bloqueo o regulación por disminución de un gen. Por ejemplo, una modificación genética en un gen que da como resultado una disminución en la función de la proteína codificada por tal gen puede ser el resultado de una delección completa del gen (es decir, el gen no existe, y por tanto la proteína no existe), una mutación en el gen que da como resultado una traducción incompleta de la proteína o que no se traduzca en absoluto (por ejemplo, la proteína no se expresa), o una mutación en el gen que disminuye o suprime la función natural de la proteína (por ejemplo, se expresa una proteína que presenta una disminución de la acción o actividad enzimática o falta de la misma). Las modificaciones genéticas que dan como resultado un aumento en la función o expresión genética se pueden denominar pueden amplificación, sobreproducción, sobreexpresión, activación, potenciación, adición o regulación por incremento de un gen.

De acuerdo con los métodos de la presente invención, un microorganismo de Thraustochytriales recombinante se puede producir usando cualquier microorganismo del orden Thraustochytriales. Los géneros preferentes de Thraustochytriales incluyen, pero no se limitan a: *Thraustochytrium*, *Labyrinthuloides*, *Japonochytrium*, y *Schizochytrium*. Las especies preferentes dentro de estos géneros incluyen, pero no se limitan a: cualquier especie de *Schizochytrium*, incluyendo *Schizochytrium aggregatum*, *Schizochytrium limacinum*; cualquier especie de *Thraustochytrium* (incluyendo cualquier especie de *Ulkenia* anterior tal como *U. visurgensis*, *U. amoeboida*, *U. sarkariana*, *U. profunda*, *U. radiata*, *U. minuta* y de *Ulkenia* sp. BP-5601), *Thraustochytrium striatum*, *Thraustochytrium aureum*, *Thraustochytrium roseum*; y cualquier especie de *Japonochytrium*. Las cepas particularmente preferentes 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 y Belsky) (ATCC 28209); *Schizochytrium limacinum* (Honda y 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).

Los siguientes ejemplos se proporcionan a modo ilustrativo y no pretenden limitar el alcance de los métodos de la presente invención.

Ejemplos

Ejemplo 1

Este ejemplo describe la producción del plásmido recombinante pTUBZE011-2.

La construcción del plásmido recombinante pTUBZE011-2 se ilustran las FIGS. 1 y 2. Este plásmido contiene el gen *ble* de *Streptoalloteichus hindustanus* acoplado de forma funcional a un promotor del gen de α -tubulina aislado de *Schizochytrium* sp. Este plásmido se produjo como sigue a continuación. Un clon de ADNc (CGNE0002-001-B6) se aisló de una biblioteca de ADNc de *Schizochytrium* sp. y se secuenció parcialmente (SEQ ID NO: 1) como parte de un proyecto de secuenciación de ADNc de *Schizochytrium* a gran escala. La secuencia de nucleótidos se determinó para codificar α -tubulina mediante búsqueda de homología BLASTX (Gish, W. y D. States. 1993. Nat. Genet. 3: 266-272). La secuencia de aminoácidos deducir a partir de las bases 116 a 550 es idéntica en un 93 % a los primeros 145 aminoácidos de α -tubulina de *Pelvetica fastigiata* (N.º de Registro en GenBank U58642).

Con el fin de aislar el promotor asociado con este gen, el ADN genómico se aisló de células de *Schizochytrium* sp. y se procesó mediante el uso de un kit "GenomeWalker™" (Clontech Laboratories, Inc., Palo Alto, CA), que implica digestión enzimática de ADN genómico con endonucleasas de restricción para generar extremos romos, seguido por ligación del ADN digerido a moléculas adaptadoras de ADN bicatenario específicas proporcionadas en el kit. El ADN cadena arriba de la secuencia codificante de α -tubulina a continuación se amplificó mediante la reacción en cadena de la polimerasa (PCR), usando el cebador adaptador externo (API) que se proporciona en el kit y el cebador PGR20 específico de α -tubulina (SEQ ID NO: 2). La amplificación adicional del gen se realizó usando el cebador adaptador anidado (AP2) proporcionado en el kit y un cebador PGR19 específico de α -tubulina anidado (SEQ ID NO: 3). Los productos de PCR resultantes se subclonaron en el plásmido pCR2.1-TOPO (Invitrogen Corp., Carlsbad,

CA). Uno de los fragmentos subclonados se secuenció; la secuencia de los 725 pb que preceden inmediatamente al codón de iniciación del gen de α -tubulina se proporcionan como la SEQ ID NO: 4.

Usando los cebadores de oligonucleótido basados en la secuencia de ADN obtenida de esta manera, Se usó PCR usando Taq ADN polimerasa (Perkin-Elmer Corp., Norwalk, CT) para generar una región promotora de α -tubulina modificada en la que un sitio de restricción *NcoI* se incorporan el extremo en la posición 3' terminal del fragmento de ADN; este sitio *NcoI* contenía un codón de inicio que estaba en la misma posición que en la región codificante de α -tubulina. Los cebadores usados en esta reacción fueron PGR33 (SEQ ID NO: 5) si PGR34 (SEQ ID NO: 6), y el molde fue ADN genómico aislado a partir de células *Schizochytrium sp.* Se usaron las siguientes condiciones de reacción: 94 °C durante 4 min; (94 °C durante 1 min, 54 °C durante 45 s, 72 °C durante 2 min) x 30; 72 °C durante 7 min. Este fragmento se clonó en el plásmido pCR2.1-TOPO para formar el plásmido p7TUB (SEQ ID NO: 7). El plásmido p7TUB se digirió con *NcoI*, un fragmento de 463 pb resultante que contenía la región promotora de α -tubulina de *Schizochytrium* se aisló mediante purificación en gel de agarosa. El plásmido pSV40/Zeo (Invitrogen Corp., Carlsbad, CA), que contiene el gen *ble* de *Streptoalloteichus hindustanus* flanqueado por un promotor y terminador de SV40, también se digirió con *NcoI* para producir un fragmento de 3201 pb y un fragmento de 314 pb. El fragmento de 3201 pb se purificó en gel de agarosa y se ligó al fragmento de 463 pb de *NcoI* de p7TUB para producir pTUBZEO-11 (SEQ ID NO: 8), que se representa en la FIG. 1.

A continuación, el plásmido pTUBZEO-11 se difirió con *SphI*, y un fragmento de 1122 pb resultante que contenía el gen *ble* flanqueado por el promotor de α -tubulina de *Schizochytrium* y el terminador de SV40 se purificó en gel de agarosa y se ligó al plásmido pUC19 (Messing, J. 1983. Meth. Enzymol. 101: 20) que se había linealizado mediante digestión con *SphI*. El plásmido resultante se denominó pTUBZEO11-2 (SEQ ID NO: 9) y se representa en las FIGS. 2 y 4. El plásmido pTUBZEO11-2 también se denomina pMON50000. En la SEQ ID NO: 9, el promotor de α -tubulina de *Schizochytrium* está contenido dentro de los nucleótidos 441-894; la región codificante del gen *ble* está contenida dentro de los nucleótidos 895-1269; y el terminador de SV40 está contenido dentro de los nucleótidos 1270-1524.

Ejemplo 2

Este ejemplo describe la producción de los plásmidos recombinantes pMON50200, pMON50201, pMON50202, y pMON50203.

El gen que codifica la acetolactato sintasa nativa (*als*) de *Schizochytrium sp.* se aisló de la siguiente manera. El clon de ADNac (LIB81-028-Q1-E1-D9) se aisló De una biblioteca de ADNc de *Schizochytrium* y se secuenció parcialmente (SEQ ID NO: 11) como parte de un proyecto de secuenciación de ADNc de *Schizochytrium sp.* a gran escala. La secuencia de nucleótidos se determinó mediante homología BLASTX que codifica acetolactato sintasa; por ejemplo, la secuencia de aminoácidos a partir de las bases 154 a 378 era idéntica en un 68 % a los aminoácidos 313 a 387 de ALS de *Schizosaccharomyces pombe* (N.º de Registro en GenBank P36620). La secuencia de longitud completa de este ADNc clonado se obtuvo a continuación, lo que indicó que el clon de ADNc no contenía la región codificante de *als* de longitud completa, una biblioteca lambda genómica de *Schizochytrium* se estudió con sonda usando protocolos convencionales (véase por ejemplo Sambrook *et. al.*, Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989) con una sonda de ADN marcada con digoxigenina (DIG) de 372 pb (denominada sonda ALS2). La sonda ALS2 se genera mediante PCR usando una mezcla de nucleótidos que incluía DIG-11-UTP (Boehringer Mannheim Biochemicals GmbH, Alemania), usando el cebador directo PGR38 (SEQ ID NO: 12) y el cebador inverso PGR39 (SEQ ID NO: 13), que se basaban en la secuencia del clon de ADNc LIB81-028-Q1-E1-D9. Uno de los clones genómicos identificados con la sonda ALS2, denominado ALS-4A, se aisló y caracterizó adicionalmente mediante transferencias de hibridación de tipo Southern usando la sonda ALS2 marcada con DIG. Se encontró que un fragmento de 4,9 kpb de ADN lambda de ALS-4A digerido con *AhdI* hibridaba con la sonda ALS2. Se aisló este fragmento mediante purificación en gel de agarosa, se trató con ADN polimerasa de T4 con el fin de generar extremos romos y entonces se ligó en pBluescriptII KS+ digerido con *SmaI* (Stratagene Corp. La Jolla, CA) para formar el plásmido pMON50200 (representado en la FIG. 3-A). La secuencia de pMON50200 se facilita como la SEQ ID NO: 14. La secuencia de la enzima acetolactato sintasa codificada por el gen *als* de *Schizochytrium* se proporciona como la SEQ ID NO: 15.

Los plásmidos pMON5020 1, pMON50202, y pMON50203 (representados en las FIGS. 3-B, 3-C, y 3-D, respectivamente) se produjeron a partir del plásmido pMON50200 mediante mutagénesis dirigida al sitio de manera que las enzimas acetolactato sintasa codificadas ya no se inhiben más por ciertos compuestos, incluyendo sulfometurón metilo (SMM). Estos plásmidos se construyeron como sigue a continuación. El kit de mutagénesis dirigida al sitio "Transformer" (Clontech Laboratories, Inc., Palo Alto, CA) se usó para introducir las siguientes mutaciones en el plásmido pMON502000 de acuerdo con las instrucciones del fabricante. Un cebador de selección de oligonucleótido, DM19 (SEQ ID NO: 16), se usó en las tres construcciones; este cebador conduce a la conversión de un único sitio EcoRV en el sitio de clonación múltiple de pMON50200 a un sitio *AatII*. El cebador DM14 (SEQ ID NO: 17) se usó para cambiar el resto de aminoácido con el número 595 en la enzima ALS codificada de triptófano a valina, mientras que al mismo tiempo se introducía un sitio *AcI* en la secuencia genética; el plásmido resultante se denomina pMON50201 (SEQ ID NO: 18). De forma análoga, el cebador DM15 (SEQ ID NO: 20) se usó para cambiar el resto de aminoácidos con el número 192 en la enzima ALS codificada a partir de una prolina a una glutamina y para introducir un sitio *BsgI* en el gen *als*, dando como resultado del plásmido pMON50202 (SEQ ID NO: 21). Para

construir el plásmido pMON50203 (SEQ ID NO: 23), se usaron los cebadores tanto DM14 como DM15, dando como resultado una enzima ALS codificada que contenía ambos reemplazamientos de restos de aminoácidos que se han descrito anteriormente. Las secuencias de las enzimas acetolactato sintasa mutantes codificadas por los plásmidos pMON50201, pMON50202, y pMON50203 se proporcionan como las SEQ ID NO: 19, SEQ ID NO: 22, y SEQ ID NO: 24, respectivamente.

Ejemplo 3

Este ejemplo describe la transformación genética de *Schizochytrium sp.* con las moléculas recombinantes que se han descrito en los Ejemplos 1 y 2.

La cepa usada en este ejemplo es N230D de *Schizochytrium sp.*, un derivado de la cepa 20888 de la Colección Americana de Cultivos Tipo (ATCC, Manassas, VA). Para los cultivos líquidos, se hicieron crecer las células axénicamente en medio M50-3 a 30 °C con agitación a 200-300 rpm. El medio M50-3 contiene los siguientes componentes: NaCl, 12,5 g; MgSO₄·7H₂O, 2,5 g; KCl, 0,5 g; CaCl₂, 0,05 g; glucosa, 30 g; Na-glutamato, 3 g; KH₂PO₄, 0,4 g; extracto de levadura, 1 g; NaHCO₃, 0,4 g; Na₂EDTA, 30 mg; FeCl₃·6H₂O, 1,2 mg; H₃BO₃, 34,2 mg; ZnSO₄·7H₂O, 0,67 mg; CoCl₂·6H₂O, 0,13 mg; NaMoO₄·2H₂O, 25 µg; CuSO₄·O5H₂O, 10 µg; NiSO₄·6H₂O, 0,26 mg; tiamina·HCl, 100 µg; biotina, 0,5 µg; cianocobalamina, 0,5 µg, y agua desionizada (hasta 1 l); el pH final se ajustó a 7,0. Para el crecimiento en medio sólido, las células se hicieron crecer a 30 °C en medio M50-3 o medio M1E-3 solidificado mediante la adición de un 1,5 % (p/v) de agar. El medio M1E-3 contiene los siguientes componentes: glucosa, 4 g; (NH₄)₂SO₄, 0,75 g; Na₂SO₄, 5 g; MgSO₄·7H₂O, 2 g; KH₂PO₄, 0,5 g; KCl, 0,5 g; CaCl₂·2 H₂O, 0,1 g; tampón MOPS, 20,9 g; FeSO₄·4H₂O, 0,3 mg; MnCl₂·4H₂O, 0,1 mg; ZnSO₄·7H₂O, 80 µg; CoCl₂·6H₂O, 2 µg; NaMoO₄·2H₂O, 1 µg; CuSO₄·5H₂O, 60 µg; NiSO₄·6H₂O, 80 µg; tiamina·HCl, 320 µg; CA-pantotenato, 320 µg; cianocobalamina, 8 µg, y agua desionizada (hasta 1 l); el pH final se ajustó a 7,0.

La sensibilidad de *Schizochytrium sp.* a Zeocin™ y SMM se determinó incluyendo estos inhibidores en medio M1E-3 solidificado diversas concentraciones y propagando las células en las placas a densidades similares a las que están presentes durante procedimientos utilizados para la selección de células recombinantes.

La transformación genética de células de *Schizochytrium sp.* se realizó mediante bombardeo de partículas (Sanford, J.C., F.D. Smith, y J.A. Russell. 1993. Meth. Enzymol. 217: 483-509) usando un sistema de suministro de partículas Biolistic PDS-1000/He de Bio-Rad (Bio-Rad Laboratories, Hercules, CA). Se hicieron crecer células N230D de *Schizochytrium sp.* en medio líquido M50-3 hasta una densidad óptica a 680 nm (DO₆₈₀) de 0,4-0,8 (longitud de trayectoria óptica de 10 mm). Una alícuota de células correspondientes a 1,0 de DO₆₈₀ se centrifugó brevemente, la solución de sobrenadante se retiró, y las células sedimentadas se volverán a suspender en 100 µl de agua estéril. A continuación las células resuspendidas se propagaron en un círculo de 4 a 6 cm sobre una placa de Petri que contenía medio solidificado con agar (por ejemplo, medio M50-3 o M1E-3) y se permitió su reposo de 30 a 60 min, de modo que el agua en exceso se pudiera absorber en el medio sólido; ésta se denomina placa diana.

Una alícuota de 1,5 mg de microvehículos de oro (diámetro nominal de 0,6 µ, disponible en Bio-Rad Laboratories, Inc., Hercules, CA) se revistió con 2,5 µg de ADN de plásmido de transformación (es decir, plásmido pTUBZEO11-2, pMON50201, pMON50202, o pMON50203) de acuerdo con las instrucciones del fabricante (manual de instrucciones del sistema de suministro de partículas Biolistic® PDS-1000/He; Bio-Rad Laboratories, Hercules, CA). Las células se bombardearon con los microvehículos de oro revestidos con ADN usando las siguientes condiciones: disco de ruptura de 7,6 MPa (1100 psi), vacío de la cámara de 7,6 MPa (25" de Hg), conjunto de lanzamiento de microvehículo colocado en el estante superior y las placas diana colocadas en el estante del medio, proporcionando una distancia desde el disco de ruptura hasta la pantalla de parada de 1,5-2 cm y una distancia desde la pantalla de parada hasta la diana de aproximadamente 7 cm. Después de bombardeo, se dejó que las células se recuperaran en las placas diana durante 4-6 horas a 30 °C. A continuación las células se separaron por enjuagado las células de las placas diana con 1,5 ml de agua estéril, se recogieron en un tubo de microcentrífuga, se centrifugaron brevemente y se resuspendieron en 400 µl de agua estéril. Se propagaron cien microlitros de la suspensión sobre cada una de cuatro placas M1E-3 que contenían o bien 150-200 µg/ml de Zeocin™ (Invitrogen Corp., Carlsbad, CA) o bien 25 µg/ml SMM. Las placas que contenían Zeocin™ se usaron para seleccionar células que se habían transformado con el plásmido pTUBZEO11-2, mientras que las placas que contenían SMM se usaron para seleccionar células que se habían transformado con los plásmidos pMON50201, pMON50202, o pMON50203. A continuación las placas se incubaron durante 7-10 días a 30 °C. Las colonias que parecían ser resistentes al agente selectivo se sembraron en parches sobre placas M1E-3 nuevas que contenían el mismo agente selectivo para confirmar la resistencia. Este protocolo generalmente da como resultado la generación de 100-1000 cepas resistentes a Zeocin™ o resistentes a SMM por bombardeo.

Ejemplo 4

El siguiente ejemplo demuestra el análisis de PCR de células de *Schizochytrium* transformadas.

La PCR se usó para confirmar la presencia de secuencias de plásmido en las cepas supuestamente transformadas que eran resistentes a los agentes selectivos Zeocin™ o SMM. El ADN de molde a partir supuestos transformantes y

células N230D de *Schizochytrium* no recombinantes se obtuvo usando un asa de inoculación de 1 µl de plástico de un solo uso para retirar una pequeña cantidad de células (1-2 mm³) de colonia resistentes que se habían sembrado en parches sobre placas de agar (tal como se describe en el Ejemplo 3). A continuación las células se volvieron a suspender en 15-20 µl de 1 % Triton X-100 en un tubo de microcentrifuga, se colocaron en un baño de agua en ebullición durante 10 minutos, y a continuación se centrifugaron durante 5 minutos a 14,000xg. Se usaron partes de estos extractos (1-3 µl) para proporcionar el ADN de molde para reacciones de PCR de 25 µl usando Taq ADN polimerasa. Para detectar la presencia de secuencias de pTUBZEO11-2 en el ADN de *Schizochytrium*, se usaron los cebadores DM20 (SEQ ID NO: 25) y DM21 (SEQ ID NO: 26); estos cebadores se hibridan con el gen *ble* en el plásmido pTUBZEO11-2 y amplifican un fragmento de ADN de 346 pb. El perfil térmico usado por el que sigue a continuación: 94 °C durante 4 min; (94 °C durante 45 s, 52 °C durante 45 s, 72 °C durante 2 min) x 30; 72 °C durante 7 min. Para detectar la presencia de secuencias de pMON50201, pMON50202, o pMON50203 en el ADN de *Schizochytrium*, se usaron los cebadores BLA1 (SEQ ID NO: 27) y BLA2 (SEQ ID NO: 28); estos cebadores se hibridan con el gen *bla* (resistencia a ampicilina) encontrado en la estructura principal del vector y amplifican un fragmento de ADN de 1229 pb. El perfil técnico usado fue el que sigue a continuación: 94 °C durante 4 min; (94 °C durante 45 s, 55 °C durante 45 s, 72 °C durante 2 min) x 30; 72 °C durante 7 min. Los productos de PCR se analizaron por mediante electroforesis en gel de agarosa convencional, seguido por tinción con bromuro de etidio.

Los resultados de estos análisis confirman que la gran mayoría de las cepas seleccionadas en estas condiciones son transformantes verdaderos que contienen ADN de plásmido. No se generaron productos de PCR del tamaño correcto cuando se usó ADN de molde de células N230D de *Schizochytrium sp.* que no se habían bombardeado con los plásmidos de transformación.

Ejemplo 5

El siguiente ejemplo describe análisis de transferencia de tipo Southern de células de *Schizochytrium* transformadas.

Las transferencias de hibridación de tipo Southern se realizaron usando ADN aislado a partir de células N230D de *Schizochytrium* precursoras y varios supuestos transformantes con el fin de confirmar la presencia de secuencias de ADN de vector de transformación dentro de las células transformadas. La transferencia de tipo Southern se realizó usando técnicas conocidas por los expertos en la materia (véase por ejemplo Sambrook *et al.*, Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press. 1989). El ADN se aisló mediante el uso de un kit de purificación de ADN de "QIAamp" (Qiagen Inc., Valencia, CA), se digirió con diversas enzimas de restricción, se separó mediante electroforesis a través de geles de agarosa (0,8 % - 1,2 % en p/v), y a continuación se transfirió a membranas de nailon mediante transferencia capilar alcalina.

La detección de ADN de vector en células transformadas con pTUBZEO11-2 mediante el uso del sistema basado en DIG "Genius" DIG (Boehringer Mannheim Biochemicals GmbH, Alemania), usando como sonda de hibridación un fragmento de gen *ble* marcado con DIG de 346 pb generado por medio de PCR con los cebadores DM20 (SEQ ID NO: 25) y DM21 (SEQ ID NO: 26) y una mezcla de nucleótidos que incluía DIG-11-UTP. La hibridación previa de la membrana se realizó a 68 °C durante 1 h en el tampón de hibridación suministrado en el kit Genius. La hibridación se realizó a 68 °C durante 18 h en tampón de hibridación que contenía la sonda de gen *ble* que se había desnaturizado térmicamente durante 5 min a 94 °C. A continuación las membranas se lavaron dos veces durante 5 min con 50 ml de 2X SSC/SDS al 0,1 % y dos veces durante 15 min en 50 ml de 0,1X SSC/SDS al 0,1 %. La detección quimioluminiscente del ADN de hibridación se realizó tal como se describe en las instrucciones del kit Genius.

El ADN de células N230D de *Schizochytrium* no transformadas no se hibridaba con la sonda de gen *ble*. Por el contrario, el ADN de células transformadas se hibridaba con la sonda tal como sigue a continuación:

SphI: El ADN digerido con *SphI* de células de *Schizochytrium* transformadas contenía un fragmento de ADN de ~1100 pb que se hibridaba con la sonda de gen *ble*; este fragmento, que también se observa en ADN de pTUBZEO11-2 digerido con *SphI*, representa el casete de expresión del gen *ble* completo (incluyendo el promotor del gen de tubulina y el terminador de SV40).

XhoI: Para cada uno de los transformantes sometidos a ensayo, la digestión con *XhoI* del ADN dio como resultado fragmentos de hibridación con un tamaño superior a 15-20 kpb. *XhoI* no se cortan dentro de pTUBZEO11-2, y por no tanto estos resultados indican que pTUBZEO11-2 no parece existir como un elemento extracromosómico en células transformadas, sino que más bien se integra en el cromosoma de *Schizochytrium*.

NcoI o HindIII: Estas enzimas se cortan ambas una vez dentro de pTUBZEO11-2. La digestión de ADN transformante con cualquiera de estas enzimas normalmente conducía a un fragmento de hibridación prominente que migraba en conjunto con el vector pTUBZEO11-2 linealizado (es decir, ~3,8 kpb). Esto sugiere que el vector se puede integrar en el cromosoma en forma de repeticiones en tándem.

Ejemplo 6

Este ejemplo demuestra la recombinación homóloga en *Schizochytrium*.

Los siguientes experimentos se realizaron para demostrar que la recombinación homóloga se puede producir en *Schizochytrium* entre secuencias de ADN nativas endógenas y secuencias de ADN homólogas presentes en moléculas de ADN recombinante introducidas en las células. Este tipo de recombinación homóloga puede ser muy beneficiosa para producir cepas recombinantes con propiedades deseables. Por ejemplo, la recombinación homóloga se puede usar para inactivar genes endógenos mediante la inserción dirigida de secuencias genéticas extrañas. Adicionalmente, la recombinación homóloga se puede usar para reemplazar un gen endógeno o parte del mismo con una forma alterada del gen de manera que las células recombinantes muestran propiedades nuevas.

Se mostró que la recombinación homóloga se producía en células de *Schizochytrium* transformadas con el plásmido pMON50202, que contiene una mutación en el gen *als* de *Schizochytrium*. Esta mutación introduce un sitio *Bsgl* en la posición de pb 571 de la región codificante de *als*. Hay un sitio *Bsgl* que se produce de manera natural en la posición de pb 1324 de la región codificante de *als*. Por tanto, se pueden usar transferencias de tipo Southern de ADN de *Schizochytrium* digerido con *Bsgl* para discernir el gen *als* nativo del gen *als* mutante recombinante. Para estos experimentos, se produjo una sonda de hibridación específica de *als* por medio de PCR usando una mezcla de nucleótidos que incluía DIG-11-UTP (Boehringer Mannheim Biochemicals GmbH, Alemania), el cebador directo PGR28 (SEC ID n°:32), el cebador inverso PGR30 (SEC ID n°:33) y una pequeña cantidad de pMON50200 como molde. La sonda de hibridación marcada con DIG de 323 pb resultante se denominó ALS1.

El ADN de células N230D de *Schizochytrium* no recombinantes se digirió con *Bsgl* y *AhdI* por separado, se sometió a electroforesis en gel de agarosa, se transfirió a una membrana de nailon y a continuación se estudió con sonda con la sonda ALS1 usando procedimientos esencialmente iguales a los que se han descrito en el Ejemplo 5. La sonda ALS 1 marcó un fragmento de 1,76 kpb de ADN digerido con *Bsgl* y un fragmento de 4,9 kpb de ADN digerido con *AhdI*.

También se estudiaron con sonda las transferencias de tipo Southern de ADN digerido con *Bsgl* y *AhdI* de diversas cepas recombinantes que se habían transformado con pMON50202 con la sonda ALS1. En algunos casos, el fragmento de *Bsgl* de 1,76 kpb no estaba presente, y en su lugar se marcó un fragmento de 0,75 kpb, correspondiente al fragmento de *Bsgl* de 753 pb presente en pMON50202. Se marcó un fragmento de *AhdI* de 4,9 kpb en estas cepas recombinantes, sin embargo, lo que indica que el gen *als* mutante, recombinante se había recombinado con el gen *als* nativo por medio de recombinación homóloga de entrecruzamiento doble.

También se observó que se producía recombinación homóloga de entrecruzamiento único en cepas recombinantes transformadas con pMON50202. En estos casos, se marcaron los fragmentos de *Bsgl* tanto de 1,76 kpb como de 0,75 kpb en las transferencias de tipo Southern de ADN de las cepas recombinantes, pero el fragmento de *AhdI* de 4,9 kpb se reemplazó por fragmentos marcados mayores, lo que indica que todo el vector pMON50202 se había insertado en el gen *als* nativo, o bien como una única copia o bien como repeticiones en tándem.

Las evidencias adicionales de la recombinación homóloga en *Schizochytrium* se obtuvieron la introducción de moléculas de ADN recombinante que contenían un gen *als* truncado, mutante de manera que la enzima ALS incompleta codificada por el gen truncado no era funcional. Este gen truncado se produjo mediante digestión de pMON50202 con *Clal* y *HindIII* para producir un fragmento de 2,8 kpb, eliminando de ese modo los últimos 388 pb de la secuencia codificante de *als* en combinación con la región terminadora de *als*. Este fragmento de 2,8 kpb se ligó en pBluescriptII KS+ (Stratagene Corp., La Jolla, CA) que se había digerido con *Clal* y *HindIII*, produciendo el plásmido pAR2. Se esperaba que el plásmido pAR2 solo confiriere resistencia a SMM en células de *Schizochytrium* transformadas si se restauraba un gen *als* funcional, mutante en cepas transformadas por medio de recombinación homóloga entre el gen *als* nativo y el gen *als* truncado mutante presente en pAR2. Esta construcción se introdujo en células N230D de *Schizochytrium* mediante bombardeo de partículas, y se aislaron cepas resistentes a SMM tal como se describe en el Ejemplo 3. El análisis de transferencia de tipo Southern de ADN digerido con *Bsgl* de los transformantes, llevado a cabo tal como se al escrito anteriormente en este ejemplo, indicó que se había producido claramente recombinación homóloga en estas cepas; es decir, un fragmento de *Bsgl* de 1,76 kpb se hibridó con la sonda ALS 1 en células no recombinantes, pero éste se reemplazó por un fragmento de hibridación de 0,75 kpb en células que se habían transformado con pAR2.

Ejemplo 7

Este ejemplo describe el uso del vector de transformación pTUBZEO11-2 o pMON50202 para producir por medio de cotransformación cepas que contenían moléculas de ADN extraño adicionales que no se unían a un gen marcador seleccionable.

La cotransformación se consiguió mediante la introducción simultánea de pTUBZEO11-2 y un plásmido adicional que contenía cualquiera de varios genes. Los plásmidos se coprecipitaron sobre las partículas de oro tal como se ha descrito en el Ejemplo 3, usando 2,5 µg de cada plásmido. Después del bombardeo de células diana con las partículas de oro recubiertas con plásmidos, las cepas recombinantes se seleccionaron sobre placas de agar que contenían Zeocin™ como se ha descrito en el Ejemplo 1. La presencia del segundo plásmido, no seleccionado se confirmó a continuación mediante análisis de PCR o mediante hibridación de transferencia de tipo Southern. Normalmente se lograban unas frecuencias de cotransformación muy elevadas (por ejemplo, 50-90 %). Por ejemplo,

el plásmido pTR202, que contiene el gen *fat-1* de *Caenorhabditis elegans* (Spychalla *et al.*, 1997. Proc. Natl. Acad. Set U. S. A. 94, 1142-1147) unido al promotor y terminador del gen de tubulina de *Schizochytrium*, se introdujo se introdujo por medio del método que se proporciona en este ejemplo, y mediante PCR se mostró que aproximadamente un 68 % de las cepas resistentes a Zeocin™ resultantes contenían el gen *fat-1* (Véase la Tabla 1). Se observan resultados similares cuando se introdujeron conjuntamente pMON50202 y un plásmido adicional, seguido por selección de células transformadas en medio sólido que contenía SMM. Este método de cotransformación se puede usar para introducir cualquier ADN extraño deseado.

Tabla 1. Eficacias de cotransformación usando el plásmido marcador seleccionable pTubZeO11-2 y plásmidos que contienen diversos genes *fad*. Los transformantes resistentes a Zeocin^R se identificaron sistemáticamente para secuencias de ADN de *fad* mediante PCR.

<u>Gen <i>fad</i> introducido</u>	<u>N.º de cepas que contienen gen <i>fad</i> gene N.º de cepas de Zeocin^R sometidas a ensayo</u>	<u>Eficacia de cotransformación</u>
syn_fat1	17/25	68 %
nat_fat1	24/25	96 %
mut_fat1	21/25	84 %
desB	20/25	80 %

Los sistemas de transformación y se describen en estos ejemplos representan un avance significativo en la capacidad para manipular genéticamente *Schizochytrium*, que es el organismo más productivo conocido para la producción fermentativa de compuestos a base de lípidos. La disponibilidad de dos sistemas de transformación independientes, junto con las altas eficacias de transformación que se producen, deben permitir el apilamiento de múltiples rasgos en cepas modificadas por ingeniería genética. Además, la presencia aparente de recombinación homóloga en esta microalga debe permitir el desarrollo de procedimientos de desactivación génica con el fin de identificar las funciones de genes desconocidos y eliminar rasgos no deseables en cepas de producción. Los presentes inventores están usando en la actualidad estos sistemas para alterar el metabolismo de ácidos grasos en *Schizochytrium*, y están explorando posibilidades para usa esta especie y microalgas relacionadas (por ejemplo, *Thraustochytrium*) para la producción de carotenoides, esteroides y otros compuestos lipídicos.

Aunque se han descrito diversas realizaciones de la presente invención en detalle, para los expertos en la materia es evidente que se producirán modificaciones y adaptaciones de dichas realizaciones. De forma expresa se debe entender, sin embargo, que tales modificaciones y adaptaciones están dentro del alcance de la presente invención, tal como se expone en las reivindicaciones que siguen a continuación.

LISTADO DE SECUENCIAS

<110> Roessler, Paul
 Matthews, T. Dave
 Ramseier, Tom
 Metz, James

<120> Producto y Proceso para Transformación de Microorganismos de Thraustochytriales

<130> 2997-23-PCT

<150> 60/284.116

<151> 16-04-2001

<160> 35

<170> PatentIn versión 3.1

<210> 1

<211> 551

<212> ADN

<213> *Schizochytrium* sp.

<220>

<221> misc_feature

<222> (520)..(520)

<223> n = a, c, g o t

ES 2 688 953 T3

<220>
 <221> misc_feature
 <222> (541)..(541)
 <223> n = a, c, g o t

5

<400> 1

gtcgtgccta	acaacaacgcc	gttctacccc	gcctttcttcg	cgccccttcg	cgccaagca	60
tccttcaagt	ttatctctct	agttcaactt	caagaagaac	aacaccacca	acaagatgog	120
tgaggtcato	tocatccaca	tgggcaggc	cggtgttcag	gtcggtaacg	cctgctggga	180
gctctactgc	ctcgagcatg	gcattccagc	ggacggccag	atgccctcgg	acaagacat	240
tggcggcggc	gatgatgcct	tcaacacctt	cttctccgag	actggcgccg	gcaagcacgt	300
gccccgcgcc	gtgctcgtcg	atctcgagcc	caccgtctgt	gacgaggctc	gcaccggcac	360
ctaccgcgct	ctttaccacc	ccgagcagat	catcacgggc	aaggaggacg	ctgccaacaa	420
ctacgctcgt	ggccaactaca	ccatcgcaaa	ggagatcgtc	gacctcgtcc	tcgaccgcat	480
ccgcaagctc	gccgacaact	gcactggctt	tcagggcttn	ctctgttca	acgccgtcgg	540
nggtgggtacc	g					551

10

<210> 2
 <211> 27
 <212> ADN
 <213> Schizochytrium sp.

15

<400> 2
 gcgccagtct cggagaagaa ggtgttg 27

20

<210> 3
 <211> 27
 <212> ADN
 <213> Schizochytrium sp.

25

<400> 3
 agtcccagc aggcgttacc gacctga 27

30

<210> 4
 <211> 725
 <212> ADN
 <213> Schizochytrium sp.

<400> 4

gagacgigct	tcgcaagacc	gctgtgctcg	cgccgcacgc	tctgtgtggt	acattaattt	60
ttttgtagat	gaagtttctc	tattctctcg	aaattctgta	gaatgttata	gtctcttcac	120
tcccgtgatt	ggagaggatt	cttgcttgtt	ccctcccgcc	cggttagcgc	ttggagcaac	180
gcttgagcgc	gcgctcgaaa	gcggacggcg	caacgagccg	tttcacgcgc	cgctgtccaa	240
gtcccatttt	tctccttacc	ccatggccgt	tgcatgccaa	ttttaggccc	ccactgacc	300
gaggtctgtc	gataatccac	ttttccattg	atcttccagg	tttctgtaac	tcattgccact	360
gagcaaaact	tcggtctttc	ctaacaaaag	ctctcctcac	aaagcatggc	gcggcaacgg	420
acgtgtcctc	atactccact	gccacacaag	gtcgataaac	taagctcctc	acaatatagag	480
gagaattcca	ctgacaaactg	aaaacaatgt	atgagagacg	atcaccactg	gagcggcgcg	540
gcggttgggc	gcggaggctcg	gcagcaaaaa	caagcgactc	gcccagcaaa	cccgaatcag	600
ccttcagacg	gtcgtgccta	acaacaacgcc	gttctacccc	gccttcttcg	cgccccttcg	660
cgccaagca	tccttcaagt	ttatctctct	agttcaactt	caagaagaac	aacaccacca	720
acaag						725

35

<210> 5
 <211> 24
 <212> ADN
 <213> Schizochytrium sp.

40

<400> 5
 caccatggt gttggtgtg ttgt 24

ES 2 688 953 T3

<210> 6
 <211> 24
 <212> ADN
 <213> Schizochytrium sp.

5

<400> 6
 aaactcgaga cgtgcttcgc aaga 24

<210> 7
 <211> 4646
 <212> ADN
 <213> Schizochytrium sp.

10

<400> 7

15

aaactcgaga	cgtgcttcgc	aagaccgctg	tgctcgcgcc	gcaogctctg	tgtgttacat	60
taattttttt	gtagatgaag	tttctctatt	ctctcgaaat	tctgtagaat	gttatagtct	120
cttcaactccc	gtgattggag	aggattcttg	cttgttccct	cccgcccggg	tagcgttgg	180
agcaacgctt	gagcgcgcgc	tcgaaagcgg	acggcgcaac	gagccgtttc	acgccgcgct	240
gtccaagtcc	catttttctc	cttaccat	ggccgttgoa	tgccaatttt	aggcccccca	300
ctgaccgagg	tctgtcgata	atccactttt	ccattgatct	tccaggtttc	gttaactcat	360
gccactgagc	aaaacttcgg	tctttcctaa	caaaagctct	cctcacaag	catggcgcg	420
caacggacgt	gtcctcatac	tccactgcc	cacaaggctg	ataaactaag	ctcctcaca	480
atagaggaga	attccactga	caactgaaa	caatgtatga	gagaogatca	ccactggagc	540
ggcgcggcgg	ttgggcgcgg	aggtcggcag	caaaaacaag	cgactcgccg	agcaaacccg	600
aatcagcctt	cagacggctg	tgccatacaa	caogccgttc	taccccgcct	tcttcgcgcc	660
ccttcgcgtc	caagcatcct	tcaagtttat	ctctctagtt	caacttcaag	agaacaaca	720
ccaccaacac	catgggtgaa	gggcgaatc	tcagatatac	catcacactg	gcggccgctc	780
gagcatgcac	ctagagggcc	caattcgccc	tatagtgagt	cgtattacaa	ttactggcc	840
gtcgttttac	aacgtcgtga	ctgggaaaac	cctggcgtta	cccaacttaa	tcgccttgca	900
gcacatcccc	ctttcgccag	ctggcgtaat	agcgaagagg	cccgcaccga	tcgccttcc	960
caacagttgc	gcagcctgaa	tggcgaatgg	gacgcgcctt	gtagcggcgc	attaagcgcg	1020

ES 2 688 953 T3

gcgggtgtgg	tggttacgcy	cagcgtgacc	gctacacttg	ccagcgcctt	agcgcgcgct	1080
cctttcgcct	tcttcccttc	ccttctcgcc	acgttcgcgc	gctttcccog	toaagctcta	1140
aatcgggggc	tccttttagg	gttccgattt	agagctttac	ggcacctcga	ccgcaaaaaa	1200
ccttgatttg	gtgatgggtc	acgtagtggg	ccatcgccct	gatagacggt	ttttcgcctt	1260
ttgacgttgg	agtccacggt	ccttaaatagt	ggactcctgt	tcoaaactgg	aacaacactc	1320
aaccctatcg	cggtctattc	ttttgattta	taagggattt	tgccgatttc	ggcctattgg	1380
ttaaaaaatg	agctgattta	acaaattcag	ggcgcaaggg	ctgctaaggg	aaccggaaca	1440
cgtagaaaagc	cagtccgcag	aaacgggtct	gaccccggat	gaatgtcagc	tactgggcta	1500
tctggacaag	ggaaaacgca	agcgcgaaaga	gaaagcaggt	agcttgacgt	gggcttacat	1560
ggcgatagct	agactggggc	gttttatgga	cagcaagcga	accggaattg	ccagctgggg	1620
cgccctctgg	taaggttggg	aagccctgca	aagtaaactg	gatggctttc	ttgcccgcaa	1680
ggatctgatg	gcgcagggga	tcaagatctg	atcaagagac	aggatgagga	tcgtttcgca	1740
tgattgaaca	agatggattg	caagcaggtt	ctccggccgc	ttgggtggag	aggctattcg	1800
gctatgactg	ggcacaacag	acaatcggct	gctctgatgc	cgccgtgttc	cgccgtcag	1860
cgacggggcg	cccggttctt	tttgtcaaga	ccgacctgtc	cggtgccctg	aatgaaactgc	1920
aggacgaggg	agcgcgggct	tcgtggctgg	ccacgacggg	cgttccttgc	gcagctgtgc	1980
tcgacgttgt	cactgaagcg	ggaagggact	ggctgctatt	gggcgaagtg	ccggggcagg	2040
atctcctgtc	atctcgcctt	gctcctgccc	agaaagtatc	catcatggct	gatgcaatgc	2100
ggcggctgca	tacgcttgat	ccggctacct	gcccattcga	ccaccaagcg	aaacatcgca	2160
tcgagcagag	acgtactcgg	atggaagccg	gtcttctcga	tcaggatgat	ctggacgaag	2220
agcatcaggg	gtcgcggcca	gccgaactgt	tcgccaggct	caaggcgcgc	atgcccgcag	2280
gcgaggatct	cgtcgtgatc	catggcgatg	cctgcttgcc	gaatatcatg	gtggaaaatg	2340
gccccttttc	tggattcaac	gactgtggcc	ggctgggtgt	ggcggaccgc	tatcaggaca	2400
tagcgttggg	taccctgtgat	attgctgaag	agcttggcgg	cgaaatgggt	gaccgcttcc	2460
tcgtgcttta	cggtatcgcc	gctcccagatt	cgacgcgcct	cgcttcttat	cgcttctttg	2520
acgagttctt	ctgaattgaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtgcgcct	2580
tattcccttt	tttgcggcat	tttgccttcc	tgtttttgct	cacccagaaa	cgctgtgtaa	2640
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	2700
cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	tgagcacttt	2760
taaagttctg	ctatgtcata	cactattatc	ccgtattgac	gcccggcaag	agcaactcgg	2820
tcgcccggcg	cggtattctc	agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	2880
tcttacggat	ggcatgacag	taagagaatt	atgcatgctt	gccataacca	tgagtgataa	2940
cactcggccc	aaacttactt	tgacaacgat	cgaggaccg	aaggagctaa	ccgctttttt	3000
gcacaacatg	ggggatcatg	taactcgcct	tgatcgttgg	gaaccggagc	tgaatgaagc	3060
cataccaaac	gacgagagtg	acaccacgat	gcctgtagca	atgccaacaa	cgttgcgcaa	3120
actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	actggatgga	3180
ggcggataaa	gttgcaggac	cacttctcgg	ctcggccctt	ccggetggct	ggtttattgc	3240
tgataaatct	ggagccgggt	agcgtgggtc	tcgcggatc	attgcagcac	tggggccaga	3300
tcggtaagcc	tcccgatcgc	tagttatcta	cacgacgggg	agtcaggcaa	ctatggatga	3360
acgaaataga	cagatcgctg	agataggtgc	ctcactgatt	aagcattggt	aactgtcaga	3420
ccaagtttac	tcatatatac	tttagattga	tttaaaactt	catttttaat	ttaaaaggat	3480
ctaggtgaag	atcctttttg	ataatctcat	gaccaaaatc	ccttaacgtg	agttttcggt	3540
ccactgagcg	tcagaccccc	tagaaaagat	caaaggatct	tcttgagatc	ccttttttct	3600
gcgcgtaatc	tgtctgcttc	aaacaaaaaa	accaccgcta	ccagcgggtg	tttgtttccc	3660
ggatcaagag	ctaccaactc	tttttcgaaa	ggttaactgg	ttcagcagag	cgcatatacc	3720
aaatactgtc	cttctagttt	agccgtagtt	agggcaccac	ttcaagaact	ctgtagcacc	3780
gcctacatac	ctcgcctcgc	taatcctggt	accagtggct	gctgccagtg	gcgataagtc	3840
gtgtcttacc	gggttggact	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	3900
aacggggggg	tcgtgcacac	agcccagcct	ggagcgaacg	acctacaccg	aactgagata	3960
cctacagcgt	gagcattgag	aaagcgcacc	gcttcccga	gggagaaagg	cggacaggta	4020
tcgggtaagc	ggcagggctg	gaacaggaga	gcgcacgagg	gagcttccag	ggggaacgcg	4080
ctggtatctt	tatagtcctg	tcgggtttcg	ccacctctga	cttgagcgtc	gatttttggt	4140
atgctcgtca	ggggggcgga	gcctatggaa	aaacgcgagc	aaocgggcct	ttttacggtt	4200
cctggcctti	tgtctgcttc	ttgctcacat	gttctttcct	gogttatccc	ctgattctgt	4260
ggataaccgt	attaccgcct	ttgagtgagc	tgataccgct	cgccgcagcc	gaacgaccga	4320
gcgcagcag	tcagtgagcg	aggaagcgga	agagcgccca	atacgcgaaac	cgctctccc	4380
cgcgcttgg	ccgattcatt	aatgcagctg	gcacgacagg	tttcccagct	ggaagcggg	4440
cagtgagcgc	aacgcaatta	atgtgagtta	gctcactcat	taggcacccc	aggctttaca	4500
ccttatgctt	ccggctcgta	tgttgtgtgg	aatgtgagc	ggataacaat	ttcacacagg	4560

aaacagctat	gaccatgatt	acgccaagct	tggtaccgag	ctcggatcca	ctagtaacgg	4620
ccgocaggtg	gctggaattc	gccctt				4646

ES 2 688 953 T3

<210> 8
 <211> 3664
 <212> ADN
 <213> Schizochytrium sp.

5

<400> 8

ccatggcogt	tgcatgccaa	ttttaggccc	cccactgacc	gaggtctgtc	gataatccac	60
ttttocattg	atottccagg	tttcgtaaac	tcatgccact	gagcaaaact	toggteittc	120
ctaacaaaag	ctctcctcac	aaagcatggc	gcggcaacgg	acgtgtcctc	atactccact	180
gccacacaag	gtcgataaac	taagctcctc	acaaatagag	gagaattcca	ctgacaactg	240
aaaacaatgt	atgagagacg	atcaccactg	gagcggcgcg	gcggttgggc	gcgagggtcg	300
gcagcaaaaa	caagcgactc	gocgagcaaa	cccgaatcag	ccttcagacg	gtcgtgccta	360
acaacacgcc	gttctacccc	gccttcttcg	cgccccctcg	cgtccaagca	tccttcaagt	420
ttatctctct	agttcaactt	caagaagaac	aacaccacca	acaccatggc	caagttgacc	480
agtgcogttc	cgggtgctcac	cgcgcgcgac	gtcgcgggag	cggtcgagtt	ctggaccgac	540
cggctcgggt	tctcccggga	cttcgtggag	gaogacttcg	cgggtgtggt	cgggacgac	600
gtgaccctgt	tcatcagcgc	ggtccaggac	caggtggtgc	cggacaacac	cctggcctgg	660
gtgtgggtgc	gogggcotgga	cgagctgtac	gocgagtggt	eggaggtcgt	gtccacgaac	720
ttccgggacg	cctccggggc	ggccatgacc	gagatcggcg	agcagccgtg	ggggcgggag	780
ttcgccctgc	gogacccggc	cggcaactgc	gtgcaacttcg	tggccgagga	gcaggactga	840
cacgtgctac	gagatttoga	ttccaccgcc	gccttctatg	aaaggttggg	cttcggaatc	900
gtttccoggg	acgcggcgtg	gatgatcctc	cagcgcgggg	atctcatgct	ggagttcttc	960
gcccacoccca	acttgtttat	tgcagcttat	aatggttaca	aataaagcaa	tagcatcaca	1020
aatttcacaa	ataaagcatt	tttttcaact	cattctagtt	gtggtttgtc	caaactcatc	1080
aatgtatcct	atcatgtctg	aattcccggg	gatcctctag	agtcgacctg	caggcatgca	1140
agcttggcac	tggccgtcgt	tttacaacgt	cgtgactggg	aaaaccctgg	cgttaccocaa	1200
cttaatcgcc	ttgcagcaca	tcccccttc	gccagctggc	gtaatagcga	agagcccgcg	1260
accgatcgcc	cttcccaca	gttgcgcagc	ctgaatggcg	aatggcgcct	gatggcgtat	1320
tttctcctta	cgcatctgtg	cggtatttca	caccgcata	atgggtgact	ctcagtacaa	1380
tctgctctga	tgccgcatag	ttaagccagc	cccgcacccc	gccaacaccc	gctgacgcgc	1440
cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca	agctgtgacc	gtctccggga	1500
gctgcatgtg	tcagaggttt	tcaccgctcat	caccgaaaacg	cgcgagacga	aagggcctcg	1560
tgatacgcct	atTTTTatag	gTTaatgtca	tgataataat	ggTTTTcttag	acgtcaggtg	1620
gcacttttcg	gggaaatgtg	cgcggaaccc	ctatttggtt	atTTTTctaa	atacattcaa	1680
atatgtatcc	gctcatgaga	caataaccct	gataaatgct	tcaataatat	tgaaaaagga	1740
agagtatgag	tattcaacat	ttcogtctcg	ccottattcc	ctTTTTtgcg	gcatttttgc	1800
ttcotgtttt	tgctcaacca	gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagttgg	1860
gtgcaacgag	gggttacatc	gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	1920
gccccgaaga	acgtttttcca	atgatgagca	cttttaaagt	tctgctatgt	ggcgcggtat	1980
tatcccgat	tgacgcgggg	caagagcaac	toggctgcgc	catacactat	tctcagaatg	2040
acttggttga	gtactcacca	gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	2100
aattatgcag	tgctgccata	accatgagtg	ataaacctgc	ggccaactta	cttctgacaa	2160
cgatoggagg	accgaaggag	ctaaccgctt	ttttgcacaa	catgggggat	catgtaatc	2220
goccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	2280
cgatgcctgt	agcaatggca	acaacgttgc	gcaaacattt	aaactggcgaa	ctacttactc	2340
tagcttcccg	gcaacaatta	atagactgga	tggaggcgga	taaagttgca	ggaccacttc	2400
tggcctcggc	ccttccggct	ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	2460
ggtctcggcg	tatcattgca	gcactggggc	cagatggtaa	gccctcccg	atcgtagtta	2520
tctacacgac	ggggagtcag	gcaactatgg	atgaaacgaa	tagacagatc	gctgagatag	2580
gtgcctcact	gattaagcat	tggtaactgt	cagaccaagt	ttactcatat	atactttaga	2640
ttgatttaaa	acttcatttt	taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	2700
tcatgaccaa	aatcccttaa	cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	2760
agatcaaaag	atcttcttga	gatccttttt	ttctgocgct	aatctgctgc	ttgcaaacaa	2820
aaaaaccacc	gctaccagcg	gtggtttgtt	tgcoggatca	agagctacca	actctttttc	2880
cgaaggtaac	tggcttcagc	agagcgcaga	taccaaatac	tgtccttcta	gtgtagccgt	2940
agttaggcca	ccacttcaag	aactctgtag	caccgcctac	atacctcgtc	ctgctaatac	3000

ES 2 688 953 T3

```

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct tacccgggttg gactcaagac 3060
gatagttacc ggataaggcg cagcggctcg gctgaaocgg gggttcgtgc acacagccca 3120
gcttgagcgc aacgacctac accgaaactga gatacctaca gcgtgagcat tgagaaagcg 3180
ccacgcttcc cgaaggggaga aaggcggaca ggtatccggg aagcggcagg gtcggaacag 3240
gagagcgcac gagggagctt ccagggggaa acgootggta tctttatagt cctgtcgggt 3300
ttcgccaact ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat 3360
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc 3420
acatgtgtgc tgggccacgc cggccagatc tgagctcgcg gccgcgatat cgtagctcg 3480
aggcaggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt 3540
ccccaggctc ccacagcagg agaagtatgc aaagcatgca tctcaattag tcagcaacca 3600
tagtcccgcg cctaactccg cccatcccgc cctaactcc gccagttcc gccattctc 3660
cgcc

```

<210> 9

<211> 3808

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Schizochytrium sp. y Streptoalloteichus hindustanus

<220>

<221> Promotor

<222> (441)..(894)

<223>

15 <220>

<221> CDS

<222> (895)..(1269)

<223>

20 <220>

<221> terminador

<222> (1270)..(1524)

<223>

25 <400> 9

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagaagggtca 60
cagcttgtct gtaagcggat gccggggagca gacaagcccg tcagggcgcg tcagggggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataaccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aaggcgcgac ggtgcggggc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgcagg ccagtgaatt cgagctcggg acccggggat 420
cctctagagt cgacctgcag gcatgccaat tttaggcccc ccaactgacc aggtctgtcg 480
ataatccaact tttccattga ttttccaggt ttogttaact catgccactg agcaaaaactt 540
cggctcttcc taacaaaagc tctcctcaca aagcatggcg cggcaacgga cgtgtcctca 600
tactccactg ccacacaagg tcgataaact aagctcctca caaatagagg agaattccac 660
tgacaactga aaacaatgta tgagagacga tcaccactgg agcggcgcgg cggttgggcg 720
cggaggctcg cagcaaaaac aagcgcactcg ccgagcaaac ccgaatcagc cttcagacgg 780
tcgtgcctaa caacacgcg ttctaccccg ctttcttcgc gcccttcgc gtccaagcat 840
ccttcaagtt tatctctcta gttcaacttc aagaagaaca acaccacca cacc atg 897

```

Met

1

```

gcc aag ttg acc agt gcc gtt ccg gtg ctc acc gcg cgc gac gtc gcc 945
Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val Ala

```

5

10

15

```

gga gcg gtc gag ttc tgg acc gac cgg ctc ggg ttc tcc cgg gac ttc 993
Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp Phe

```

20

25

30

ES 2 688 953 T3

gtg gag gac gac ttc gcc ggt gtg gtc cgg gac gac gtg acc ctg ttc	1041
Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu Phe	
35 40 45	
atc agc gog gtc cag gac cag gtg gtg ccg gac aac acc ctg gcc tgg	1089
Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala Trp	
50 55 60 65	
gtg tgg gtg cgc ggc ctg gac gag ctg tac gcc gag tgg tcg gag gtc	1137
Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu Val	
70 75 80	
gtg tcc acg aac ttc cgg gac gcc tcc ggg ccg gcc atg acc gag atc	1185
Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu Ile	
85 90 95	
ggc gag cag ccg tgg ggg cgg gag ttc gcc ctg cgc gac ccg gcc ggc	1233
Gly Glu Gln Pro Trp Gly Arg Glu Glu Phe Ala Leu Arg Asp Pro Ala Gly	
100 105 110	
aac tgc gtg cac ttc gtg gcc gag gag cag gac tga cacgtgctac	1279
Asn Cys Val His Phe Val Ala Glu Glu Gln Asp	
115 120	
gagatttoga ttccaaccgc gccttctatg aaaggttggg cttoggaatc gttttccggg	1339
acgccgctg gatgatctc cagcgcgggg atctcatgct ggagttcttc gcccaacca	1399
actgtttat tgcagcttat aatggttaca aataaagcaa tagcatcaca aatttcacaa	1459
ataaagcatt tttttcactg cattctagtt gtggtttgtc caaactcatc aatgtatctt	1519
atcatgtctg aattcccggg gatcctctag agtcgacctg caggcatgca agcttggcgt	1579
aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt ccacacaaca	1639
taogagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc taactcacat	1699
taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt	1759
aatgaatcgg ccaacgogcg gggagagggc gtttgcgtat tgggcgctct tccgcttct	1819
cgtcactga ctgcctgcgc tccgtcgttc ggctgcggcg agcggtatca gctcactcaa	1879
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa	1939
aaggccagca aaaggccagg aaccgtaaaa aggcgcgctt gctggcgttt ttccataggc	1999
tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaccoga	2059
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc	2119
cgaccctgcc gcttacggga tacctgtccg cctttctccc ttccgggaagc gtggcggtt	2179
ctcaatgctc acgctgtagg tatctcagtt cgggtgtaggt cgttcgtcc aagctgggct	2239
gtgtgcaoga acccccggtt cagcccgacc gctgcgectt atccggtaac tatcgtcttg	2299
agtccaacc ccgtaagacac gacttatcgc cactggcagc agccactggt aacaggatta	2359
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct	2419
acactagaag gacagtatctt ggtatctgog ctctgctgaa gccagttacc ttccgaaaaa	2479
gagttggtag ctcttgatcc ggcaaaaaaa ccaccgctgg tagcgggtgg tttttgttt	2539
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta	2599
cggggtctga cgctcagtgg aacgaaaact cacgtaagg gattttggtc atgagattat	2659
caaaaaggat ctccacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa	2719
gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct	2779
cagcgatctg tctatttctg tcatccatag ttgcctgaet ccccgctcgt tagataacta	2839
cgatacggga gggcttacca tctggcccc gtgctgcaat gataccgga gccccagct	2899
caccgctcc cagattatca gcaataaacc agccagcgg aagggccgag cgcagagtg	2959
gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgcccggaa gctagagtaa	3019
gtagttcgcc agttaatagt ttgcgcaacg ttggttgcct tgctacaggc atcgtggtgt	3079
cacgctcgtc gtttgggtat gcttcatcca gctccggttc ccaacgatca aggcgagtta	3139
catgatcccc catgttctgc aaaaaagcgg ttagctcctt cggctcctcc atcgttctca	3199
gaagtaagtt ggcgcagtg ttatcactca ttggtatggc agcactgcat aattctctta	3259
ctgtcatgcc atccgtaaga tgctttctg tgactcgtga gtaactcaacc aagtcattct	3319
gagaatagtg tatgcggcga ccgagttgct ctggcccgc gtcaatacgg gataatacgg	3379
cgccacatag cagaacttta aaagtgtcca tcatgggaaa acgttctctg gggcgaaaac	3439
tctcaaggat ctaccgctg ttgagatcca gttcgatgta acccaactcgt gcacccaact	3499
gatcttcagc atotttfact ttaccagcg tttctgggtg agcaaaaaa ggaaggcaaa	3559
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttcttt	3619
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat	3679
gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg	3739
acgtotaaga aaccattatt atcatgacat taacctataa aatagggct atcacgaggc	3799
cctttcgtc	3808

ES 2 688 953 T3

<210> 10
 <211> 124
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Schizochytrium sp. y Streptoalloteichus hindustanus

<400> 10

10

```

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

<210> 11
 <211> 1416
 <212> ADN
 <213> Schizochytrium sp.

15

<400> 11

```

gcaaaagggtc gagcttttcc acaaggagcg cattggcgct cctggcacgg ccgacttcaa      60
gtcattgccc gagatgatca accgtgogga gcgaccogtc atctatgctg gccagggtgt      120
catgcagagc ccggttgaatg gcccgctgt gctcaaggag ttccgaggaga agccaacat      180
tccogtgacc accaccatgc agggctctgg cggctttgac gagcgtagtc coctctccct      240
caagatgctc ggcatgcacg gctctgccta cgccaactac tcgatgcaga acgccgatct      300
tacctggcg ctcgggtgccc gctttgatga tcgtgtgacg ggccgcggtg acgcctttgc      360
tccggaggct ccgocgtgccg agcgcgaggg ccgocggtggc atcgttcact ttgagatttc      420
ccccagaac ctccacaagg tcgtccagcc caccgtcgcg gtccctcggcg acgtggtcga      480
gaacctcgcc aacgtcacgc cccacgtgca gcgccaggag cgcgagccgt ggtttgogca      540
gatcgccgat tgggaaggaga agcacccttt tctgctcgag tctggtgatt cggacgacaa      600
ggttotcaag ccgcagcagg tctcaogga gcttaacaag cagattctcg agattcagga      660
gaaggacgcc gaccaggagg tctacatcac cacgggogtc ggaagccacc agatgcaggc      720
agcgcagttc cttacctgga ccaagcogcg ccagtggtatc tctcgggtg ggcocggocac      780
tatgggetac ggccttccct cggccattgg cgccaagatt gcgaagcccg atgctattgt      840
tattgacatc gatgggtgatg cttcttattc gatgaccggt atggaattga tcacagcagc      900
cgaattcaag gttggcgtga agattcttct tttgcagaac aactttcagg gcatggtcaa      960
gaactggcag gatctctttt acgacaagcg ctactcgggc accgccatgt tcaaccocgg      1020
cttcgacaag gtcgcccgatg cgatgcgtgc caagggtctc tactgcgcga aacagtcgga      1080
    
```

20

ES 2 688 953 T3

```

gctcaaggac aagatcaagg agtttctcga gtacgatgag ggtcccgtcc tcctcgaggt 1140
tttctgtggac aaggacacgc tcgtcttgcc catggteccc gctggctttc cgctccacga 1200
gatggtcctc gagcctccta agcccaagga cgcctaagtt cttttttcca tggcgggoga 1260
gcgagcgagc gcgcgagcgc gcaagtgcgc aagcgccttg ccttgctttg cttcgcttgc 1320
ctttgctttg cttcacacaa cctaagtatg aattcaagtt ttcttgcttg tcggcgaaaa 1380
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1416

```

<210> 12

<211> 20

<212> ADN

5 <213> Schizochytrium sp.

<400> 12

ggatctcttt tacgacaagc 20

10 <210> 13

<211> 18

<212> ADN

<213> Schizochytrium sp.

15 <400> 13

ggtgtgtga agcaaagc 18

<210> 14

<211> 7847

20 <212> ADN

<213> Schizochytrium sp.

<220>

<221> ceb_transcripción

25 <222> (1)..(1259)

<223>

<220>

<221> CDS

30 <222> (1260)..(3314)

<223>

<220>

<221> ceb_transcripción

35 <222> (3315)..(4887)

<223>

<400> 14

```

ttgtcgacag caacttgcaa gttatacgcg accaccaggc aatctcagca cgcccagcga 60
gcacggagct tgogaagagg gtttacaagt cgtcgttcat tcgctctcaa gctttgcctc 120
aacgcaacta ggcccaggcc tactttcaet gtgtcttgtc ttgcctttca caccgaccga 180
gtgtgcacaa ccgtgttttg cacaaagcgc aagatgctca ctcgactgtg aagaaagggt 240
gcgcgcaagc gactgcgact gogaggatga ggatgactgg cagcctgttc aaaaactgaa 300
aatccgcgat gggtcagtgc cattcgcgca tgaagcctgc gagagacaag ttaactcgtg 360
tcaactggcat gtccctagcat ctttacgcga gaaaattca atcgctttat tttttcagtt 420
tcgtaacott ctgcgaaccg cgaatcgccg ttccagcctg actaatctgc agctgcgtgg 480
cactgtcagt cagtcagtca gtctgtgcgc ctgttccagc accgaggctc gcgctgcgcc 540
cccttgacc gctgtgcta ctgctagtgg cacggcaggg aggagcttgt tgccggaaca 600
ccagcagccg ccagtcgacg ccagccaggg gaaagtccgg cgtcgaaggg agaggaagyc 660
ggcgtgtgca aactaacgtt gaccactggc gcccgccgac acgagcagga agcaggaagc 720
tgcaagcgc agcgcgcaag tgcagaatgc gcgaaagatc cacttgccgc cggcgggggc 780
gcacttgcyg gcgcggcgcg gaacagtgcg gaaagggagc gtgcagacgg cgcgcagtga 840
cagtgggcgc aaagccgcgc agtaagcagc ggcggggaac ggtatacgca gtgccgcggg 900

```

40

ccgcccaca cagaagtata cgcgggccga agtggggcgt cgcgcgcggg aagtgcggaa 960
 tggcgggcaa ggaaaggagg agacggaaaag agggcgggaa agagagagag agagagtgaa 1020
 aaaagaaaga aagaaagaaa gaaagaaaga aagotcggag ccacgccgcg gggagagaga 1080
 gaaatgaaag cacggcacgg caaagcaaag caaagcagac ccagccagac ccagccgagg 1140
 gaggagcgcg cgcaggaccc gcgcggcgag cgagcgcgca cggcgcgca gcgagcgcg 1200
 gagcgcgcg gcgagcgcg aaggcttgct gcgagcgcg gagcgcgca gcgggaagg 1259
 atg agc gcg acc cgc gcg gcg acg agg aca gcg gcg gcg ctg tcc tcg 1307
 Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
 1 5 10 15
 gcg ctg acg acg cct gta aag cag cag cag cag cag ctg cgc gta 1355
 Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Gln Leu Arg Val
 20 25 30
 ggc gcg gcg tcg gca cgg ctg gcg gcc gcg gcg ttc tcg tcc ggc acg 1403
 Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Phe Ser Ser Gly Thr
 35 40 45
 ggc gga gac gcg gcc aag aag gcg gcc gcg gcg agg gcg ttc tcc acg 1451
 Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Arg Ala Phe Ser Thr
 50 55 60
 gga cgc ggc ccc aac gcg aca cgc gag aag agc tcg ctg gcc acg gtc 1499
 Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
 65 70 75 80
 cag gcg gcg acg gac gat gcg cgc ttc gtc gcc ctg acc ggc gcc caa 1547
 Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
 85 90 95
 atc ttt cat gag ctc atg cgc gag cac cag gtg gac acc atc ttt ggc 1595
 Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
 100 105 110
 tac cct ggc ggc gcc att ctg ccc gtt ttt gat gcc att ttt gag agt 1643
 Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
 115 120 125
 gac gcc ttc aag ttc att ctc gct cgc cac gag cag ggc gcc ggc cac 1691
 Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
 130 135 140
 atg gcc gag ggc tac gcg cgc gcc acg gcc aag ccc ggc gtt gtc ctc 1739
 Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu
 145 150 155 160
 gtc acc tcg ggc cct gga gcc acc aac acc atc acc ccg atc atg gat 1787
 Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp
 165 170 175
 gct tac atg gac ggt acg ccg ctg ctc gtg ttc acc ggc cag gtg ccc 1835
 Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Pro
 180 185 190
 acc tct gct gtc ggc acg gac gct ttc cag gag tgt gac att gtt ggc 1883
 Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly
 195 200 205
 atc agc cgc gcg tgc acc aag tgg aac gtc atg gtc aag gac gtg aag 1931
 Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys
 210 215 220
 gag ctc ccg cgc cgc atc aat gag gcc ttt gag att gcc atg agc ggc 1979
 Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly
 225 230 235 240
 cgc ccg ggt ccc gtg ctc gtc gat ctt cct aag gat gtg acc gcc gtt 2027
 Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val
 245 250 255
 gag ctc aag gaa atg ccc gac agc tcc ccc cag gtt gct gtg cgc cag 2075
 Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln
 260 265 270
 aag caa aag gtc gag ctt ttc cac aag gag cgc att ggc gct cct gcc 2123
 Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly

ES 2 688 953 T3

Lys	Asn	Trp	Gln	Asp	Leu	Phe	Tyr	Asp	Lys	Arg	Tyr	Ser	Gly	Thr	Ala	
	595						600				605					
atg	ttc	aac	ccg	cgc	ttc	gac	aag	gtc	gcc	gat	gcg	atg	cgt	gcc	aag	3131
Met	Phe	Asn	Pro	Arg	Phe	Asp	Lys	Val	Ala	Asp	Ala	Met	Arg	Ala	Lys	
	610					615				620						
ggt	ctc	tac	tgc	gcg	aaa	cag	tcg	gag	ctc	aag	gac	aag	atc	aag	gag	3179
Gly	Leu	Tyr	Cys	Ala	Lys	Gln	Ser	Glu	Leu	Lys	Asp	Lys	Ile	Lys	Glu	
	625				630					635					640	
ttt	ctc	gag	tac	gat	gag	ggt	ccc	gtc	ctc	ctc	gag	ggt	ttc	gtg	gac	3227
Phe	Leu	Glu	Tyr	Asp	Glu	Gly	Pro	Val	Leu	Leu	Glu	Val	Phe	Val	Asp	
				645					650					655		
aag	gac	acg	ctc	gtc	ttg	ccc	atg	gtc	ccc	gct	ggc	ttt	cgg	ctc	cac	3275
Lys	Asp	Thr	Leu	Val	Leu	Pro	Met	Val	Pro	Ala	Gly	Phe	Pro	Leu	His	
		660						665					670			
gag	atg	gtc	ctc	gag	cct	cct	aag	ccc	aag	gac	gcc	taa	gttctttttt			3324
Glu	Met	Val	Leu	Glu	Pro	Pro	Lys	Pro	Lys	Asp	Ala					
	675						680									
ccatggcggg	cgagcgagcg	agcgcgcgag	cgcgcaagtg	cgcaagcgcc	ttgccttget											3384
ttgcttcgct	tgcctttgct	ttgcttcaca	caacctaatg	atgaattcaa	gttttcttgc											3444
ttgtcgggca	tgccctgctg	ccaaccagcc	agccatccgg	ccggccgctc	ttgacgcctt											3504
cgcttccggc	gggcccatcg	attcaattca	cccataccgat	acgttccggc	ccctcacgtc											3564
cgtctgcgca	gcacccctgc	acgaccacgc	caaggccaac	gcgcccgtca	gctcagcttg											3624
tcgacgagtc	gcaogtcaca	tatctcagat	gcatttggac	tgtgagtgtt	attatgccac											3684
tagcaogcaa	cgatcttcgg	ggtcctcgtc	cattgcatcc	gttcggggcc	tcgaggcgtg											3744
gacgcgagtc	gcgcgcgaga	cgctgcagca	ggccgctccg	acgcgagggc	tcgagctcgc											3804
cgcccccggg	cgatgtctgc	ctggcgccga	ctgatctctg	gagcgcaagg	aagacacggc											3864
gacgcgagga	ggaccgaaga	gagacgctgg	ggtatgcagg	atatacccg	ggcgggacat											3924
tcgttccgca	tacactcccc	cattcgagct	tgctcgtcct	tggcagagcc	gagcgcgaac											3984
ggttcogaac	goggcaagga	ttttggtct	ggtgggtgga	ctccgatcga	ggcgcaggtt											4044
ctccgcaggt	tctcgcaggc	cggcagtggt	cgttagaaat	agggagtgcc	ggagtcttga											4104
cgccgcttag	ctcactctcc	gcccacgcgc	gcatacgcgc	catgccgcgc	tcocgctcgt											4164
cgctgcgctg	gocgcgaccg	gctgcgccag	agtacgacag	tgggacagag	ctcagggcga											4224
cgogaatcgc	togggttgta	agggtttcaa	gggtcgggcg	tcgtcgcgtg	ccaaagtga											4284
aatagtaggg	gggggggggg	gtaccaccc	cgggcaggtt	ctcctcgcca	gcttaagtgc											4344
ctaagggagc	gtaggggttt	cgttgaccag	agaagcggag	aacctgcgc	ggcgcggaga											4404
acctatcggc	ggagaacttg	ccaggcgcga	ggcagttctc	caatttgogg	acagcggcgc											4464
gcccacgcga	ggcggccgcg	tggcgataca	gcgagggcag	cgccgggggc	cgcgctggcg											4524
cacagctgcg	cgccggagtcg	gctgcgagaa	ggcttctcgc	tggcttgggt	ggggctcggg											4584
gtggcagggg	atggatgccc	aggtacgtcg	gcgtgcgcgc	gcccagggag	aaaaggacag											4644
acgcgcgggc	ctgcgatgcg	agcacgcgat	gcgagcacgc	gatgcgagca	cgcgatgcga											4704
gcacgcgagc	gagcgcgccg	gcaaatgcca	cggaacacgc	gttttttgtt	tgggtgattc											4764
tatgtatgcg	gggagacttc	gatggccgaa	aggggtgcaa	ggccaaaaga	tgctgacagc											4824
ttogatcggg	ctacggcgcg	agcaggaaag	ggagcaaggg	gcggaattct	tctgccttga											4884
cccgggggat	ccactagttc	tagagcggcc	gccaccgcgg	tggagctcca	attcgcctta											4944
tagtgagtcg	tattacgcgc	gctcactggc	cgctcgtttt	caacgtcgtg	actgggaaaa											5004
ccctggcggt	acccaactta	atcgccttgc	agcaeatccc	ccittcgcca	gctggcgtaa											5064
tagcgaagag	gcocgcaccg	atcgccttcc	ccaacagttg	cgcagcctga	atggcgaaatg											5124
ggacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	gcagcgtgac											5184
cgctacactt	gccagcgcgc	tagcgcocgc	tcctttcgtc	ttcttccctt	cctttctcgc											5244
caogttcggc	ggctttcccc	gtcaagctct	aaatcggggg	ctcccttag	ggttccgatt											5304
tagtgcttta	cggcacctcg	accccaaaaa	acttgattag	ggtgatgggt	cacgtagtgg											5364
gccatcgcgc	tgatagacgg	tttttcgccc	tttgaccttg	gagtccacgt	tctttaatag											5424
tggactcttg	ttccaaaactg	gaacaacact	caaccctatc	tcggtctatt	ccttttgattt											5484
ataagggatt	ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt											5544
taacgcgaat	tttaacaaaa	tattaaogct	tacaatttag	gtggcacttt	tcggggaaat											5604
gtgcgcggaa	cccctatttg	tttatttttc	taaatacatt	caaataatgta	tcocgctcatg											5664
agacaataac	cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtatccaa											5724
catttccggtg	tcgcccttat	tccttttttt	gcggcatttt	gccttctcgt	ttttgctcac											5784
ccagaaacgc	tggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac											5844

```

atcgaactgg atctcaacag cggtaaagatc cttgagagtt ttcgccccga agaacgtttt 5904
ccaatgatga gcaacttttaa agttctgcta tgtggcgcggt tattateccg tattgaecgc 5964
gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggg tgagtactca 6024
ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtctgc 6084
ataaccatga gtgataacac tgcggccaac ttactttctga caacgatcgg aggaccgaag 6144
gagctaaccg cttttttgca caacatgggg gatcatgtaa ctccgcttga tcgttgggaa 6204
ccggagctga atgaagccat accaaaacgac gagcgtgaca ccacgatgcc thtagcaatg 6264
gcaacaacgt tgcgcaaac attaaactggc gaactactta ctctagcttc cgggcaacaa 6324
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttcg 6384
gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt 6444
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 6504
caggcaacta tggatgaacg aatagacag atcgtgaga taggtgcctc actgattaag 6564
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaaacttcat 6624
tttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 6684
taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 6744
tgagatcctt tttttctgcg cgtaatctgc tgottgcaaa caaaaaaac accgctacca 6804
gcggtgggtt gtttgccgga tcaagagcta ccaactcttt ttccgaagg aactggcttc 6864
agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc 6924
aagaactctg tagcaccgoc tacatacctc gctctgctaa tctgttacc agtggctgct 6984
gccagtgccg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag 7044
gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttgya gcgaacgacc 7104
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tccggaagg 7164
agaaaggcgg acaggtatcc ggtaagcggc agggctggaa caggagagcg cacgagggag 7224
cttccagggg gaaacgcctg gtatctttat agtctgtcgc ggttccgca cctctgactt 7284
gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgcagcaac 7344
gcgccctttt tacggttcct ggcccttttg tggccttttg ctcacatggt ctttctgog 7404
ttatccctg attctgtgga taaccgtatt accgccttg agtgagctga taccgctcgc 7464
cgcagccgaa cgaccgagcg cagcagatca gtgagcaggg aagcggaaaga gcgcccaata 7524
cgcaaaccgc ctctcccgc gcgttggccg attcattaat gcagctggca cgacaggtt 7584
cccgactgga aagcgggagc tgagcgcaac gcaattaatg tgagttagct cactcattag 7644
gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgg 7704
taacaatttc acacaggaaa cagctatgac catgattacg ccaagcgcgc aattaaccct 7764
cactaaaggg aacaaaagct ggtaccggg cccccctcg aggtcgacgg tatcgataag 7824
cttgatatcg aattcctgca gcc 7847

```

<210> 15
 <211> 684
 <212> PRT
 <213> Schizochytrium sp.

5

<400> 15

```

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

```

10

ES 2 688 953 T3

Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100 105 110

Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
115 120 125

Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
130 135 140

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

Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp
165 170 175

Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Pro
180 185 190

Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly
195 200 205

Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys
210 215 220

Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly
225 230 235 240

Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val
245 250 255

Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln
260 265 270

Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly
275 280 285

Thr Ala Asp Phe Lys Leu Ile Ala Glu Met Ile Asn Arg Ala Glu Arg
290 295 300

Pro Val Ile Tyr Ala Gly Gln Gly Val Met Gln Ser Pro Leu Asn Gly
305 310 315 320

Pro Ala Val Leu Lys Glu Phe Ala Glu Lys Ala Asn Ile Pro Val Thr
325 330 335

Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser
340 345 350

Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met
355 360 365

Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg
370 375 380

Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu
385 390 395 400

Arg Glu Gly Arg Gly Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn
405 410 415

ES 2 688 953 T3

Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val
 420 425 430

Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
 435 440 445

Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
 450 455 460

Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
 465 470 475 480

Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala
 485 490 495

Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln
 500 505 510

Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525

Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540

Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560

Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575

Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590

Lys Asn Trp Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605

Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620

Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640

Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655

Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670

Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680

<210> 16
 <211> 30
 <212> ADN
 <213> Schizochytrium sp.

5

<400> 16
 tatcgataag cttgacgtcg aattcctgca 30

10

<210> 17
 <211> 32
 <212> ADN
 <213> Schizochytrium sp.

15

ES 2 688 953 T3

```

<400> 17
  catggtcaag aacgttcagg atctctttta eg      32

5
  <210> 18
  <211> 7847
  <212> ADN
  <213> Schizochytrium sp.

10
  <220>
  <221> ceb_transcripción
  <222> (1)..(1259)
  <223>

15
  <220>
  <221> CDS
  <222> (1260)..(3314)
  <223>

20
  <220>
  <221> ceb_transcripción
  <222> (3315)..(4887)
  <223>

25
  <220>
  <221> mutación
  <222> (3042)..(3044)
  <223>

30
  <400> 18
  ttgtcgacag caacttgcaa gttatacgcg accaccaggc aatctcagca cgcccagcga      60
  gcacggagct tgcgaagagg gtttacaagt cgtcgttcat tcgctctcaa gctttgcctc      120
  aacgcaacta ggcccaggcc taactttcaat gtgtcttgtc ttgcctttca caccgaccga      180
  gtgtgcacaa ccgtgttttg cacaaagcgc aagatgctca ctgactgtg aagaaaggtt      240
  gcgcgcgaagc gactgcgact gcgaggatga ggatgactgg cagcctgttc aaaaactgaa      300
  aatccgcgat gggtcagtgc cattcgcgca tgacgcctgc gagagacaag ttaactcgtg      360
  tcaactggcat gtcctagcat ctttacgcga gcaaaattca atogctttat tttttcagtt      420
  tcgtaacctt ctgcgaaccg cgaatcgccg tttcagcctg actaatctgc agctgocgtg      480
  cactgtcagt cagtcagtca gtcgtgcgcg ctgttccagc accgaggtcg cgcgtgcgcg      540
  cccttgacc gctgctgcta ctgctagtgg cacggcaggt aggagcttgt tgccggaaca      600
  ccagcagccg ccagtcgacg ccagccaggg gaaagtcgg cgtcgaaggg agaggaaggc      660
  ggctgtgca aactaacgtt gaccactggc gccgcgcgac acgagcagga agcaggcagc      720
  tgcagagcgc agcgcgcaag tgcagaatgc gcgaaagatc cacttgcgcg cggcgggocg      780
  gcaacttgcg gcgcggcgcg gaacagtgcg gaaaggagcg gtgcagacgg cgcgcagtga      840
  cagtgggocg aaagccgcgc agtaagcagc ggccgggaaac ggtatacgcg gtgccgcggg      900
  ccgcccacaa cagaagtata cgcgggcccga agtggggcgt cgcgcgcggg aagtgcggaa      960
  tggcgggcaa ggaaaggagg agacgggaaag agggcgggaa agagagagag agagagtgaa      1020
  aaaagaaaga aagaaagaaa gaaagaaaga aagctcggag ccacgcgcg gggagagaga      1080
  gaaatgaaag cacggcacgg caaagcaaag caaagcagac ccagccagac ccagccaggg      1140
  gaggagcgcg cgcaggacc cgcgggcgag cgagcagca cggcgcgcga gcgagcgcgc      1200
  gagcgcgcgc gcgagcgcgc aaggcttgct gcgagcgcgc gagcgcgcga gcgggaagg      1259
  atg agc gcg acc cgc gcg gcg acg agg aca gcg gcg gcg ctg tcc tcg      1307
  Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
  1          5          10          15
  gcg ctg acg acg cct gta aag cag cag cag cag cag cag ctg cgc gta      1355
  
```


ES 2 688 953 T3

acc acc atg cag ggt ctc ggc ggc ttt gac gag cgt agt ccc ctc tcc	2315
Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser	
340 345 350	
ctc aag atg ctc ggc atg cac ggc tct gcc tac gcc aac tac tcg atg	2363
Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met	
355 360 365	
cag aac gcc gat ctt atc ctg gcg ctc ggt gcc cgc ttt gat gat cgt	2411
Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg	
370 375 380	
gtg acg ggc cgc gtt gac gcc ttt gct ccg gag gct cgc cgt gcc gag	2459
Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu	
385 390 395 400	
cgc gag ggc cgc ggt ggc atc gtt cac ttt gag att tcc ccc aag aac	2507
Arg Glu Gly Arg Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn	
405 410 415	
ctc cac aag gtc gtc cag ccc acc gtc gcg gtc ctc gcc gac gtg gtc	2555
Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val	
420 425 430	
gag aac ctc gcc aac gtc acg ccc cac gtg cag cgc cag gag cgc gag	2603
Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu	
435 440 445	
ccg tgg ttt gcg cag atc gcc gat tgg aag gag aag cac cct ttt ctg	2651
Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu	
450 455 460	
ctc gag tot gtt gat tcg gac gac aag gtt ctc aag ccg cag cag gtc	2699
Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val	
465 470 475 480	
ctc acg gag ctt aac aag cag att ctc gag att cag gag aag gac gcc	2747
Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala	
485 490 495	
gac cag gag gtc tac atc acc acg ggc gtc gga agc cac cag atg cag	2795
Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln	
500 505 510	
gca gcg cag ttc ctt acc tgg acc aag ccg cgc cag tgg atc tcc tcg	2843
Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser	
515 520 525	
ggt ggc gcc ggc act atg ggc tac ggc ctt ccc tcg gcc att ggc gcc	2891
Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala	
530 535 540	
aag att gcc aag ccc gat gct att gtt att gac atc gat ggt gat gct	2939
Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala	
545 550 555 560	
tot tat tog atg acc ggt atg gaa ttg atc aca gca gcc gaa ttc aag	2987
Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys	
565 570 575	
gtt ggc gtg aag att ctt ctt ttg cag aac aac ttt cag ggc atg gtc	3035
Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val	
580 585 590	
aag aac gtt cag gat ctc ttt tac gac aag cgc tac tcg gcc acc gcc	3083
Lys Asn Val Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala	
595 600 605	
atg ttc aac ccg cgc ttc gac aag gtc gcc gat gcg atg cgt gcc aag	3131
Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys	
610 615 620	
ggt ctc tac tgc gcg aaa cag tog gag ctc aag gac aag atc aag gag	3179
Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu	
625 630 635 640	
ttt ctc gag tac gat gag ggt ccc gtc ctc cag gtt ttc gtg gac	3227
Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp	

ES 2 688 953 T3

	645		650		655	
aag gac acg ctc gtc ttg ccc atg gtc ccc gct ggc ttt cgg ctc cac						3275
Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His						
	660		665		670	
gag atg gtc ctc gag cct cct aag ccc aag gac gcc taa gttctttttt						3324
Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala						
	675		680			
ccatggcggg cgagcgcgag agcgcgcgag cgcgcgaagt gcgaagcgcc ttgccttgct						3384
ttgcttcgct tgcctttgct ttgcttcaca caacctaat atgaattcaa gttttcttgc						3444
ttgtcggcga tgccctgcctg ccaaccagcc agccatccgg cggcccgcc itgacgcctt						3504
cgcttcgggc gggccatcg attcaattca cccatccgat acgttcggcc cccctacgctc						3564
cgctcgcgca cgaaccctgc acgaccacgc caaggccaac ggcgcgctca gctcagcttg						3624
tcgacgagtc gcaogtcaca tatctcagat goatttgagc tgtgagtgtt attatgccac						3684
tagcagcгаа cgatcttcgg ggtccctcgt cattgcatcc gttcggggccc tgcaggcgtg						3744
gacgcgagtc gccgcgcgaga cgtcgcagca ggccgctccg acgcgcggggc tgcagctcgc						3804
cgcgcccgcg cgatgtctgc ctggcgcgca ctgatctctg gagcgcgaagg aagacacggc						3864
gacgcgagga ggaccgaaga gagacgctgg ggtatgcagg atataaccgg ggccggacat						3924
tcgttcggca taactcctcc cattcagact tgctcgtcct tggcagagcc gagcgcgaac						3984
ggttccgaac ggcgcgaagga ttttgctct ggtgggtgga ctccgatcga ggccgaggtt						4044
ctccgcaggt tctcgcaggc cggcagtggt cgttagaaat agggagtgcc ggagtcttga						4104
cgcgccttag ctactctcc gccacgcgc gcacgcgcgc catgcccgcg tcccgtctgt						4164
cgctgcgctg gccgcgaccg cctcgcgcag cgtcgcgcag agtacgcag tgggacagag						4224
cgcgaaatcg tcgggttgta agggtttcaa ggtcggggcg tcgtcgcgtg ccaaaagtga						4284
aatagtaggg gggggggggg gtaccacccc cgggcaggtt ctctcgcga gcctaagtgc						4344
ctaagggagc gtaggggttt cgttgaccag agaagcggag aacctgccgc ggccgcggaga						4404
acctatcggc ggagaacttg ccaggcgcga ggcagttctc caatttgccg acagcggcgc						4464
gccacgcgga gccggccgcg tggcgataca gcgagggcag cgcgcggggc cgcgtggcga						4524
cacagctggc cgcggagtcg gctgcgagaa ggcttctcgc tggcttggtt ggggtcgcgg						4584
gtggcagggg atggatgccc aggtacgtcg gcgtgcgcgc gcccagggag aaaagacag						4644
acgcgcgggc ctgcgatgcg agcacgcgat gcgagcacgc gatgcgagca cgcgatgcga						4704
gcacgcgagc gagcgcgccg gcaaatgcc cggaacacgc gttttttggt tggtgatttc						4764
tatgtatgcg gggagacttc gatggccgaa aggggtgcaa ggccaaaaga tgctgacagc						4824
ttcgatcggc ctacggcgcg agcaggaaag ggagcaaggg ggggaattct tctgccttga						4884
cccgggggat ccactagttc tagagcggcc gccacgcgcg tcgagctoca attcgccta						4944
tagtgagtcg tattacgcgc gctcactggc cgtcgtttta caactcgtg actgggaaaa						5004
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttgcga gctggcgtaa						5064
tagcgaagag gcccgcaccg atcgccttc ccaacagttg cgcagcctga atggcgaatg						5124
ggaagcgcoc tgtagcggcg cattaagcgc ggcgggtgtg gtggttaocg gcagcgtgac						5184
cgctacactt gccagcgcgc tagcgcgcgc tcccttctgc cctttccctt cctttctcgc						5244
cagttcgcgc ggccttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt						5304
tagtgcctta cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg						5364
gccatcgccc tgatagacgg ttttgcgcc tttgacgttg gagtccactt tctttaatag						5424
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt						5484
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt						5544
taacgcgaaat ttttaacaaa tattaacgct tacaatttag gtggcacttt tcggggaaat						5604
gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg						5664
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa						5724
catttcgctg tcgcccctat tccctttttt gcggcatttt gccttcctgt ttttgctcac						5784
ccagaaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac						5844
atcgaactgg atctcaacag cggtaagatc cttgagagtt ttgcccccg agaaogtttt						5904
ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc						5964
gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggg tgagtactca						6024
ccagtacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgcgcc						6084
ataaccatga gtgataaacac tgcggccaac ttacttctga caacgatcgg aggaccggaag						6144
gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga tcggtgggaa						6204
ccggagctga atgaagccat accaaaagac gagcgtgaca ccacgatgcc tgtagcaatg						6264
gcaacaacgt tcgcgaaact attaactggc gaactactta ctctagcttc ccggcaacaa						6324
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttcgg						6384
gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatcattt						6444

ES 2 688 953 T3

```

gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 6504
caggcaacta tggatgaacg aatagacag atcgtcgaga taggtgcctc actgattaag 6564
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 6624
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 6684
taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 6744
tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca 6804
gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc 6864
agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc 6924
aagaactctg tagcaccgcc tacatacctc gctctgctaa tectgttacc agtggctgct 6984
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag 7044
gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg gcgaacgacc 7104
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg 7164
agaaaggcgg acaggtatcc ggtaagcggc agggctggaa caggagagcg cagagggag 7224
cttccagggg gaaacgcctg gtatctttat agtccctgtc ggtttcgcca cctctgactt 7284
gagcgtogat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac 7344
gocggcctttt tacggttcct ggccttttgc tggccttttg ctacatgkt ctttctgcg 7404
ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc 7464
cgcagccgaa cgaccgagcg cagcaggtca gtgagcaggg aagcgggaaga gcgcccaata 7524
cgcaaaccgc ctctcccgc gcgttggccg attcattaat gcagctggca cgacaggttt 7584
cccgaactgga aagcgggcag tgagcgaac gcaattaatg tgagttagct cactcattag 7644
gcaccocagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga 7704
taacaatttc acacaggaaa cagctatgac catgattacg ccaagcgcgc aattaaccct 7764
cactaaaggg aacaaaagct ggtaccggg cccccctcg aggtcgacgy tatcgataag 7824
cttgacgtcg aattcctgca gcc 7847

```

<210> 19
 <211> 684
 <212> PRT
 <213> Schizochytrium sp.

5

<400> 19

```

Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1          5          10          15

Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Leu Arg Val
20          25          30

Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
35          40          45

Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
50          55          60

Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
65          70          75          80

Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
85          90          95

Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100         105         110

Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
115         120         125

Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
130         135         140

Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu

```

10

ES 2 688 953 T3

145					150					155				160	
Val	Thr	Ser	Gly	Pro	Gly	Ala	Thr	Asn	Thr	Ile	Thr	Pro	Ile	Met	Asp
				165					170					175	
Ala	Tyr	Met	Asp	Gly	Thr	Pro	Leu	Leu	Val	Phe	Thr	Gly	Gln	Val	Pro
			180					185					190		
Thr	Ser	Ala	Val	Gly	Thr	Asp	Ala	Phe	Gln	Glu	Cys	Asp	Ile	Val	Gly
		195					200					205			
Ile	Ser	Arg	Ala	Cys	Thr	Lys	Trp	Asn	Val	Met	Val	Lys	Asp	Val	Lys
	210					215					220				
Glu	Leu	Pro	Arg	Arg	Ile	Asn	Glu	Ala	Phe	Glu	Ile	Ala	Met	Ser	Gly
225					230					235					240
Arg	Pro	Gly	Pro	Val	Leu	Val	Asp	Leu	Pro	Lys	Asp	Val	Thr	Ala	Val
				245					250					255	
Glu	Leu	Lys	Glu	Met	Pro	Asp	Ser	Ser	Pro	Gln	Val	Ala	Val	Arg	Gln
			260					265					270		
Lys	Gln	Lys	Val	Glu	Leu	Phe	His	Lys	Glu	Arg	Ile	Gly	Ala	Pro	Gly
		275					280					285			
Thr	Ala	Asp	Phe	Lys	Leu	Ile	Ala	Glu	Met	Ile	Asn	Arg	Ala	Glu	Arg
	290					295					300				
Pro	Val	Ile	Tyr	Ala	Gly	Gln	Gly	Val	Met	Gln	Ser	Pro	Leu	Asn	Gly
305					310					315					320
Pro	Ala	Val	Leu	Lys	Glu	Phe	Ala	Glu	Lys	Ala	Asn	Ile	Pro	Val	Thr
				325					330					335	
Thr	Thr	Met	Gln	Gly	Leu	Gly	Gly	Phe	Asp	Glu	Arg	Ser	Pro	Leu	Ser
			340					345					350		
Leu	Lys	Met	Leu	Gly	Met	His	Gly	Ser	Ala	Tyr	Ala	Asn	Tyr	Ser	Met
		355					360					365			
Gln	Asn	Ala	Asp	Leu	Ile	Leu	Ala	Leu	Gly	Ala	Arg	Phe	Asp	Asp	Arg
	370					375					380				
Val	Thr	Gly	Arg	Val	Asp	Ala	Phe	Ala	Pro	Glu	Ala	Arg	Arg	Ala	Glu
385					390					395					400
Arg	Glu	Gly	Arg	Gly	Gly	Ile	Val	His	Phe	Glu	Ile	Ser	Pro	Lys	Asn
				405					410					415	
Leu	His	Lys	Val	Val	Gln	Pro	Thr	Val	Ala	Val	Leu	Gly	Asp	Val	Val
			420					425					430		
Glu	Asn	Leu	Ala	Asn	Val	Thr	Pro	His	Val	Gln	Arg	Gln	Glu	Arg	Glu
		435					440					445			
Pro	Trp	Phe	Ala	Gln	Ile	Ala	Asp	Trp	Lys	Glu	Lys	His	Pro	Phe	Leu
	450					455					460				

ES 2 688 953 T3

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

<210> 20
 <211> 26
 <212> ADN
 <213> Schizochytrium sp.

<400> 20
 ccggccaggt gcagacctct gctgtc 26

<210> 21
 <211> 7847
 <212> ADN
 <213> Schizochytrium sp.

<220>
 <221> ceb_transcripción
 <222> (1)..(1259)
 <223>

<220>
 <221> mutación
 <222> (1334)..(1835)
 <223>

ES 2 688 953 T3

<220>
 <221> CDS
 <222> (1260)..(3314)
 <223>

5

<220>
 <221> ceb_transcripción
 <222> (3315)..(4887)
 <223>

10

<400> 21

```

ttgtcgacag caacttgcaa gttatacgcg accaccaggc aatctcagca cgcccagcga      60
gcacggagct tgcgaagagg gtttacaagt cgtcgttcat tcgctctcaa gctttgcctc      120
aacgcaacta ggcccaggcc tactttcact gtgtcttgtc ttgcctttca caccgaccga      180
gtgtgcacaa ccgtgttttg cacaaagcgc aagatgctca ctcgactgtg aagaaaggtt      240
gcgcgcaagc gactgcgact gcgaggatga ggatgactgg cagcctgttc aaaaactgaa      300
aatccgcgat gggtcagtgcc cattcgcgca tgaogcctgc gagagacaag ttaactcgtg      360
tcaactggcat gtcctagcat ctttaocgga gcaaaattca atcgttttat tttttcagtt      420
tcgtaacctt ctgcgaaccg cgaatcgcgc tttcagcctg actaatctgc agctgcgtgg      480
cactgtcagt cagtcagtca gtcgtgocgc ctgttccagc accgaggtcg cgcgtcgcgc      540
cgccctggacc gctgctgcta ctgctagtgg cacggcaggt aggagcttgt tgccggaaca      600
ccagcagccg ccagtcgacg ccagccaggg agagtcctcg cgtcgaaggg agaggaaggg      660
ggcgtgtgca aactaacgtt gaccactggc gccgcgcgac acgagcagga agcagggcagc      720
tgcagagcgc agcgcgcaag tgcagaatgc gcgaaagatc caettgcgcg cggcggggcgc      780
gcacttgccg gcgcggcgcg gaacagtgcg gaaaggagcg gtgcagacgg cgcgcagtga      840
cagtgggcgc aaagccgcgc agtaagcagc ggccggggaac ggtatacgcg gtgcccgggg      900
ccgcccacaca cagaagtata cgcgggcoga agtggggcgt cgcgcgcggg aagtgcggaa      960
tggcggggcaa ggaaggagg agacggaaag agggcgggaa agagagagag agagagtgaa     1020
aaaagaaaga aagaaagaaa gaaagaaaga aagctcggag ccacgcgcgc gggagagaga     1080
gaaatgaaag cacggcaagg caaagcaaag caaagcagac ccagccagac ccagccaggg     1140
gaggagcgcg cgcaggaccc gcgcggcgcg cgagcgcgca cggcgcgcga gcgagcgcgc     1200
gagcgcgcgc gcgagcgcgc aagccttget gcgagcgcgc gagcgcgcga gcggggaagg     1259
atg agc gcg acc cgc gcg gcg acg agg aca gcg gcg gcg ctg tcc tcg     1307
Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1      5      10      15
gcg ctg acg acg cct gta aag cag cag cag cag cag cag ctg cgc gta     1355
Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Gln Leu Arg Val
20     25     30
ggc gcg gcg tcg gca cgg ctg gcg gcc gcg gcg ttc tcg tcc ggc acg     1403
Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
35     40     45
ggc gga gac gcg gcc aag aag gcg gcc gcg gcg agg gcg ttc tcc acg     1451
Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
50     55     60
gga cgc gcc ccc aac gcg aca cgc gag aag agc tcg ctg gcc acg gtc     1499
Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
65     70     75     80
cag gcg gcg acg gac gat gcg cgc ttc gtc gcc ctg acc ggc gcc caa     1547
Gln Ala Ala Thr Asp Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
85     90     95
atc ttt cat gag ctc atg cgc gag cac cag gtg gac acc atc ttt ggc     1595
Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100    105    110
tac cct gcc gcc att ctg ccc gtt ttt gat gcc att ttt gag agt     1643
    
```


ES 2 688 953 T3

Tyr	Pro	Gly	Gly	Ala	Ile	Leu	Pro	Val	Phe	Asp	Ala	Ile	Phe	Glu	Ser		
		115					120					125					
gac	gcc	ttc	aag	ttc	att	ctc	gct	cgc	cac	gag	cag	ggc	gcc	ggc	cac	1691	
Asp	Ala	Phe	Lys	Phe	Ile	Leu	Ala	Arg	His	Glu	Gln	Gly	Ala	Gly	His		
		130				135					140						
atg	gcc	gag	ggc	tac	gcg	cgc	gcc	acg	ggc	aag	ccc	ggc	gtt	gtc	ctc	1739	
Met	Ala	Glu	Gly	Tyr	Ala	Arg	Ala	Thr	Gly	Lys	Pro	Gly	Val	Val	Leu		
		145			150					155					160		
gtc	acc	tcg	ggc	cct	gga	gcc	acc	aac	acc	atc	acc	ccg	atc	atg	gat	1787	
Val	Thr	Ser	Gly	Pro	Gly	Ala	Thr	Asn	Thr	Ile	Thr	Pro	Ile	Met	Asp		
				165					170					175			
gct	tac	atg	gac	ggt	acg	ccg	ctg	ctc	gtg	ttc	acc	ggc	cag	gtg	cag	1835	
Ala	Tyr	Met	Asp	Gly	Thr	Pro	Leu	Leu	Val	Phe	Thr	Gly	Gln	Val	Gln		
			180					185					190				
acc	tot	gct	gtc	ggc	acg	gac	gct	ttc	cag	gag	tgt	gac	att	gtt	ggc	1883	
Thr	Ser	Ala	Val	Gly	Thr	Asp	Ala	Phe	Gln	Glu	Cys	Asp	Ile	Val	Gly		
		195					200					205					
atc	agc	cgc	gcg	tgc	acc	aag	tgg	aac	gtc	atg	gtc	aag	gac	gtg	aag	1931	
Ile	Ser	Arg	Ala	Cys	Thr	Lys	Trp	Asn	Val	Met	Val	Lys	Asp	Val	Lys		
		210				215					220						
gag	ctc	ccg	cgc	cgc	atc	aat	gag	gcc	ttt	gag	att	gcc	atg	agc	ggc	1979	
Glu	Leu	Pro	Arg	Arg	Ile	Asn	Glu	Ala	Phe	Glu	Ile	Ala	Met	Ser	Gly		
		225			230					235					240		
cgc	ccg	ggt	ccc	gtg	ctc	gtc	gat	ctt	cct	aag	gat	gtg	acc	gcc	gtt	2027	
Arg	Pro	Gly	Pro	Val	Leu	Val	Asp	Leu	Pro	Lys	Asp	Val	Thr	Ala	Val		
				245				250						255			
gag	ctc	aag	gaa	atg	ccc	gac	agc	tcc	ccc	cag	gtt	gct	gtg	cgc	cag	2075	
Glu	Leu	Lys	Met	Pro	Asp	Ser	Ser	Pro	Gln	Val	Ala	Val	Arg	Gln			
			260				265						270				
aag	caa	aag	gtc	gag	ctt	ttc	cac	aag	gag	cgc	att	ggc	gct	cct	ggc	2123	
Lys	Gln	Lys	Val	Glu	Leu	Phe	His	Lys	Glu	Arg	Ile	Gly	Ala	Pro	Gly		
		275				280						285					
acg	gcc	gac	ttc	aag	ctc	att	gcc	gag	atg	atc	aac	cgt	gcg	gag	cga	2171	
Thr	Ala	Asp	Phe	Lys	Leu	Ile	Ala	Glu	Met	Ile	Asn	Arg	Ala	Glu	Arg		
		290			295						300						
ccc	gtc	atc	tat	gct	ggc	cag	ggt	gtc	atg	cag	agc	ccg	ttg	aat	ggc	2219	
Pro	Val	Ile	Tyr	Ala	Gly	Gln	Gly	Val	Met	Gln	Ser	Pro	Leu	Asn	Gly		
				305	310					315				320			
ccg	gct	gtg	ctc	aag	gag	ttc	gcg	gag	aag	gcc	aac	att	ccc	gtg	acc	2267	
Pro	Ala	Val	Leu	Lys	Glu	Phe	Ala	Glu	Lys	Ala	Asn	Ile	Pro	Val	Thr		
				325					330					335			
acc	acc	atg	cag	ggt	ctc	ggc	ggc	ttt	gac	gag	cgt	agt	ccc	ctc	tcc	2315	
Thr	Thr	Met	Gln	Gly	Leu	Gly	Gly	Phe	Asp	Glu	Arg	Ser	Pro	Leu	Ser		
			340					345					350				
ctc	aag	atg	ctc	ggc	atg	cac	ggc	tot	gcc	tac	gcc	aac	tac	tcg	atg	2363	
Leu	Lys	Met	Leu	Gly	Met	His	Gly	Ser	Ala	Tyr	Ala	Asn	Tyr	Ser	Met		
		355					360					365					
cag	aac	gcc	gat	ctt	atc	ctg	cgc	ctc	ggt	gcc	cgc	ttt	gat	gat	cgt	2411	
Gln	Asn	Ala	Asp	Leu	Ile	Leu	Ala	Leu	Gly	Ala	Arg	Phe	Asp	Asp	Arg		
		370			375						380						
gtg	acg	ggc	cgc	gtt	gac	gcc	ttt	gct	ccg	gag	gct	cgc	cgt	gcc	gag	2459	
Val	Thr	Gly	Arg	Val	Asp	Ala	Phe	Ala	Pro	Glu	Ala	Arg	Arg	Ala	Glu		
		385			390					395				400			
cgc	gag	ggc	cgc	ggt	ggc	atc	gtt	cac	ttt	gag	att	tcc	ccc	aag	aac	2507	
Arg	Glu	Gly	Arg	Gly	Gly	Ile	Val	His	Phe	Glu	Ile	Ser	Pro	Lys	Asn		
				405					410					415			
ctc	cac	aag	gtc	gtc	cag	ccc	acc	gtc	gcg	gtc	ctc	ggc	gac	gtg	gtc	2555	
Leu	His	Lys	Val	Val	Gln	Pro	Thr	Val	Ala	Val	Leu	Gly	Asp	Val	Val		
			420					425					430				

ES 2 688 953 T3

gag aac ctc gcc aac gtc acg ccc cac gtg cag cgc cag gag cgc gag 2603
 Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
 435 440 445
 ccg tgg ttt gcg cag atc gcc gat tgg aag gag aag cac cct ttt ctg 2651
 Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
 450 455 460
 ctc gag tot gtt gat tcg gac gac aag gtt ctc aag ccg cag cag gtc 2699
 Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
 465 470 475 480
 ctc acg gag ctt aac aag cag att ctc gag att cag gag aag gac gcc 2747
 Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala
 485 490 495
 gac cag gag gtc tac atc acc acg gcc gtc gga agc cac cag atg cag 2795
 Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln
 500 505 510
 gca gcg cag ttc ctt acc tgg acc aag ccg cgc cag tgg atc tcc tcg 2843
 Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525
 ggt gcc gcc gcc act atg gcc tac gcc ctt ccc tcg gcc att gcc gcc 2891
 Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540
 aag att gcc aag ccc gat gct att gtt att gac atc gat ggt gat gct 2939
 Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560
 tct tat tcg atg acc ggt atg gaa ttg atc aca gca gcc gaa ttc aag 2987
 Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 gtt gcc gtg aag att ctt ctt ttg cag aac aac ttt cag gcc atg gtc 3035
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 aag aac tgg cag gat ctc ttt tac gac aag cgc tac tcg gcc acc gcc 3083
 Lys Asn Trp Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605
 atg ttc aac ccg cgc ttc gac aag gtc gcc gat ccg atg cgt gcc aag 3131
 Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620
 ggt ctc tac tgc gcg aaa cag tcg gag ctc aag gac aag atc aag gag 3179
 Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640
 ttt ctc gag tac gat gag ggt ccc gtc ctc ctc gag gtt ttc gtg gac 3227
 Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655
 aag gac acg ctc gtc ttg ccc atg gtc ccc gct gcc ttt ccg ctc cac 3275
 Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670
 gag atg gtc ctc gag cct cct aag ccc aag gac gcc taa gttctttttt 3324
 Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680
 ccattggcggg cgagcgagcg agcgcgcgag cgcgcaagtg cgcaagcgcc ttgccttgct 3384
 ttgcttcgct tcgctttgct ttgcttcaca caacctaatg atgaattcaa gttttcttgc 3444
 ttgtcggcga tgccctgctg ccaaccagcc agccatccgg ccggccgctc ttgacgcctt 3504
 cgcttcgggc gcggccatcg attcaattca cccatccgat acgttccgcc cctcaccgtc 3564
 cgtctgcgca cgaccctgc acgaccacgc caaggccaac gcgccgctca gtcagcttg 3624
 tcgacgagtc gcacgtcaca tatctcagat gcatttggac tgtgagtgtt attatgccac 3684
 tagcacgcaa cgatcttcgg ggtcctogct cattgcatcc gitcgggccc tcgaggcgtg 3744
 gacgcgagtc gcgcgcgaga cgcctgcagca ggccgctccg acgcgagggc tcgagctcgc 3804
 cgcgcccgcg cgatgtctgc ctggcgcoga ctgatctctg gagcgcaagg aagacacggc 3864
 gacgcgagga ggaccgaaga gagacgctgg ggtatgcagg atataccgg ggccggacat 3924
 tcgttccgca tacaactccc cattcagact tgctcgtcct tggcagagcc gagcgcgaac 3984

ggttcogaac	goggaagga	ttttggctct	ggtgggtgga	ctccgatoga	ggcgcaggtt	4044
ctccgcaggt	tctgcaggo	oggcagtggt	cgtagaagt	agggagtgc	ggagtcttga	4104
cgccgcttag	ctcactctcc	gcccacgcgc	gcatgcgcgc	catgcgcgcg	tcccgtctgt	4164
cgctgccttg	gcccgcgaccg	gctgcgccag	agtacgacag	tgggacagag	ctcgaggcga	4224
cgcgaaatcgc	tcgggttgta	agggtttcaa	gggtcgggcg	tcgtcgcgtg	ccaaagtgaa	4284
aatagtaggg	gggggggggg	gtacccaccc	cgggcaggtt	ctcctcgcca	gcctaagtgc	4344
ctaagggagc	gtagggggtt	cgttgaccag	agaagcggag	aacctgcgcg	ggcgcggaga	4404
acctatcggc	ggagaacttg	ccaggcgcga	ggcagttctc	caatttgccg	acagcggcgc	4464
gcccacgcga	ggcggcgcgc	tggcgataca	gcgaggogac	cgccgggggc	cgcgtggcga	4524
cacagctgcg	cgccgagtcg	gctgcgagaa	ggcttctcgc	tggcttgggt	ggggtcgcgg	4584
gtggcagggg	atggatgccc	aggtacgtcg	gcgtgcgcgc	gcccaggag	aaaaggacag	4644
acgcgcgggc	ctgcgatgcg	agcacgcgat	gcgagcacgc	gatgcgagca	cgcgatgcga	4704
gcacgcgagc	gagcgcgccg	gcaaatgcc	cggaacacgc	gttttttggt	tgggtatttc	4764
tatgatgcg	gggagacttc	gatggccgaa	aggggtgcaa	ggccaaaaga	tgctgacagc	4824
ttcgatcggg	ctacggcgcg	agcaggaaag	ggagcaagg	gcggaattct	tctgccttga	4884
cccgggggat	ccactagttc	tagagcggcc	gcccgcggcg	tggagctcca	atcgcoccta	4944
tagtgcgtog	tattacgcgc	gctcactggc	cgctgtttta	caacgtcgtg	actgggaaaa	5004
ccctggcgtt	acccaactta	atcgccttgc	agcacatccc	cccttcgcca	gctggcgtaa	5064
tagcgaagag	gcccgcaccg	atcgccttcc	ccaacagttg	cgcagcctga	atggcgaatg	5124
ggacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	gcagcgtgac	5184
cgctacactt	gcccagcgcg	tagcgcgcgc	tcctttcgct	ttcttccctt	cccttctcgc	5244
caegtctcgc	ggctttcccc	gtcaagctct	aaatcgggg	ctccctttag	ggttcogatt	5304
tagtgcctta	gcccacctcg	accccaaaaa	acttgattag	ggtgatgggt	cacgtatggt	5364
gccatcgcgc	tgatagacgg	tttttcgccc	tttgacgttg	gagtcacagt	tctttaatag	5424
tggactcttg	ttccaaactg	gaacaacact	caacctatc	tcggctctatt	cttttgattt	5484
ataagggatt	ttgccgattt	cgccctattg	gttaaaaaat	gagctgattt	aacaaaaatt	5544
taacycgaat	tttaacaaaa	tattaacgct	tacaatttag	gtggcacttt	tcggggaaat	5604
gtgcgcggaa	ccctattttg	tttatttttc	taaatacatt	caaataatgta	tcgcctcatg	5664
agacaataaac	ccctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagttatcaa	5724
catttcogtg	tcgccccttat	tccttttttt	gocgcatttt	gccttctcgt	ttttgctcac	5784
ccagaaaacgc	tgggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	5844
atcgaactgg	atctcaacag	cggtaaagtc	cttgagaggt	ttcgcocccg	agaacgttt	5904
ccaatgatga	gcactttttaa	agttctgcta	tgtggcgcgg	tattatcccg	tattgacgcc	5964
gggcaagagc	aaactcggctg	ccgcatacac	tattctcaga	atgacttggg	tgagctctca	6024
ccagtcacag	aaaagcatct	tacggatggc	gagacagtaa	gagaattatg	cagtgcctcc	6084
ataaccatga	gtgataaac	tgccgccaac	ttacttctga	caacgatcgg	aggaccgaag	6144
gagctaaacg	cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	6204
ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	6264
gcaacaacgt	tgccgaaact	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	6324
ttaatagact	ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgcctc	ggcccttcgg	6384
gctggctggt	ttattgctga	taaatctgga	gcgggtcagc	gtgggtctcg	cggtatcatt	6444
gcagcactgg	ggccagatgg	taagccctcc	cgtagctgag	ttatctacac	gacggggagt	6504
caggcaacta	tggatgaacg	aatagacag	atcgcctgaga	taggtgcctc	actgattaag	6564
cattggtaac	tgtcagacca	agtttactca	tatatacttt	agattgattt	aaaacttcct	6624
ttttaattta	aaaggatcta	ggtgaagatc	ctttttgata	atctcatgac	caaaatccct	6684
taacgtgagt	tttctgtcca	ctgagcgtca	gaccccgtag	aaaagatcaa	aggatcttct	6744
tgatctcctt	tttttctgcg	cgtaatctgc	tgcttgcaaa	caaaaaaac	acogctacca	6804
gcggtgggtt	gtttgcggga	tcaagagcta	ccaactcttt	ttccgaaggt	aaactggctt	6864
agcagagcgc	agataccaaa	tactgtcctt	ctagtgtagc	cgtagttagg	ccaccacttc	6924
aagaactctg	tagcaccgcc	tacatacctc	gotctgctaa	tcctgttacc	agtggctgct	6984
gccagtggcg	ataagtcgtg	tcttaccggg	ttggactcaa	gacgatagtt	accggataag	7044
gcgcagcggg	cgggctgaac	ggggggttcc	tgcacacagc	ccagcttggg	gcgaacgacc	7104
tacaocgaa	tgagatacct	acagcgtgag	ctatgagaaa	gcgccacgct	tcccgaaggg	7164
agaaagggcg	acaggatccc	ggtaaagggc	agggctggaa	caggagagcg	cacgagggag	7224
cttccagggg	gaaacgcctg	gtatctttat	agtcctgctg	ggtttcgcca	cctctgactt	7284
gagcgtogat	ttttgtgatg	ctcgtcaggg	ggcgggagcc	tatggaaaaa	cgccagcaac	7344
goggcctttt	tacggttcct	ggccttttgc	tggccttttg	ctcacatggt	ctttcctgcg	7404
ttatccoctg	attctgtgga	taaccgtatt	accgcctttg	agtgagctga	taaccgtcgc	7464
cgcagccgaa	cgaccgagcg	cagcaggtca	gtgagcggag	aagcgggaaga	gcgcccaata	7524
cgcaaacgcg	ctctccccgc	gcgttggccg	attcattaat	gcagctggca	cgacaggttt	7584
ccogactgga	aagcgggag	tgagcgaac	gcaattaatg	tgagttagct	cactcattag	7644
gcaccccagg	ctttacactt	tatgcttccg	gctcgtatgt	tgtgtggaat	tgtgagcggg	7704
taacaatttc	acacaggaaa	cagctatgac	catgattacg	ccaagcgcgc	aattaacct	7764
cactaaaggg	aacaaaagct	ggtaaccggg	ccccccctcg	aggtcgcagc	tatcgataag	7824
cttgacgtcg	aattcctgca	gcc				7847

ES 2 688 953 T3

<210> 22
 <211> 684
 <212> PRT
 <213> Schizochytrium sp.

5

<400> 22

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

ES 2 688 953 T3

				245						250					255	
Glu	Leu	Lys	Glu	Met	Pro	Asp	Ser	Ser	Pro	Gln	Val	Ala	Val	Arg	Gln	
			260					265					270			
Lys	Gln	Lys	Val	Glu	Leu	Phe	His	Lys	Glu	Arg	Ile	Gly	Ala	Pro	Gly	
		275					280					285				
Thr	Ala	Asp	Phe	Lys	Leu	Ile	Ala	Glu	Met	Ile	Asn	Arg	Ala	Glu	Arg	
	290					295					300					
Pro	Val	Ile	Tyr	Ala	Gly	Gln	Gly	Val	Met	Gln	Ser	Pro	Leu	Asn	Gly	
305					310					315					320	
Pro	Ala	Val	Leu	Lys	Glu	Phe	Ala	Glu	Lys	Ala	Asn	Ile	Pro	Val	Thr	
				325					330					335		
Thr	Thr	Met	Gln	Gly	Leu	Gly	Gly	Phe	Asp	Glu	Arg	Ser	Pro	Leu	Ser	
			340					345					350			
Leu	Lys	Met	Leu	Gly	Met	His	Gly	Ser	Ala	Tyr	Ala	Asn	Tyr	Ser	Met	
		355					360					365				
Gln	Asn	Ala	Asp	Leu	Ile	Leu	Ala	Leu	Gly	Ala	Arg	Phe	Asp	Asp	Arg	
	370					375					380					
Val	Thr	Gly	Arg	Val	Asp	Ala	Phe	Ala	Pro	Glu	Ala	Arg	Arg	Ala	Glu	
385					390					395					400	
Arg	Glu	Gly	Arg	Gly	Gly	Ile	Val	His	Phe	Glu	Ile	Ser	Pro	Lys	Asn	
				405					410					415		
Leu	His	Lys	Val	Val	Gln	Pro	Thr	Val	Ala	Val	Leu	Gly	Asp	Val	Val	
			420					425					430			
Glu	Asn	Leu	Ala	Asn	Val	Thr	Pro	His	Val	Gln	Arg	Gln	Glu	Arg	Glu	
		435					440					445				
Pro	Trp	Phe	Ala	Gln	Ile	Ala	Asp	Trp	Lys	Glu	Lys	His	Pro	Phe	Leu	
	450					455					460					
Leu	Glu	Ser	Val	Asp	Ser	Asp	Asp	Lys	Val	Leu	Lys	Pro	Gln	Gln	Val	
465					470					475					480	
Leu	Thr	Glu	Leu	Asn	Lys	Gln	Ile	Leu	Glu	Ile	Gln	Glu	Lys	Asp	Ala	
				485					490					495		
Asp	Gln	Glu	Val	Tyr	Ile	Thr	Thr	Gly	Val	Gly	Ser	His	Gln	Met	Gln	
			500					505					510			
Ala	Ala	Gln	Phe	Leu	Thr	Trp	Thr	Lys	Pro	Arg	Gln	Trp	Ile	Ser	Ser	
		515					520					525				
Gly	Gly	Ala	Gly	Thr	Met	Gly	Tyr	Gly	Leu	Pro	Ser	Ala	Ile	Gly	Ala	
	530					535					540					
Lys	Ile	Ala	Lys	Pro	Asp	Ala	Ile	Val	Ile	Asp	Ile	Asp	Gly	Asp	Ala	
545					550					555					560	

ES 2 688 953 T3

Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 Lys Asn Trp Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605
 Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620
 Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640
 Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655
 Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670
 Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680

5 <210> 23
 <211> 7847
 <212> ADN
 <213> Schizochytrium sp.

10 <220>
 <221> ceb_transcripción
 <222> (1)..(1259)
 <223>

15 <220>
 <221> CDS
 <222> (1260)..(3314)
 <223>

20 <220>
 <221> ceb_transcripción
 <222> (3315)..(4887)
 <223>

25 <220>
 <221> mutación
 <222> (1834)..(1835)
 <223>

30 <220>
 <221> mutación
 <222> (3042)..(3044)
 <223>

<400> 23

35 ttgtogacag caacttgcaa gttatacgcg accaccaggc aatctcagca cgcccagcga 60
 gcacggagct tgccaagagg gtttacaagt cgtogttcat tcgctctcaa gctttgcctc 120
 aacgcaacta ggcccaggcc tactttcaact gtgtcttgc ttgcctttca caccgacoga 180
 gtgtgcacaa ccgtgttttg cacaaaagcgc aagatgctca ctcgactgtg aagaaaggtt 240

ES 2 688 953 T3

```

gcgcgcaagc gactgcgact gogaggatga ggatgactgg cagcctgttc aaaaactgaa 300
aatccgcgat gggtcagtgat cattcgcgca tgacgcctgc gagagacaag ttaactcgtg 360
tcaactggcat gtcctagcat ctttacgcga gcaaaattca atcgctttat tttttcagtt 420
tcgtaacctt ctgcgaaccg cgaatcgccg tttcagcctg actaatctgc agctgcgtgg 480
cactgtcagt cagtcaagtca gtogtgccgc ctgttccagc accgaggtcg cgcgtcgccg 540
cgcctggacc gctgctgcta ctgctagtgg cacggcaggt aggagcttgt tgccggaaca 600
ccagcagccg ccagtcgacg ccagccaggg gaaagtcggg cgtcgaaggg agaggaaggc 660
ggcgtgtgca aactaacggt gaccactggc gcccgccgac acgagcagga agcaggcagc 720
tgcagagcgc agcgcgcaag tgcagaatgc gogaagatc cacttgccgc cggcgggggc 780
gcacttgccg gcgcggcgcg gaacagtgcg gaaaggagcg gtgcagacgg cgcgcagtga 840
cagtgggcgc aaagccgcgc agtaagcagc ggccgggaac ggtatacgc gtgcccggg 900
ccgcgcgaca cagaagtata cgcgggccga agtggggcgt cgcgcgcggg aagtgcggaa 960
tggcgggcaa ggaaaggagg agacggaaag agggcgggaa agagagagag agagagtgaa 1020
aaaagaaaga aagaaagaaa gaaagaaaga aagctcggag ccacgccgcg gggagagaga 1080
gaaatgaaag cacggcacgg caaagcaaag caaagcagac ccagccagac ccagccgagg 1140
gaggagcgc cgcaggacc cgcgcggcgc cgagcgagca cggcgcgca cgcagcgcgc 1200
gagcgcgcgc gcgagcgagc aaggcttgcg gcgagcgatc gagcgcgca gcgggaagg 1259
atg agc gcg acc cgc gcg gcg acg agg aca gcg gcg gcg ctg tcc tcg 1307
Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1 5 10 15
gcg ctg acg acg cct gta aag cag cag cag cag cag ctg cgc gta 1355
Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Leu Arg Val
20 25 30
ggc gcg gcg tcg gca cgg ctg gcg gcc gcg gcg ttc tcg tcc gcc acg 1403
Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr
35 40 45
ggc gga gac gcg gcc aag aag gcg gcc gcg gcg agg gcg ttc tcc acg 1451
Gly Gly Asp Ala Ala Lys Lys Ala Ala Ala Ala Arg Ala Phe Ser Thr
50 55 60
gga cgc gcc ccc aac gcg aca cgc gag aag agc tcg ctg gcc acg gtc 1499
Gly Arg Gly Pro Asn Ala Thr Arg Glu Lys Ser Ser Leu Ala Thr Val
65 70 75 80
cag gcg gcg acg gac gat gcg cgc ttc gtc gcc ctg acc gcc gcc caa 1547
Gln Ala Ala Thr Asp Ala Arg Phe Val Gly Leu Thr Gly Ala Gln
85 90 95
atc ttt cat gag ctg atg cgc gag cac cag gtg gac acc atc ttt gcc 1595
Ile Phe His Glu Leu Met Arg Glu His Gln Val Asp Thr Ile Phe Gly
100 105 110
tac cct gcc gcc gcc att ctg ccc gtt ttt gat gcc att ttt gag agt 1643
Tyr Pro Gly Gly Ala Ile Leu Pro Val Phe Asp Ala Ile Phe Glu Ser
115 120 125
gac gcc ttc aag ttc att ctg gct cgc cac gag cag gcc gcc gcc cac 1691
Asp Ala Phe Lys Phe Ile Leu Ala Arg His Glu Gln Gly Ala Gly His
130 135 140
atg gcc gag gcc tac gcg cgc gcc acg gcc aag ccc gcc gtt gtc ctg 1739
Met Ala Glu Gly Tyr Ala Arg Ala Thr Gly Lys Pro Gly Val Val Leu
145 150 155 160
gtc acc tcg gcc cct gga gcc acc aac acc atc acc ccg atc atg gat 1787
Val Thr Ser Gly Pro Gly Ala Thr Asn Thr Ile Thr Pro Ile Met Asp
165 170 175
gct tac atg gac ggt acg ccg ctg ctg gtg ttc acc gcc cag gtg cag 1835
Ala Tyr Met Asp Gly Thr Pro Leu Leu Val Phe Thr Gly Gln Val Gln
180 185 190
acc tct gct gtc gcc acg gac gct ttc cag gag tgt gac att gtt gcc 1883
Thr Ser Ala Val Gly Thr Asp Ala Phe Gln Glu Cys Asp Ile Val Gly
195 200 205
atc agc cgc gcg tgc acc aag tgg aac gtc atg gtc aag gac gtg aag 1931
Ile Ser Arg Ala Cys Thr Lys Trp Asn Val Met Val Lys Asp Val Lys
210 215 220

```

ES 2 688 953 T3

gag ctc ccg cgc cgc atc aat gag gcc ttt gag att gcc atg agc ggc	1979
Glu Leu Pro Arg Arg Ile Asn Glu Ala Phe Glu Ile Ala Met Ser Gly	
225 230 235 240	
cgc ccg ggt ccc gtg ctc gtc gat ctt cct aag gat gtg acc gcc gtt	2027
Arg Pro Gly Pro Val Leu Val Asp Leu Pro Lys Asp Val Thr Ala Val	
245 250 255	
gag ctc aag gaa atg ccc gac agc tcc ccc cag gtt gct gtg cgc cag	2075
Glu Leu Lys Glu Met Pro Asp Ser Ser Pro Gln Val Ala Val Arg Gln	
260 265 270	
aag caa aag gtc gag ctt ttc cac aag gag cgc att ggc gct cct ggc	2123
Lys Gln Lys Val Glu Leu Phe His Lys Glu Arg Ile Gly Ala Pro Gly	
275 280 285	
acg gcc gac ttc aag ctc att gcc gag atg atc aac cgt gcg gag cga	2171
Thr Ala Asp Phe Lys Leu Ile Ala Glu Met Ile Asn Arg Ala Glu Arg	
290 295 300	
ccc gtc atc tat gct ggc cag ggt gtc atg cag agc ccg ttg aat ggc	2219
Pro Val Ile Tyr Ala Gly Gln Gly Val Met Gln Ser Pro Leu Asn Gly	
305 310 315 320	
ccg gct gtg ctc aag gag ttc gcg gag aag gcc aac att ccc gtg acc	2267
Pro Ala Val Leu Lys Glu Phe Ala Glu Lys Ala Asn Ile Pro Val Thr	
325 330 335	
acc acc atg cag ggt ctc ggc ggc ttt gac gag cgt agt ccc ctc tcc	2315
Thr Thr Met Gln Gly Leu Gly Gly Phe Asp Glu Arg Ser Pro Leu Ser	
340 345 350	
ctc aag atg ctc ggc atg cac ggc tct gcc tac gcc aac tac tcg atg	2363
Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met	
355 360 365	
cag aac gcc gat ctt atc ctg gcg ctc ggt gcc cgc ttt gat gat cgt	2411
Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg	
370 375 380	
gtg acg ggc cgc gtt gac gcc ttt gct ccg gag gct cgc cgt gcc gag	2459
Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu	
385 390 395 400	
cgc gag ggc cgc ggt ggc atc gtt cac ttt gag att tcc ccc aag aac	2507
Arg Glu Gly Arg Gly Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn	
405 410 415	
ctc cac aag gtc gtc cag ccc acc gtc gcg gtc ctc ggc gac gtg gtc	2555
Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val	
420 425 430	
gag aac ctc gcc aac gtc acg ccc cac gtg cag cgc cag gag cgc gag	2603
Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu	
435 440 445	
ccg tgg ttt gcg cag atc gcc gat tgg aag gag aag cac cct ttt ctg	2651
Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu	
450 455 460	
ctc gag tct gtt gat tcg gac gac aag gtt ctc aag ccg cag cag gtc	2699
Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val	
465 470 475 480	
ctc acg gag ctt aac aag cag att ctc gag att cag gag aag gac gcc	2747
Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala	
485 490 495	
gac cag gag gtc tac atc acc acg ggc gtc gga agc cac cag atg cag	2795
Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln	
500 505 510	
gca gcg cag ttc ctt acc tgg acc aag ccg cgc cag tgg atc tcc tcg	2843
Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser	
515 520 525	
ggg ggc gcc ggc act atg ggc tac ggc ctt ccc tcg gcc att ggc gcc	2891
Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala	

530	535	540	
aag att gcc aag ccc gat gct att gtt att gac atc gat ggt gat gct			2939
Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala			
545	550	555	560
tct tat tcg atg acc ggt atg gaa ttg atc aca gca gcc gaa ttc aag			2987
Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys			
	565	570	575
gtt ggc gtg aag att ctt ctt ttg cag aac aac ttt cag ggc atg gtc			3035
Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val			
	580	585	590
aag aac gtt cag gat ctc ttt tac gac aag cgc tac tcg ggc acc gcc			3083
Lys Asn Val Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala			
	595	600	605
atg ttc aac cag cgc ttc gac aag gtc gcc gat gcg atg cgt gcc aag			3131
Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys			
	610	615	620
ggt ctc tac tgc gcg aaa cag tcg gag ctc aag gac aag atc aag gag			3179
Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu			
	625	630	635
ttt ctc gag tac gat gag ggt ccc gtc ctc ctc gag gtt ttc gtg gac			3227
Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp			
	645	650	655
aag gac acg ctc gtc ttg ccc atg gtc ccc gct ggc ttt cag ctc cac			3275
Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His			
	660	665	670
gag atg gtc ctc gag cct cct aag ccc aag gac gcc taa gttctttttt			3324
Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala			
	675	680	
ccatggcggg cgagcgagcg agcgcgcgag cgcgcgaagt cgcaagcgcc ttgccttget			3384
ttgcttcgct tcgctttgct ttgcttcaca caacctaat atgaattcaa gttttcttgc			3444
ttgtcggcga tgcctgcctg ccaaccagcc agccatcccg ccggccgtcc ttgacgcctt			3504
cgcttcgggc ggggccatcg attcaattca cccatccgat acgttccgcc cccacagctc			3564
cgtctgogca cgacccctgc acgaccacgc caaggccaac gcgccgctca gctcagcttg			3624
tcgacgagtc gcacgtcaca tatctcagat gcatttggac tctgagtggt attatgccac			3684
tagcacgcaa cgatcttogg ggtcctcget cattgcatcc gttcggggccc tgcaggcgtg			3744
gacgcgagtc gccgcccaga cgtctcagca ggccgctccg acgcgagggc tcgagctcgc			3804
cgcgcccggc cgtatgtctgc ctggcgcga ctgatctctg gaggcgcaagg aagacacggc			3864
gacgcgagga ggaccgaaga gagacgctgg gttatgcagg atatacccg ggccggacat			3924
tcgttcgca tacaactccc cattcagact tctctgctct tggcagagcc gagcgcgaac			3984
ggtccgaac gcggcaagga ttttgctct ggtgggtgga ctccgatcga ggccgaggtt			4044
ctcgcaggt tctcgcaggc cggcagtggt cgttagaaat agggagtgcc ggagtcttga			4104
cgcgccttag ctcaactctc gccacgcgc gcacgcgcgc catgccgcgc tcccgtctgt			4164
cgtcgcgctg gccgcgaccg gctgcgccag agtacgacag tgggacagag ctcgaggcga			4224
cgcgaatcgc togggttcta agggtttcaa ggtcggggcg tcgtcgcgtg ccaaagtga			4284
aatagtaggg gggggggggg gtacccaccc cgggcaggtt ctccctcgcca gcctaagtgc			4344
ctaagggagc gtaggggttt cgttgaccag agaagcggag aacctgccgc ggccgggaga			4404
acctatcggc ggagaacttg ccaggcgcga ggcagttctc caatttgccg acagcggcgc			4464
gcccacgcga gccggccgcg tggcgataca gcgagggcag ccgcgggggc cgcgtggcga			4524
cacagctgcg ccgggagtcg gctgcgagaa ggcttctcgc tggcttgggt ggggtcgcg			4584
gtggcagggg atggatgccc aggtacgtcg gcgtgcgcgc gccagggag aaaaggacag			4644
acgcgcgggc ctgcgatgcg agcacgcgat gcgagcaagc gatgcgagca cgcgatgcga			4704
gcacgcgagc gagcgcgccg gcaaatgcc cggaaacagc gttttttgtt tggtgatttc			4764
tatgatgcg gggagacttc gatggcgcga aggggtgcaa ggccaaaaga tgctgacagc			4824
ttcgatcggc ctacggcgcg agcaggaaag ggagcaaggg gcggaattct tctgccttga			4884
cccgggggat ccaactagttc tagagcggcc gccaccgcgg tggagctcca attcgcctta			4944
tagtgagtcg tattacgcgc gctcactggc cgtcgtttta caacgtcgtg actgggaaa			5004
ccctggcggt acccaactta atcgccttgc agcagatccc cctttcgcca cctggcgtaa			5064
tagcgaagag gccgcaccg atcgccttc ccaacagttg cgcagcctga atggcgaatg			5124
ggacgcgccc tgtagcggcg catlaagcgc ggcgggtgtg gtggttacgc gcagcgtgac			5184

```

cgctacactt gccagcgccc tagcgccgc tcccttcgct ttcttccctt cctttctcgc 5244
cacgttgcgc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 5304
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatgggt cacgtagtgg 5364
gccatcgccc tgatagacgg tttttogccc tttgacgttg gagtccacgt tctttaatag 5424
tggactcttg ttccaaactg gaacaacact caacocctatc tcggtctatt cttttgattt 5484
ataagggatt ttgccgattt oggocctattg gttaaaaaat gagctgattt aacaaaaatt 5544
taacgcgaat tttacaaaaa tattaacgct tacaatttag gtggcacttt tcggggaat 5604
gtgocgggaa cccctatttg tttatttttc taatacatt caaataatgta tccgctcatg 5664
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa 5724
catttccgtg tcgcccttat tccctttttt ggggcatttt gccttccctg ttttgctcac 5784
ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcaog agtgggttac 5844
atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt 5904
ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgaogcc 5964
gggcaagagc aactcggctg ccgcatacac tattctcaga atgacttggg tgagtactca 6024
ccagtcacag aaaagcactc tacggatggc atgacagtaa gagaattatg gaggctgoc 6084
ataacatga gtgataacac tgcggccaac ttactttctga caacgatcgg aggaccgaag 6144
gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa 6204
cgggagctga atgaagccat accaaaacgac gagcgtgaca ccacgatgcc tgtagcaatg 6264
gcaacaacgt tgcgcaact attaactggo gaactactta ctctagcttc ccggcaacaa 6324
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg 6384
gotggctggt ttattgctga taaatctgga gcgggtgagc gtgggtctcg oggtatcatt 6444
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 6504
caggcaacta tggatgaacg aaatagacag atcgcctgaga taggtgcctc actgattaag 6564
catttgtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 6624
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 6684
taacgtgagt tttogttcca ctgagcgtca gaccocgtag aaaagatcaa aggatcttct 6744
tgagatcctt tttttctgcg cgtaatotgc tgcttgcaaa caaaaaaac accgctacca 6804
goggtggttt gtttgccgga tcaagagcta ccaactcttt tccgaagggt aactggcttc 6864
agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc 6924
aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct 6984
gccagtgccg ataagtctgt tottaocggg ttggactcaa gacgatagtt accggataag 7044
gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttggg gcgaacgacc 7104
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg 7164
agaaaggcgg acaggtatcc ggtaagcggc agggctggaa caggagagcg cacgagggag 7224
cttccagggg gaaacgcctg gtatctttat agtccctgct ggtttgcca cctctgactt 7284
gagcgtcgat ttttgatg ctogtcaggg gggcgagcc tatggaaaaa cgcagcaac 7344
goggcctttt tacggttcct ggccttttgc tggccttttg ctacatggt ctttccctg 7404
ttatccctg atctgtgga taaccgtatt acegccttg agtgagctga taccgctcgc 7464
cgcaaccgaa gcaccgagcg cagcagctca gtgagcagg aagcgggaaga gcgcccaata 7524
cgcaaacgc ctctcccg cgtttggcgg attcattaat gcagctggca cgcaggttt 7584
cccactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag 7644
gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgg 7704
taacaatttc acaacggaaa cagctatgac catgattacg ccaagcgcgc aattaaccct 7764
cactaaaggg aacaaaagct gggtaaccgg cccccctcg aggtcgacgg tatcgataag 7824
cttgacgtcg aattcctgca gcc 7847

```

<210> 24
 <211> 684
 <212> PRT
 <213> Schizochytrium sp.

<400> 24

```

Met Ser Ala Thr Arg Ala Ala Thr Arg Thr Ala Ala Ala Leu Ser Ser
1           5           10           15
Ala Leu Thr Thr Pro Val Lys Gln Gln Gln Gln Gln Gln Leu Arg Val
20           25           30
Gly Ala Ala Ser Ala Arg Leu Ala Ala Ala Ala Phe Ser Ser Gly Thr

```

5

10

ES 2 688 953 T3

Leu Lys Met Leu Gly Met His Gly Ser Ala Tyr Ala Asn Tyr Ser Met
 355 360 365
 Gln Asn Ala Asp Leu Ile Leu Ala Leu Gly Ala Arg Phe Asp Asp Arg
 370 375 380
 Val Thr Gly Arg Val Asp Ala Phe Ala Pro Glu Ala Arg Arg Ala Glu
 385 390 395 400
 Arg Glu Gly Arg Gly Gly Ile Val His Phe Glu Ile Ser Pro Lys Asn
 405 410 415
 Leu His Lys Val Val Gln Pro Thr Val Ala Val Leu Gly Asp Val Val
 420 425 430
 Glu Asn Leu Ala Asn Val Thr Pro His Val Gln Arg Gln Glu Arg Glu
 435 440 445
 Pro Trp Phe Ala Gln Ile Ala Asp Trp Lys Glu Lys His Pro Phe Leu
 450 455 460
 Leu Glu Ser Val Asp Ser Asp Asp Lys Val Leu Lys Pro Gln Gln Val
 465 470 475 480
 Leu Thr Glu Leu Asn Lys Gln Ile Leu Glu Ile Gln Glu Lys Asp Ala
 485 490 495
 Asp Gln Glu Val Tyr Ile Thr Thr Gly Val Gly Ser His Gln Met Gln
 500 505 510
 Ala Ala Gln Phe Leu Thr Trp Thr Lys Pro Arg Gln Trp Ile Ser Ser
 515 520 525
 Gly Gly Ala Gly Thr Met Gly Tyr Gly Leu Pro Ser Ala Ile Gly Ala
 530 535 540
 Lys Ile Ala Lys Pro Asp Ala Ile Val Ile Asp Ile Asp Gly Asp Ala
 545 550 555 560
 Ser Tyr Ser Met Thr Gly Met Glu Leu Ile Thr Ala Ala Glu Phe Lys
 565 570 575
 Val Gly Val Lys Ile Leu Leu Leu Gln Asn Asn Phe Gln Gly Met Val
 580 585 590
 Lys Asn Val Gln Asp Leu Phe Tyr Asp Lys Arg Tyr Ser Gly Thr Ala
 595 600 605
 Met Phe Asn Pro Arg Phe Asp Lys Val Ala Asp Ala Met Arg Ala Lys
 610 615 620
 Gly Leu Tyr Cys Ala Lys Gln Ser Glu Leu Lys Asp Lys Ile Lys Glu
 625 630 635 640
 Phe Leu Glu Tyr Asp Glu Gly Pro Val Leu Leu Glu Val Phe Val Asp
 645 650 655
 Lys Asp Thr Leu Val Leu Pro Met Val Pro Ala Gly Phe Pro Leu His
 660 665 670
 Glu Met Val Leu Glu Pro Pro Lys Pro Lys Asp Ala
 675 680

ES 2 688 953 T3

<210> 25
 <211> 18
 <212> ADN
 <213> Schizochytrium sp.
 5
 <400> 25
 gttgaccagt gccgttc 18

 <210> 26
 <211> 18
 <212> ADN
 <213> Schizochytrium sp.
 10
 <400> 26
 cgaagtgcac gcagttgc 18
 15
 <210> 27
 <211> 31
 <212> ADN
 <213> Schizochytrium sp.
 20
 <400> 27
 gcgccatgg gacgtcaggt ggcacttttc g 31
 25
 <210> 28
 <211> 36
 <212> ADN
 <213> Schizochytrium sp.
 30
 <400> 28
 gcgccatgg ccgcggaag cagcagatta cgcgca 36

 <210> 29
 <211> 1263
 <212> ADN
 <213> Caenorhabditis elegans
 35
 <220>
 <221> CDS
 <222> (23)..(1228)
 <223>
 40
 <400> 29

 gcggccgcga attcagatct cc atg gtc gct cac tcg tcg gag ggt ctc tcg 52
 Met Val Ala His Ser Ser Glu Gly Leu Ser
 1 5 10
 gcc acc gcc ccg gtc acc ggc ggc gac gtc ctc gtc gac gcc cgc gcc 100
 Ala Thr Ala Pro Val Thr Gly Gly Asp Val Leu Val Asp Ala Arg Ala
 15 20 25
 tcg ctc gag gag aag gag gcc ccg cgc gac gtc aac gcc aac acc aag 148
 Ser Leu Glu Glu Lys Glu Ala Pro Arg Asp Val Asn Ala Asn Thr Lys
 30 35 40
 cag gcc acc acc gag gag ccc cgc atc cag ctg ccc acc gtc gac gcc 196
 Gln Ala Thr Thr Glu Glu Pro Arg Ile Gln Leu Pro Thr Val Asp Ala
 45 50 55
 ttc cgc cgc gcc atc ccc gcc cac tgc ttc gag cgc gac ctc gtc aag 244
 45

ES 2 688 953 T3

Phe	Arg	Arg	Ala	Ile	Pro	Ala	His	Cys	Phe	Glu	Arg	Asp	Leu	Val	Lys	
	60					65					70					
tcg	atc	cgc	tac	ctc	gtc	cag	gac	ttc	gcc	gcc	ctc	acc	atc	ctc	tac	292
Ser	Ile	Arg	Tyr	Leu	Val	Gln	Asp	Phe	Ala	Ala	Leu	Thr	Ile	Leu	Tyr	
75				80					85					90		
ttc	gcc	ctc	ccc	gcc	ttc	gag	tac	ttc	ggc	ctc	ttc	ggc	tac	ctc	gtc	340
Phe	Ala	Leu	Pro	Ala	Phe	Glu	Tyr	Phe	Gly	Leu	Phe	Gly	Tyr	Leu	Val	
			95						100					105		
tgg	aac	atc	ttc	atg	ggc	gtc	ttc	ggc	ttc	gcc	ctc	ttc	gtc	gtc	ggc	388
Trp	Asn	Ile	Phe	Met	Gly	Val	Phe	Gly	Phe	Ala	Leu	Phe	Val	Val	Gly	
		110						115					120			
cac	gac	tgc	ctc	cac	gga	agc	ttc	tcg	gac	aac	cag	aac	ctc	aac	gac	436
His	Asp	Cys	Leu	His	Gly	Ser	Phe	Ser	Asp	Asn	Gln	Asn	Leu	Asn	Asp	
		125					130					135				
ttc	atc	ggc	cac	atc	gcc	ttc	tcg	ccc	ctc	ttc	tcg	ccc	tac	ttc	ccc	484
Phe	Ile	Gly	His	Ile	Ala	Phe	Ser	Pro	Leu	Phe	Ser	Pro	Tyr	Phe	Pro	
	140					145						150				
tgg	cag	aag	tcg	cac	aag	ctc	cac	cac	gcc	ttc	acc	aac	cac	atc	gac	532
Trp	Gln	Lys	Ser	His	Lys	Leu	His	His	Ala	Phe	Thr	Asn	His	Ile	Asp	
155					160					165					170	
aag	gac	cac	ggc	cac	gtc	tgg	atc	cag	gac	aag	gac	tgg	gag	gcc	atg	580
Lys	Asp	His	Gly	His	Val	Trp	Ile	Gln	Asp	Lys	Asp	Trp	Glu	Ala	Met	
			175						180					185		
ccc	tcg	tgg	aag	cgc	tgg	ttc	aac	ccc	atc	ccc	ttc	tcg	ggc	tgg	ctc	628
Pro	Ser	Trp	Lys	Arg	Trp	Phe	Asn	Pro	Ile	Pro	Phe	Ser	Gly	Trp	Leu	
			190					195					200			
aag	tgg	ttc	ccc	gtc	tac	acc	ctc	ttc	ggc	ttc	tgc	gac	ggc	tcg	cac	676
Lys	Trp	Phe	Pro	Val	Tyr	Thr	Leu	Phe	Gly	Phe	Cys	Asp	Gly	Ser	His	
		205					210					215				
ttc	tgg	ccc	tac	tcg	tcg	ctc	ttc	gtc	cgc	aac	tcg	gac	cgc	gtc	cag	724
Phe	Trp	Pro	Tyr	Ser	Ser	Leu	Phe	Val	Arg	Asn	Ser	Asp	Arg	Val	Gln	
	220					225						230				
tgc	gtg	atc	agc	ggc	atc	tgc	tgc	tgc	gtc	tgc	gcc	tac	atc	gcc	ctc	772
Cys	Val	Ile	Ser	Gly	Ile	Cys	Cys	Cys	Val	Cys	Ala	Tyr	Ile	Ala	Leu	
235					240					245				250		
acc	atc	gcc	ggc	tcg	tac	tcg	aac	tgg	ttc	tgg	tac	tac	tgg	gtc	ccg	820
Thr	Ile	Ala	Gly	Ser	Tyr	Ser	Asn	Trp	Phe	Trp	Tyr	Tyr	Trp	Val	Pro	
			255						260					265		
ctc	tcg	ttc	ttc	ggc	ctc	atg	ctc	gtc	atc	gtc	acc	tac	ctg	cag	cac	868
Leu	Ser	Phe	Phe	Gly	Leu	Met	Leu	Val	Ile	Val	Thr	Tyr	Leu	Gln	His	
			270					275					280			
gtc	gac	gac	gtc	gcc	gag	gtc	tac	gag	gcc	gac	gag	tgg	tcg	ttc	gtc	916
Val	Asp	Asp	Val	Ala	Glu	Val	Tyr	Glu	Ala	Asp	Glu	Trp	Ser	Phe	Val	
			285				290						295			
cgc	ggc	cag	acc	cag	acc	atc	gac	cgc	tac	tac	ggc	ctc	ggc	ctc	gac	964
Arg	Gly	Gln	Thr	Gln	Thr	Ile	Asp	Arg	Tyr	Tyr	Gly	Leu	Gly	Leu	Asp	
	300					305					310					
acc	acc	atg	cac	cac	atc	acc	gac	ggc	cac	gtg	gcc	cac	cac	ttc	ttc	1012
Thr	Thr	Met	His	His	Ile	Thr	Asp	Gly	His	Val	Ala	His	His	Phe	Phe	
					320					325				330		
aac	aag	atc	ccg	cac	tac	cac	ctc	atc	gag	gcc	acc	gag	ggc	gtc	aag	1060
Asn	Lys	Ile	Pro	His	Tyr	His	Leu	Ile	Glu	Ala	Thr	Glu	Gly	Val	Lys	
					335				340					345		
aag	gtc	ctc	gag	ccc	ctc	tcg	gac	acc	cag	tac	ggc	tac	aag	tcg	cag	1108
Lys	Val	Leu	Glu	Pro	Leu	Ser	Asp	Thr	Gln	Tyr	Gly	Tyr	Lys	Ser	Gln	
			350					355					360			
gtc	aac	tac	gac	ttc	ttc	gcc	cgc	ttc	ctc	tgg	ttc	aac	tac	aag	ctc	1156
Val	Asn	Tyr	Asp	Phe	Phe	Ala	Arg	Phe	Leu	Trp	Phe	Asn	Tyr	Lys	Leu	
		365					370					375				

ES 2 688 953 T3

```

gac tac ctc gtg cac aag acc gcc ggc atc atg cag ttc cgc acc acc      1204
Asp Tyr Leu Val His Lys Thr Ala Gly Ile Met Gln Phe Arg Thr Thr
380                               385                               390
ctc gag gag aag gcc aag gcc aag taaccocgggg gtacccttaa ggcacgcgcg      1258
Leu Glu Glu Lys Ala Lys Ala Lys
395                               400
gccgc                                                                    1263

```

<210> 30
 <211> 402
 <212> PRT
 <213> Caenorhabditis elegans

5

<400> 30

```

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

```

10

ES 2 688 953 T3

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

5 <210> 31
 <211> 570
 <212> ADN
 <213> Caenorhabditis elegans

<400> 31
 agcagcagat tgcccgaggc cggcggaagg gacgaggccc aggcggctcg tgaagcgca 60
 ttccgaagg cgggctcggc gacgacgcc ggcggcgac gacggccctg ccggaccggg 120
 cctggggtgg acggcgaggc taactaggac ttggggaagc cgagctgagc gacttgagcg 180
 ggttgagggg acgaactggt taggcgcggc cgagtcgtca gagccagcct gtggagaaag 240
 aggcgcggcc gagtgcgacg gggaaagctg cgcggacctc gcattgcacc gcatcgcawt 300
 cgcaccgcaw tcgcaccgca ccgcatcgca ccgcatcgca tcgagaccg acgcagcgag 360
 acgcgacgct gggccttccc ggcgaaaaaa agtgatctgg cttacaaatc ccgagacgag 420
 acagacgtcg gcagcagaaa cgaatcagtc gacgacgagc tgcagcagca gcagcagcag 480
 cagcagccca tcgcgagcaa gggctcagcc agcagaacac caatcaggcc aagaatcgca 540
 10 cgggaagcaag ccttgacatc ctttgcacaac 570

15 <210> 32
 <211> 18
 <212> ADN
 <213> Schizochytrium sp.

<400> 32
 gaccgctcat ctatgctg 18

20 <210> 33
 <211> 18
 <212> ADN
 <213> Schizochytrium sp.

25

ES 2 688 953 T3

<400> 33
ctcaaagtga acgatgcc 18

5
<210> 34
<211> 4244
<212> ADN
<213> Schizochytrium sp.

10
<400> 34

tttctctctc	togagctggt	gctgctgctg	ctgctgctgc	tgcttcocttg	ctggttctca	60
cgtccgttgc	atcaagcgc	cgctcgctcg	accgatcggg	gcgtgcgtgc	gtgcgtgagt	120
cttgttgcca	ggcagccgca	ggctgtctgt	ctgtttgtgt	agttttacc	tcggggttcg	180
gggtctgcct	gcctcccgc	cccgcccgc	gccgcccgt	tccaaccgc	tcgcctccgc	240
ccatcgggcc	tgcctcctc	gcgcgcgacg	catcgcgcgc	atcgcgatga	tcatgctgcc	300
acgcacgggg	ggacgcgcgc	cccgctccc	ccgcgcgcgc	cgctcgtcgt	tggcgatgcc	360
gtcgcgcgcc	tccttcoctc	cctcgcctcc	tcttccctcc	gagccccct	gtcttccctc	420
gccccgcgag	cggcgcgcgag	gaagcgagga	gagcggggag	gagagaagaa	aagaaaagaa	480
aagaaaagaa	aataacagcg	ccgtctcgcg	cagacgcgcg	cggccgcgctg	cgaggcggcg	540
tgatggggct	tctcgtggcg	cggctcggcg	ctggcccggc	ctcgcctttg	aggtgcaggc	600
tttgggagag	aagagtggga	cgcgggagaag	ataagatggg	gccatggcgc	aggacggaga	660
ggttgctgaa	acttcttcga	gcggcacagg	cgatggcgag	agaccgacag	ctgccggcgc	720
ggaggggatg	gatacctccc	gaggctggca	tggacgagct	ggccgcgcgg	atctggctgg	780
ccgcgcggcg	gtgggtccgg	aggcgcgagg	ttggttttct	tcatacctga	taccatacgg	840
tattcattct	tcctctccag	gaaggaagca	agtcaatag	agtatcacta	gcctaataat	900
ggactctatg	ttttaggcca	cgtcggagca	gaaggcgcga	gcgattcgaa	tgcgagcgat	960
agatacagca	cagagacctt	gccggcgacg	cggatgcagg	cgagcacgca	cgcaccgca	1020
gcacggcagc	ggtgcacgcg	ctcctcggca	gatgcacggg	tctgcgcgcg	gcctttacat	1080
tttttgatgt	taggtgggtg	gcctgccact	ttgaacatca	tccacaagtc	aacgcagcat	1140
caagaggcaa	gcaagtacat	acatccattc	gaattcaagt	tcaagagacg	cagcaacagc	1200
cgccgctccg	ctcaagctgc	agctagctgg	ctgacagggc	tcgctggctg	tagtggaaaa	1260
ttccattcac	ttttctgcat	ccgcggccag	caggcccgt	cgcaagttct	ctcgtttgtt	1320
tgttcgttcg	tgcgtgcgtg	cgtgcgtccc	agctgcctgt	ctaactctgc	gcgcgatcca	1380
acgaccctcg	gtcgtgcgcg	caagcgaaac	ccgacgcgga	cctggccaat	gccgcaagaa	1440
tgctaagcgc	gcagcaatgc	tgagagtaat	cttcagccca	ccaagtcatt	atcgcgtccc	1500
aagtctccat	cgcagccaca	ttcaggcttt	ctctctctct	ccctccctct	ctttctgcgg	1560
ggagagaagg	aaagaccgcg	cgccgcggcc	tctgcgcctg	tgacgggctg	tcogttgtaa	1620
gacctcttag	acagttccta	ggtgcggggc	gcgcgcgcgc	ctccgtcgcg	ggcaacagta	1680
ggcggccacg	ggttcccccc	gcacctcca	caocttcttc	ccccgcagcc	ggaccgcgcg	1740
ccgtctgctt	acgcaactcg	cgcggccgcc	gcccgcgaa	ccgagcgcgt	gctgtggggc	1800
ccgtcttcgc	gcccgcgtcg	aggtcgtccc	cgcgcgcgcg	tactccgggt	cctgtgcggg	1860
acgtacttaa	tattaacagt	gggacctcgc	acaggacctg	acggcagcac	agacgtcgc	1920
gcctcgcato	gctggggagc	caggcgcggc	atcccggcgc	ggccccgcac	cggggagggt	1980
gogggggcgc	ctcttcgggc	cggcggccgc	atcaggcggg	tgacgcaaga	gcctcgcag	2040
tcgctcgcct	gcgggagcgc	agcgcgcgcg	cagcgtggcc	aagctcccgc	ccctctggc	2100
tggctgcatg	cctgcctgcc	tgctgcctg	cgtgcgtgcg	tgcgtgcgtg	ccttcgtgcg	2160
tgctgcctt	cgtgcgtgcg	tgctgcgtg	cggcgggaaga	gggatcatgc	gaggatcaat	2220
caccgcgcgc	acctcgactt	ttgaagaagc	cgcgatgcga	tgcgatgcga	tgcgatgcga	2280
cgcgataccg	tgcgaggcta	cgaagcaggt	ctggccggcc	gtcatacaac	gcacgttttc	2340
gagaaggagg	gctggcggag	gcgtgcctgc	cggcgaccat	tgcgaacgcg	gcgtctcgtg	2400
gctggcgaag	gtccctggag	gatctaaocg	tcgctgctat	gatgctatag	ctgtctgat	2460
ccccggtcca	ttccaccacg	tctgtgcctg	ccgcctgacc	tgcgcttggc	tttccctcaa	2520
gttctctctc	gcccggcctt	caggaccgag	acgagacctg	cagctgcagc	tagactcgcg	2580
ctcgcctcgc	gaggattcgc	cggccgcggg	gcccggacggg	actcgcgagg	tcacacggcc	2640
gcccgcgatc	gcgatggctg	tgctgacgta	ctcgtgcgtg	gcagccgtac	gtcagcgcag	2700
ccgcctccgt	attgtggatt	cgttagttgg	ttgttggttg	atttgttgat	taattttttt	2760

ES 2 688 953 T3

gttcgtaggc	ttggttatag	ctaatagttt	agtttatact	ggtgctcttc	ggtgctgatt	2820
tagctcgact	tgggtccaca	ccactgcccc	totactgtga	atggatcaat	ggacgcacga	2880
cgggcgcag	aaagtgcgcg	agtgaggtaa	cctaagcaac	ggcggctctc	agaggggacg	2940
cacgccctcc	gtcgcagtca	gtccagacag	gcagaaaagc	gtcttaggga	ccacgcacgc	3000
acgcacgcac	gcacgcacgc	ccgcacgcac	gctccctccc	tcgcgtgcct	atTTTTtag	3060
gcttcccttc	gcacgggcct	acctctcgct	ccctcgccctc	gccgcaccag	gcggcagcag	3120
cgatacctgc	cggtgcccgc	tccgtcacgc	gctcagccgc	agctcagccc	agccgcgagc	3180
tagggtttgt	tcgtcctgaa	ttgtttgatt	tgatttgatt	tgatttgatc	cgatccgac	3240
cgatctgato	tgatttgctt	tgctttgctt	tgctcccttc	ccggcgcgga	ccaagcgtcc	3300
gtctgcgcgc	cgcagcttcc	cttcttctcc	cagccctcct	tctgctcccg	cctctcgcgc	3360
aagcacgcag	cttcgcgcgc	gcatccggtc	ggtcggctgg	tcgatcgacc	cgctcgcgc	3420
tgctgctgtg	gccgggcttt	tctccatcgg	cgactcttct	ttctccatac	gtcctaactac	3480
gtacatacat	actgccggct	tctcctctct	ccagcgcggc	gacggcggca	ggctgogacg	3540
tcgtgcgcgc	cgcgggcgc	gcgcgcgcgc	ccgcgcgcgc	ccgcgtcgca	ggcctcgtc	3600
gccgcgcgcg	ctccgcctcc	ctccgagggc	gcgagagggc	ccgcgcggcg	cgatggatgg	3660
atggatggat	ggatggatgg	atggattttg	ttgatccgat	gcggcgcgat	ggcggagatg	3720
agcagggacg	agcgcgcgag	cgccgcgcgc	ggattcgcag	ggcctcgtct	gcctcgcgc	3780
cgctgcgcgc	ccgccttgc	gagcctgcgc	cgcgagcgag	cgagcgcgag	agcgggctt	3840
tctttgtctc	gcgcgcgcgc	tgccctcgtg	tgtcttgcgc	ttgcgtagcg	ggcgcgcgcg	3900
tggaagatgg	ctcattcaat	cgaccatttc	acgcacgcac	tccggcgcgc	agagaaggcc	3960
gaggaggagc	agcaagcaaa	ccaaaagctc	tcgogctcgc	ggtctcgggc	tcgagcggtc	4020
tcggagagag	agtcttgccg	cgaccaccgc	cagcagcagc	agcagcagca	gcctgtgca	4080
gcaocgacac	gagcacgcagc	acgagcacga	gcattcgcagc	aagaggacag	acagcgttgt	4140
cagcgcctag	ctcgtctgat	acagaaagag	gcgggttggg	cgtaaaaaaa	aaggagcagc	4200
caagccgccca	gccagccagc	tagctagcca	gcctgcctgc	caaa		4244

<210> 35
 <211> 4244
 <212> ADN
 <213> Schizochytrium sp.

5

<400> 35

ttctctctc	tcgagctggt	gctgctgctg	ctgctgctgc	tgcttctctg	ctggttctca	60
cgctcgttgc	atcaagcgc	cgctcgtcgc	accgatcgg	ggctgctgct	gtgctgagct	120
cttgttgcca	ggcagccgca	ggctgtctgt	ctgtttgtgt	agttttacc	tcggggttcg	180
gggtctgcct	gcctcccgct	ccgcgcgcgc	gcgcgcgcgc	tccaccccg	tcgcctccgc	240
ccatcggggc	tcgctccttc	gcgcgcgcgc	catcgcgcgc	atcgcgatga	tcgatctgcc	300
acgcacgggg	ggcgcgcgc	ccgcgcgcgc	ccgcgcgcgc	cgctcgtcgc	tcggcgtgcc	360
gtcgcgcgc	tccttctctc	cctcgcctcc	tcttctctcc	gagcccccct	gtcttctctc	420
gccccgcag	cggcgcgcag	gaagcgagga	gagcggggag	gagagaagaa	aagaaaagaa	480
aagaaaagaa	aataacagcg	ccgtctcgcg	cagacgcgcg	cgcccgctg	cgaggcggcg	540
tgatggggct	tctcgtggcg	cggtcgcgc	ctggcccggc	ctcgcctttg	aggtgcaggc	600
tttgggagag	aagagtggga	cgcgagaga	ataagatggt	gccaatggcg	aggacggaga	660
ggttgcctgaa	acttcttctga	goggcacagg	cgatggcgag	agaccgacag	ctgcccgcgc	720
ggaggggatg	gatacctccc	gaggtggca	tggaacgagct	ggccgcgcgc	atctggctgg	780
ccgcgcggcg	gtgggtccgg	aggcgcgagg	ttggttttct	tcatacctga	taccatacgg	840
tattcattct	tcctctccag	gaaggaagca	agtcacatag	agtatcacta	gcctaattgat	900
ggactctatg	ttttagggca	cgctcgagca	gaagggcgcga	gcgattcgaa	tcgagcgcgat	960
agatacagca	cagagacctt	gccggcgcgc	cgatgcaggg	cgagcacgca	cgcaaccgac	1020
gcacggcagc	ggtgcacgcg	ctcctcggca	gatgcacggg	tctgcgcgcg	gcctttacat	1080
tttttgattt	taggtgggtg	gcctgccact	ttgaacatca	tccacaagtc	aacgcagcat	1140
caagaggcaa	gcaagtacat	acatccattc	gaattcaagt	tcaagagacg	cagcaacagc	1200
cgccgctccg	ctcaagctgc	agctagctgg	ctgacagggc	tcgctggctg	tagtggaaaa	1260
ttccattcac	ttttctgcat	ccgcggccag	cagggccgta	cgcacgttct	ctcgtttggt	1320
tgctcgttgc	tcgctcgcgc	cgctcgcctc	agctgcctgt	ctaattctgcc	gcgcgatcca	1380
acgaccctcg	gtcgtcgcgc	caagcgaaac	ccgacgcgcga	cctggccaat	gccgcaagaa	1440
tgctaagcgc	gcagcaatgc	tcgagtaaat	cttcagccca	ccaagtcatt	atcgttgcgc	1500
aagtctccat	cgcagccaca	ttcaggcttt	ctctctctct	ccctccctct	cttctcgcgc	1560
ggagagaagg	aaagaccgcg	cgccgcgcgc	tctgcgcctg	tgacgggctg	tcctgtgtaa	1620

10

ES 2 688 953 T3

gcccccttag	acagttccta	ggtgccggggc	gcccgcgcgc	ctccgtcgca	ggcacacgta	1680
ggcggccaacg	ggttcccccc	gcaccttcca	caccttcttc	ccccgcagcc	ggaccgcgcg	1740
ccgtctgctt	acgcacttcg	cggggccgcc	gccccggaac	ccgagcgcgt	gctgtgggcg	1800
ccgtcttccg	gccgcgtcgg	aggctcgtccc	cgcgcgcgc	tactccgggt	cctgtgcggg	1860
acgtacttaa	tattaacagt	gggacctcgc	acaggacctg	acggcagcac	agacgtcgc	1920
gcctcgcac	gotggggacg	caggcgaggc	atccccgcgc	ggccccgcac	cggggaggct	1980
gccccggcgc	ctcttcgggc	cgggggcgc	atcaggcgga	tgacgcaaga	gccctcgcag	2040
tcgctcgc	gccccggcgc	agcgcggcgc	cagcgtggcc	aagctccgc	ccctctggc	2100
tggtcgcacg	octgcctgcc	tgctgcctg	cgtgcgtgcg	tgctgcgtg	ccctcgtgcg	2160
tgctgcctt	cgtgcgtgcg	tgctgagtg	cggcggaaga	gggatcatgc	gaggatcaat	2220
caccgcgcgc	acctcgcactt	ttgaagaagc	cgcgatgcga	tgcatgcga	tgogatgcga	2280
cgcgataccg	tgcgaggcta	cgaagcgagt	ctggccccgc	gtcatacaac	gcacgtttc	2340
gagaaggag	gctggcgggag	gogtgcagtc	cggcgacct	tgcaaacgcg	gcgtctcgtg	2400
tgcgcggaag	gtccctggag	gatctaacga	tcgctgctat	gatgctatag	ctgtcgtgat	2460
ccccggtcca	ttccaccacg	tctgtgcctg	ccgctgacc	tgcccttggc	tttcttcaa	2520
gttctcctcc	gccccggcctt	caggaccgag	acgagacctg	cagctgcagc	tagactcgcg	2580
ctcgcctcgc	gaggatcgc	cggccccgcg	gccccgaggg	actcgcgagg	tcacacggcc	2640
gccccgcgac	gcatggctg	tgctgacgta	ctcgtgcgtg	gcagccgtac	gtcagcgaag	2700
ccgctccgt	attgtggatt	cgttagttg	ttgttggtg	atttgttgat	taattttttt	2760
ttcgttaggc	ttggttatag	ctaatagttt	agtttatact	ggtgctcttc	ggtgctgatt	2820
tagctcgact	tggtccaca	ccactgcccc	tctactgtga	atggatcaat	ggacgcaoga	2880
cggccgacg	aaagtgcgcg	agtgaggtaa	cctaagcaac	ggcggctctc	agaggggaag	2940
cagccctcc	gtcgcagtc	gtccagacag	gcagaaaagc	gtcttaggga	ccacgcacgc	3000
acgcacgcac	gcacgcacgc	ccgcacgcac	gtccccccc	tcgctgcct	atttttttag	3060
gcttccctcc	gcacgggct	acctctcgt	ccctcgcctc	gccccaccag	gccccagcag	3120
cgatacctgc	cgggtgcgcgc	tcgctcagc	gctcagccgc	agctcagccc	agccccgagc	3180
tagggtttgt	tcgtcccgaa	ttgtttgatt	tgatttgatt	tgatttgatc	cgatccgatc	3240
cgatctgac	tgatttgctt	tgctttgctt	tgtctccctc	ccccgcgga	ccaagcgtcc	3300
gtctgcgcgc	cgcagcttcc	cttcttctcc	cagccctcct	tctgctcccg	ctctcgcgc	3360
aagcagcag	cttcgcgcgc	gcctccggtc	ggtcggctcg	tcgatcgacc	cgcctgcgc	3420
tgctgctgtg	gccccgcttt	tctccatcgg	cgactctttc	ttctccatac	gtcctaactac	3480
gtacatacat	actgccccgt	tctctctctt	ccagcgcggc	gccccgggca	ggtcgcgacg	3540
tcgtcgcgc	cgcggggcgc	gcgcgcgcgc	ccccgcgcgc	ccccgcgcgc	ggcctcgtc	3600
gccccgcgc	ctccgctccg	ctccgagccc	gcgagagggc	ccccgcgcgc	cgatggatgg	3660
atggatggat	ggatggatgg	atggatttg	ttgatcgatg	gccccgcag	ggcggagatg	3720
agcagggacg	agcgcgcgag	cgcggcagcc	ggattcgcag	ggcctcgcct	gcctcgcgc	3780
cgctccgcgc	cccccttgc	gagcctgcgc	cgcgagcgag	cgagcagcgc	agccccgctt	3840
tctttgtctc	gccccgcgc	tgccctcgtg	tgtcttctgc	ttgcgtagcg	ggccccgcgc	3900
tggaagatgg	ctcattcaat	cgaccattc	acgcacgcac	tcgggcgcgc	agagaaggcc	3960
gaggaggagc	agcaagcaaa	ccaaaagctc	tcgcgctcgc	ggtctcgggc	tcgagcggtc	4020
tcggagagag	agtcttgcgg	cgaccaccg	cagcagcagc	agcagcagca	gcgctgtcga	4080
gcacagcagc	gagcagcagc	acgagcagc	gcattcgcagc	aagaggacag	acacggttgt	4140
cagcgcctag	ctcgcctcgc	acagaaagag	gccccgttgg	cgtaaaaaaa	aaggagcagc	4200
caagcgcaca	gcccagccagc	tagctagcca	gcctgcctgc	caaa		4244

REIVINDICACIONES

- 5 1. Un método para producción de una proteína, que comprende: cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína, para producir la proteína; y recuperar la proteína.
- 10 2. El método de la reivindicación 1, en el que la molécula de ácido nucleico aislada comprende adicionalmente una o más secuencias de control de transcripción y el gen extraño está unido de forma operativa a la una o más secuencias de control de transcripción.
- 15 3. El método de la reivindicación 2, en el que la una o más secuencias de control de transcripción comprenden una secuencia de control de transcripción de Thraustochytriales seleccionada entre el grupo que consiste en secuencias de promotor de Thraustochytriales, secuencias de amplificación de Thraustochytriales, secuencias de operador de Thraustochytriales, secuencias de represor de Thraustochytriales, y combinaciones de las mismas.
- 20 4. El método de la reivindicación 3, en el que al menos una secuencia de control de transcripción de Thraustochytriales es una secuencia de promotor de Thraustochytriales seleccionada entre el grupo que consiste en una secuencia de promotor de acetolactato sintasa de Thraustochytriales, secuencia de promotor de α -tubulina de Thraustochytriales, secuencia de promotor de sistema de policétido sintasa de Thraustochytriales, y secuencia de promotor de ácido graso desaturasa de Thraustochytriales.
- 25 5. El método de cualquier reivindicación precedente, en el que la molécula de ácido nucleico aislada comprende adicionalmente una secuencia de segmento de señal de ácido nucleico.
- 30 6. El método de cualquier reivindicación precedente, en el que la proteína se secreta en el medio, y en el que la proteína se recupera del medio.
- 35 7. El método de cualquier reivindicación precedente, en el que el microorganismo es de un género seleccionado entre el grupo que consiste en *Thraustochytrium*, *Labyrinthuloides*, *Japonochytrium*, y *Schizochytrium*.
- 40 8. Un método para producción de una proteína, que comprende:
cultivar un microorganismo recombinante del Orden Thraustochytriales en un medio, en el que el microorganismo recombinante comprende una molécula de ácido nucleico aislada que comprende un gen extraño que codifica la proteína, para producir la proteína; y
extraer el microorganismo que comprende la proteína.
9. El método de la reivindicación 8, en el que el microorganismo se extrae como una biomasa.

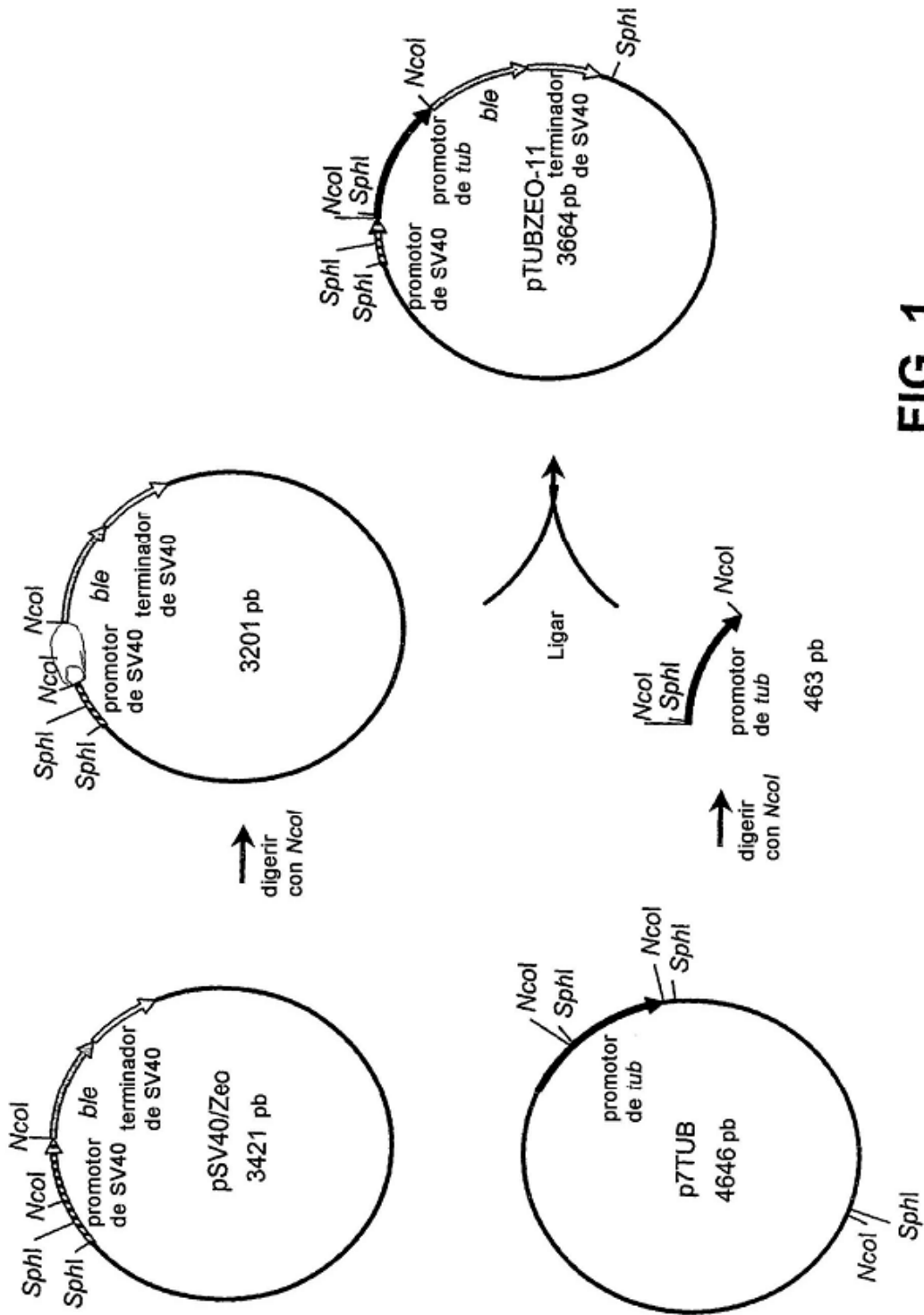


FIG. 1

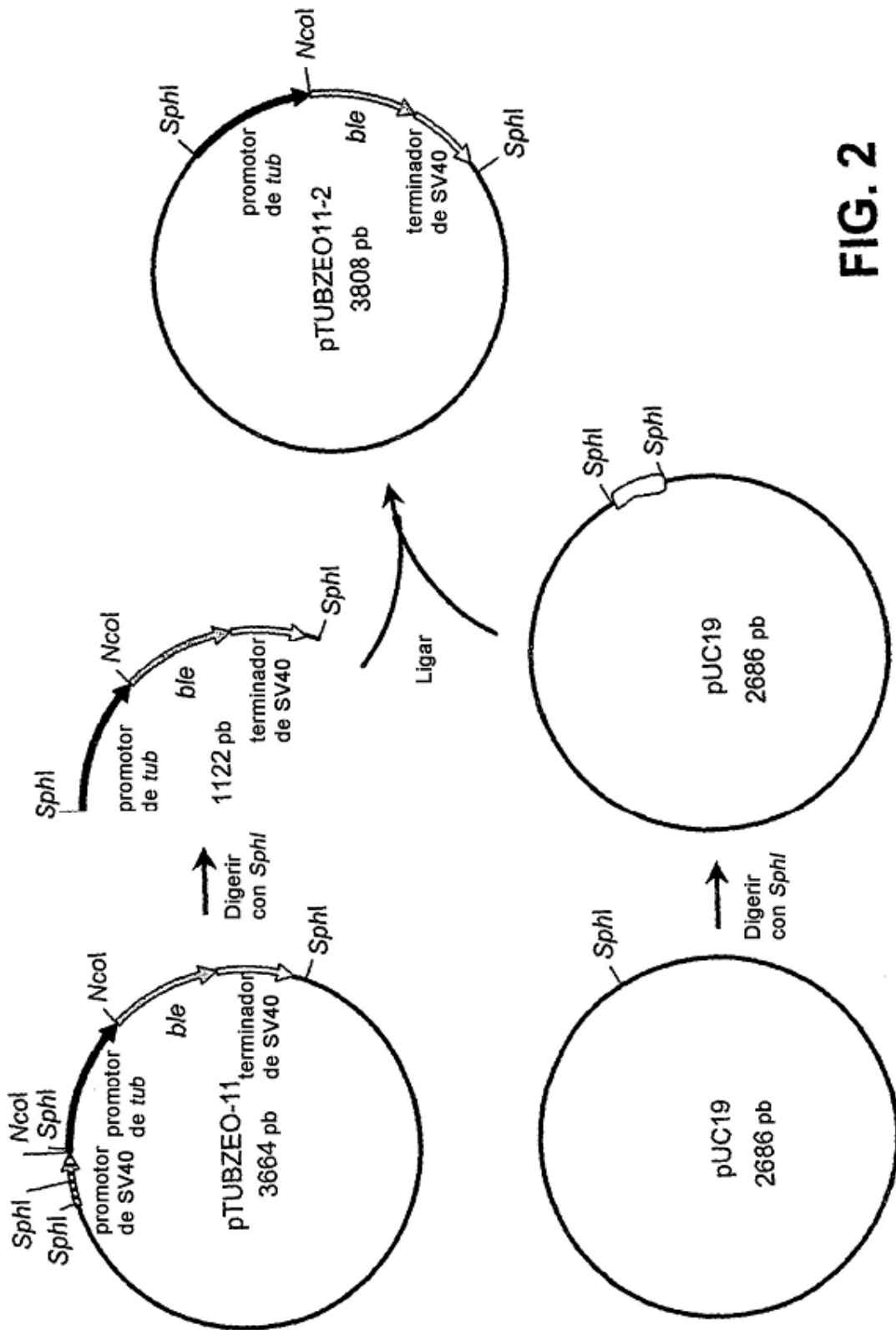


FIG. 2

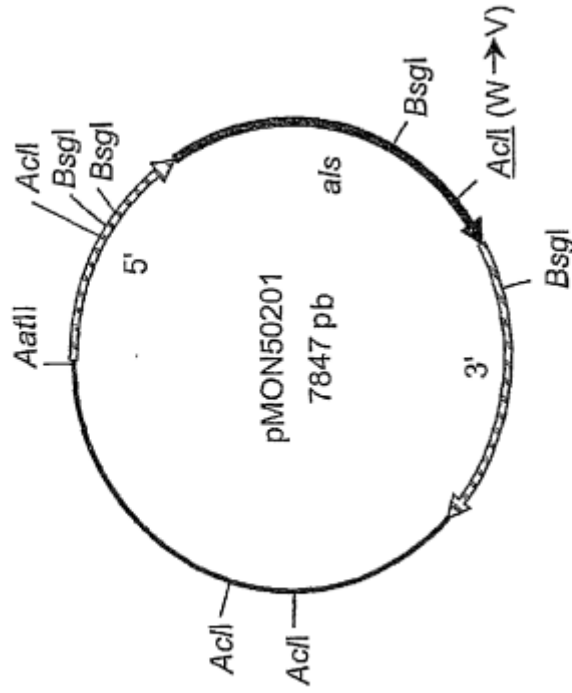


FIG. 3B

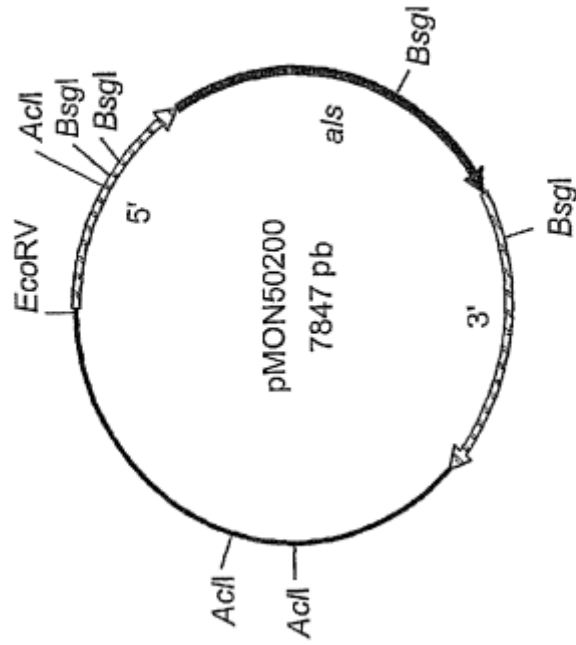


FIG. 3A

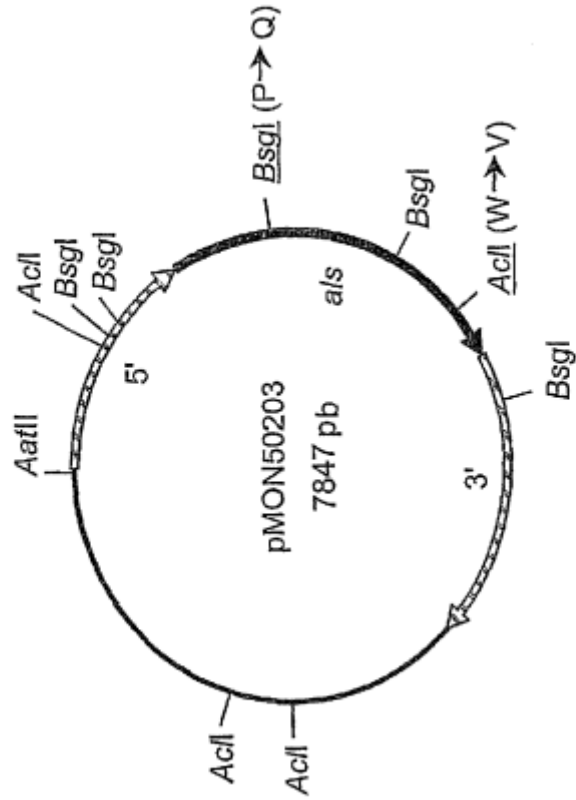


FIG. 3D

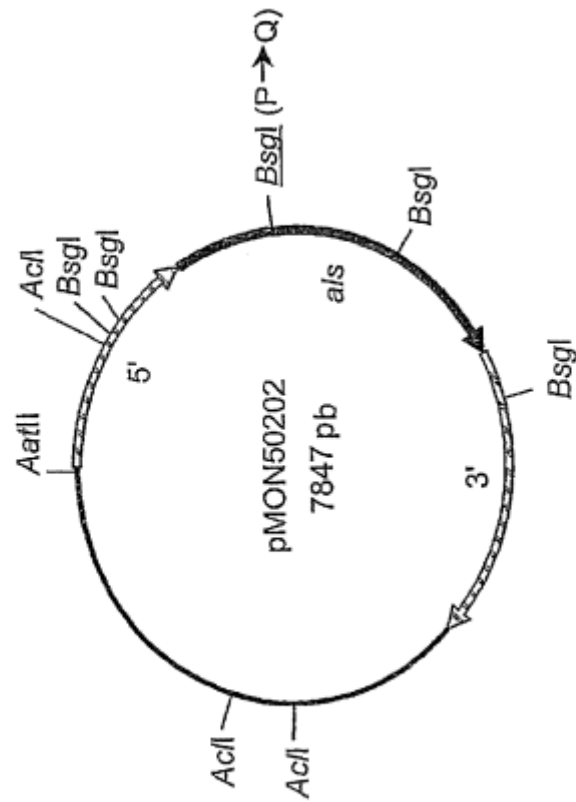


FIG. 3C