

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 724 998**

51 Int. Cl.:

C12N 1/14 (2006.01)

C12N 15/53 (2006.01)

C12N 15/81 (2006.01)

C12P 7/64 (2006.01)

C12R 1/645 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.08.2015** **PCT/SG2015/050273**

87 Fecha y número de publicación internacional: **17.03.2016** **WO16039685**

96 Fecha de presentación y número de la solicitud europea: **21.08.2015** **E 15840158 (8)**

97 Fecha y número de publicación de la concesión europea: **17.04.2019** **EP 3191577**

54 Título: **Métodos para la producción eficiente de ácidos grasos poliinsaturados (PUFA) en especies de Rhodosporidium y Rhodotorula**

30 Prioridad:

08.09.2014 US 201462047300 P

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

18.09.2019

73 Titular/es:

**TEMASEK LIFE SCIENCES LABORATORY
LIMITED (100.0%)
1 Research Link The National University Of
Singapore
Singapore 117604, SG**

72 Inventor/es:

**LIU, YANBIN;
KOH, CHONG MEI y
JI, LIANGHUI**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 724 998 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

Métodos para la producción eficiente de ácidos grasos poliinsaturados (PUFA) en especies de *Rhodosporidium* y *Rhodotorula*.

Referencia cruzada a la solicitud relacionada

La presente solicitud se refiere a y reivindica la prioridad a la solicitud de patente provisional de Estados Unidos núm. de serie 62/047,300 presentada el 8 de septiembre de 2014.

Presentación de secuencias

La presente solicitud se presenta junto con un Listado de Secuencias en formato electrónico. El Listado de Secuencias se titula 2577237PCTSequenceListing.txt, creado el 2 de julio de 2015 y tiene un tamaño de 321 kb. La información en el formato electrónico del Listado de Secuencias se incorpora en la presente descripción como referencia en su totalidad.

Antecedentes de la invención

La presente invención se refiere al campo de la biotecnología de hongos, más en particular a métodos de ingeniería genética para la producción de ácidos grasos poliinsaturados (PUFA) omega-3 en huéspedes fúngicos seleccionados del género *Rhodosporidium* y *Rhodotorula*.

Los ácidos grasos omega-3 (denominados además ácidos grasos ω -3 o ácidos grasos n-3) se refieren a ácido alfa-linolénico (ALA) [ácido (9Z,12Z,15Z)-9,12,15-Octadecatrienoico], a EPA (ácido eicosapentaenoico, o [ácido (5Z,8Z,11Z,14Z,17Z)-5,8,11,14,17-Eicosapentaenoico]) y a DHA [ácido docosahexaenoico, o ácido (4Z,7Z,10Z,13Z,16Z,19Z)-docosa-4,7,10,13,16,19-hexaenoico]. Las fuentes comunes de ácidos grasos EPA y DHA omega-3 de animales incluyen aceites de pescado, aceite de huevo, aceites de calamar y aceite de krill mientras que algunos aceites vegetales, tales como aceite de semilla y baya de espinillo amarillo, células de algas, semilla de lino, semilla de Chia y semilla de cáñamo, contienen altos niveles de ALA.

El ácido linoleico [ácido (9Z,12Z)-9,12-Octadecadienoico], el ácido gamma-linolénico (GLA, o ácido all-cis-6,9,12-octadecatrienoico) y el ácido araquidónico [ácido (5Z,8Z,11Z,14Z)-5,8,11,14-Eicosatetraenoico]) son ácidos grasos omega-6. GLA es un ácido graso omega-6 que se encuentra principalmente en aceites de origen vegetal tales como aceite de semilla de borraja, aceite de onagra y aceite de semilla de grosella negra.

Los ácidos grasos omega-3 son vitales para el metabolismo normal. Los omega-3 se consideran ácidos grasos esenciales, es decir, no pueden ser sintetizados por el cuerpo humano excepto por los mamíferos que tienen una capacidad limitada, cuando la dieta incluye el ácido graso omega-3 de cadena más corta ALA, de formar los ácidos grasos omega-3 de cadena larga más importantes, EPA y después de EPA, el más importante, DHA con una eficiencia aún mayor. En la actualidad se acepta que los ácidos grasos poliinsaturados omega-3, especialmente EPA y DHA juegan papeles importantes en un número de aspectos de la salud humana. Sin embargo, la pesca excesiva y las preocupaciones sobre la contaminación del ambiente marino indican una necesidad de desarrollar fuentes alternativas, sostenibles de ácidos grasos poliinsaturados de cadena muy larga (VLC-PUFA) tales como EPA y DHA [1]. Los ácidos grasos omega-6 se consideran ácidos grasos esenciales: Estos son necesarios para la salud humana. Junto con los ácidos grasos omega-3, el ácido graso omega-6 juega un papel importante en la función del cerebro, así como el crecimiento y desarrollo normal. Los ácidos grasos omega-3 y el ácido graso omega-6 ayudan a estimular el crecimiento de la piel y el cabello, mantienen la salud ósea, regulan el metabolismo y mantienen el sistema reproductivo [2]. Algunas investigaciones clínicas preliminares sugieren que GLA puede ser útil para la Neuropatía diabética, Artritis reumatoide, Alergias, Eczema, Presión sanguínea alta (Hipertensión), Síntomas menopáusicos, etc. La relación de la ingesta dietética de ácidos grasos esenciales omega-6 y omega-3 se considera importante para la salud de los seres humanos [3].

Hasta la fecha se han reportado una gran cantidad de microorganismos oleaginosos. El aceite que producen, frecuentemente denominado Aceite Microbiano (SCO), es similar a los de las plantas y pueden usarse para la producción de biodiesel, alimentos y productos industriales [4-6]. En la actualidad el SCO es ampliamente aceptado en el mercado y existe una conciencia creciente de los beneficios para la salud de los PUFA, tales como ácido γ -linolénico (GLA), ácido araquidónico (ARA), DHA y EPA. ARA y DHA se han usado además para la fortificación de fórmulas infantiles en muchas partes del mundo. Los aceites de pescado son fuentes ricas en DHA y EPA y una cantidad limitada de semillas oleaginosas vegetales son buenas fuentes de otros PUFA. Los protistas y dinoflagelados marinos, tales como las especies de *Thraustochytrium*, *Schizochytrium* y *Cryptocodinium* son las fuentes ricas en DHA, mientras que las microalgas como *Phaeodactylum Monodus* son buenas fuentes de EPA. Las especies de hongos inferiores *Mortierella* acumulan un alto porcentaje de ARA en la fracción lipídica [7].

Si bien la levadura *Yarrowia lipolytica* tal vez ha disfrutado de una larga historia de investigación y desarrollo como un huésped para transformar por bioingeniería para SCO [8-12], *Rhodosporidium toruloides* (conocido además como *Rhodotorula glutinis*) ha atraído una mayor atención debido a su capacidad de realizar una fermentación de mayor

densidad celular a una rápida tasa de crecimiento, produciendo eficientemente una masa celular con un contenido de aceite de > 67 % (p/p de masa celular seca) [13-16].

El *Pucciniomycotina* es un subfilo de hongos en el filo de *Basidiomycota* [17]. Posee muchas especies que tienen importantes aplicaciones industriales. Por ejemplo, un número de especies en los géneros *Rhodospordium* y *Sporidiobolus*, tales como *Rhodospordium toruloides* (conocido además como *Rhodotorula gracilis*, *Rhodospordium glutinis*, *Rhodotorula glutinis*, *Torula koishikawensis* y *Torula rubescens*) y *Sporobolomyces salmonicolor*, son levaduras unicelulares ricas en aceite capaces de una fermentación de alta densidad [6, 18]. Estas especies poseen gran potencial como un huésped para la producción de hidrocarburos de cadena larga, tales como triacilglicerol (TAG, o grasa), ésteres de ácidos grasos (biodiesel), alcoholes grasos, alcoholes, lactonas, terpenoides y vitaminas [14, 19-21].

Los genomas de *Rhodospordium* y *Rhodotorula* son altamente ricos en GC, lo que se ha encontrado que influye profundamente en la transformación genética y la expresión de proteínas [22-24]. La ingeniería metabólica es una técnica eficaz para mejorar la producción de metabolitos en plantas y microbios. En términos de bioingeniería de ácidos grasos omega-3, la expresión de diversas desaturasas y elongasas, tanto en plantas como en levadura oleaginosa, es crítica para la producción de PUFA [25]. GLA se sintetiza a partir de ácido linoleico (LA; C18:2Δ9,12 *cis*) por la Δ6-desaturasa. El aceite de semilla de cártamo (*Carthamus tinctorius*) contiene alto contenido de LA y se ha modificado mediante la transformación con Δ6-desaturasas de *Mortierella alpina* y *Saprolegnia diclina* para lograr más del 50 % (v/v) de GLA respectivamente [26].

Tanto ALA como GLA son precursores para la producción de ácidos grasos omega-3 de cadenas más largas, tales como ácido araquidónico (AA), EPA y ácido docosahexaenoico (DHA) [7, 27]. Además, la capacidad de producir altos niveles de ALA y GLA con una alta productividad volumétrica es crucial para la bioingeniería de los PUFA de cadena más larga en *Rhodospordium toruloides*. Por lo tanto, existe una necesidad de desarrollar especies fúngicas del género *Rhodospordium* y *Rhodotorula* que producen altos niveles de ALA y GLA que después están disponibles en las especies fúngicas para la producción de los PUFA de cadena más larga.

Resumen de la invención

La presente invención se refiere al campo de la biotecnología de hongos, más en particular a métodos de ingeniería genética para la producción de ácidos grasos poliinsaturados omega-3 (PUFA) en huéspedes fúngicos seleccionados de los géneros *Rhodospordium* y *Rhodotorula*.

En un primer aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido α-linolénico (ALA) presente en una cantidad de al menos 9 % de ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa que usa un aldehído de ácido graso como un sustrato. La ALD nativa es codificada por un gen de ALD nativa. En una modalidad el gen de ALD nativa codifica una aldehído deshidrogenasa (ALD) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:3. En otra modalidad, el gen de ALD nativa codifica una ALD que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de aminoácidos expuesta en la SEQ ID NO:3. En una modalidad, el gen de ALD nativa tiene la secuencia de nucleótidos genómica expuesta en la SEQ ID NO:1. En otra modalidad, el gen de ALD nativa tiene la secuencia de nucleótidos de ADNc expuesta en la SEQ ID NO:2. En una modalidad adicional, el gen de ALD nativa tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos expuesta en la SEQ ID NO:1 o SEQ ID NO:2.

En algunas modalidades, la actividad reducida de la ALD nativa es provocada por la expresión reducida del gen de ALD nativa. La expresión reducida puede ser provocada por cualquier mecanismo genético o epigenético. En una modalidad, la expresión reducida es provocada por un mecanismo del ARNi, tales como siARN, shARN, miARN y similares. En otra modalidad, la expresión reducida es provocada por un represor de la transcripción artificial. En una modalidad adicional, la expresión reducida es provocada por un mecanismo antisentido. En una modalidad, la expresión reducida es provocada por la supresión con sentido. En una modalidad adicional, la actividad reducida es provocada por una mutación del gen nativo. En una modalidad, la mutación puede ser una de sustituciones, delección, inserciones, adición, o inversión y similares lo que da como resultado una actividad reducida. En otra modalidad, la mutación puede ser provocada mediante recombinación homóloga. En una modalidad adicional, la mutación puede ser provocada mediante la inserción del T-ADN o transposones.

En un segundo aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido α-linolénico (ALA) presente en una cantidad de al menos 49 % de ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa como se describe en la presente descripción. En otras modalidades, el genoma del huésped fúngico se ha modificado para incluir establemente dos o más genes que codifican proteínas que están implicadas en la biosíntesis de ácidos grasos. Los ejemplos de tales proteínas son una acil-CoA delta-12 desaturasa, una estearoil-CoA-delta-9-desaturasa, una omega-3 desaturasa, un ácido graso elongasa, una acetil-CoA carboxilasa (ACC), una ATP: citrato liasa (ACL), una diacilglicerol aciltransferasa (DGA) o una enzima málica (MAE). En algunas modalidades, las secuencias

codificantes de tales genes se han modificado para que contengan al menos 55 % de contenido de G y C, preferentemente 60 %-70 % de contenido de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición.

En una modalidad, una ATP:citrato liasa (ACL) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:88. En otra modalidad, esta ATP:citrato liasa (ACL1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:86. En otra modalidad, esta ATP:citrato liasa (ACL1) es codificada por un ácido nucleico de ADNc que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:87. En otra modalidad, el gen de la ATP:citrato liasa (ACL1) tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una diacilglicerol aciltransferasa (DGA1) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:82. En otra modalidad, esta diacilglicerol aciltransferasa (DGA1) es codificada por un ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:80. En una modalidad adicional, esta diacilglicerol aciltransferasa (DGA1) es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:81. En otra modalidad, el gen de la diacilglicerol aciltransferasa (DGA1) tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una enzima málica (MAE1) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:85. En otra modalidad, esta enzima málica (MAE1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:83. En una modalidad adicional, esta enzima málica (MAE1) es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:84. En otra modalidad, el gen de la enzima málica (AME1) tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En otro aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido gamma-linolénico (GLA) presente en una cantidad de al menos el 30 % de los ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa como se describe en la presente descripción. En otras modalidades, el genoma del huésped fúngico se ha modificado para expresar establemente dos o más genes adicionales que codifican proteínas que están implicadas en la biosíntesis de ácidos grasos. Los ejemplos de tales proteínas son una acil-CoA delta-12 desaturasa, una estearoil-CoA-delta-9-desaturasa, acil-CoA delta-6 desaturasa, un ácido graso elongasa, una acetil-CoA carboxilasa (ACC), una ATP: citrato liasa (ACL), una diacilglicerol aciltransferasa (DGA) o una enzima málica (MAE). En algunas modalidades, las secuencias codificantes de tales genes se han modificado para que contengan al menos 55 % de contenido de G y C, preferentemente 60 %-70 % de contenido de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición.

En una modalidad, una acetil-CoA carboxilasa (ACC1) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:91. En otra modalidad, esta acetil-CoA carboxilasa (ACC1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:89. En una modalidad adicional, esta acetil-CoA carboxilasa (ACC1) es codificada por un ácido nucleico de ADNc que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:90. En otra modalidad, el gen de la acetil-CoA carboxilasa (ACC1) tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una acil-CoA delta-12 desaturasa tiene una secuencia aminoacídica expuesta en las SEQ ID NO:5 y 94. En otra modalidad, esta acil-CoA delta-12 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en las SEQ ID NO:4, 92 y NO:93. En otra modalidad, los genes de la acil-CoA delta-12 desaturasa tienen una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de las secuencias de nucleótidos descritas en la presente descripción.

En una modalidad, una estearoil-CoA-delta-9-desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:8. En una modalidad, esta estearoil-CoA-delta-9-desaturasa es codificada por un ácido nucleico genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:6. En otra modalidad, esta estearoil-CoA-delta-9-desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:7. En otra modalidad, el gen de la estearoil-CoA-delta-9-desaturasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una omega-3 desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:10. En una modalidad, esta omega-3 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:9. En una modalidad adicional, una omega-3 desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:12. En una modalidad, esta última omega-3 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:11. En otra modalidad, el gen de la omega-3 desaturasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una acil-CoA delta-6 desaturasa tiene una secuencia aminoacídica expuesta en las SEQ ID NO:96 y 98 en donde las secuencias son codificadas por un ADN que contiene al menos G y C, preferentemente 60 %-70 % de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición. En una modalidad, la acil-CoA delta-6 desaturasa es codificada por un ácido nucleico expuesto en las SEQ ID NO:95 y 97.

En una modalidad, un ácido graso elongasa tiene la secuencia de aminoácidos expuesta en las SEQ ID NO:101 y 104. En otra modalidad, este ácido graso elongasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en las SEQ ID NO:99, 100, 102 y 103. En otra modalidad, el gen del ácido graso elongasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En algunas modalidades, los genes descritos en la presente descripción que se han incorporado establemente en el genoma fúngico se unen operativamente a un promotor que permite la expresión eficiente en especies del género *Rhodospordium* y el género *Rhodotorula*. Los promotores para cada gen incorporado pueden ser iguales o diferentes. En algunas modalidades, los promotores son promotores encontrados en especies del género *Rhodospordium* y el género *Rhodotorula*. Los ejemplos de promotores adecuados incluyen, pero no se limitan a, promotores de los siguientes genes que codifican las siguientes proteínas: gliceraldehído 3-fosfato deshidrogenasa (GPD), proteína portadora de acil-CoA (ACP), ácido graso desaturasa, factor de alargamiento de la traducción (TEF), piruvato descarboxilasa (PDC), enolasa (2-fosfoglicerato deshidratasa) (ENO), peptidilprolil isomerasa (PPI), acetil-CoA carboxilasa (ACC) o transaldolasa. En otras modalidades, los genes descritos en la presente descripción incluyen además un terminador transcripcional de ARNm que puede encontrarse en cualquier especie eucariota y los virus de ADN.

En otro aspecto, la presente invención proporciona un método para producir ácidos grasos poliinsaturados (PUFA) omega-3 y omega-6, que comprende cultivar una célula huésped fúngica descrita en la presente descripción en condiciones adecuadas para producir los PUFA. Puede usarse cualquier medio con al menos una fuente de carbono al 5 %, como glucosa, manosa, glicerol, sacarosa. El ejemplo del medio es Medio MinLG que contiene 30-100 g de glucosa, 1.5 g de extracto de levadura, 0.5 g de $(\text{NH}_4)_2\text{SO}_4$, 2.05 g de K_2HPO_4 , 1.45 g de KH_2PO_4 , 0.6 g de MgSO_4 , 0.3 g de NaCl, 10 mg de CaCl_2 , 1 mg de FeSO_4 , 0.5 mg de ZnSO_4 , 0.5 mg de CuSO_4 , 0.5 mg de H_3BO_4 , 0.5 mg de MnSO_4 , 0.5 mg de NaMoO_4 (por litro). El pH del medio se ajusta a 6-7. El cultivo celular se realiza a 25 °C - 32 °C.

Breve descripción de las figuras

Las Figuras 1A-1E muestran la organización de T-ADN en constructos usados en esta invención. Todos los vectores binarios tienen la misma cadena principal de pPZP200 [28]. Figura 1A: pEC3GPD-GUS. Figura 1B: pEC3Pxxx-HPT3. Figura 1C: pRH2034. Figura 1D: pRHDGA1 y pRHMAE1. Figura 1E: pRH201. LB: borde izquierdo del T-ADN; RB: borde derecho del T-ADN; *Pgpd*: promotor de *Umgpd1* de 595 pb; *P_{GPD1}*: promotor de *RtGPD1* de 795 pb; Pxxx: diversos promotores; *hpt-3*: gen de resistencia a higromicina con codón optimizado en base al sesgo en el uso de codones en *R. toruloides*; *GUS*: gen de la β -glucuronidasa de *E. coli*; *T_{35S}*: terminador del gen de 35S del virus del mosaico de la coliflor; *Tnos*: terminador del gen de la nopalina sintasa de *A. tumefaciens*. Se muestran los sitios de corte de enzimas de restricción únicas en el plásmido.

Las Figuras 2A-2C muestran la organización del T-ADN en los constructos usados en esta invención. Todos los vectores binarios tienen la misma cadena principal de pPZP200 [28]. Figura 2A: pRHE001. Figura 2B: pRHE002. Figura 2C: pRHE003. LB: borde izquierdo del T-ADN; RB: borde derecho del T-ADN; *Pgpd*: promotor de *Umgpd1* de 595 pb; *P_{GPD1}*: promotor de *RtGPD1* de 795 pb; *hpt-3*: gen de resistencia a higromicina con codón optimizado en base al sesgo en el uso de codones en *R. toruloides*; *T_{35S}*: terminador del gen de 35S del virus del mosaico de la coliflor; *Tnos*: terminador del gen de la nopalina sintasa de *A. tumefaciens*. Se muestran los sitios de corte de enzimas de restricción únicas en el plásmido.

Las Figuras 3A-3C muestran la organización del T-ADN en los constructos usados en esta invención. Todos los vectores binarios tienen la misma cadena principal de pPZP200 [28]. Figura 3A: pRHE004. Figura 3B: pRHE005. Figura 3C: pRHE006. LB: borde izquierdo del T-ADN; RB: borde derecho del T-ADN; *Pgpd*: promotor de *Umgpd1* de 595 pb; *P_{GPD1}*: promotor de *RtGPD1* de 795 pb; *hpt-3*: gen de resistencia a higromicina con codón optimizado en base al sesgo en el uso de codones en *R. toruloides*; *T_{35S}*: terminador del gen de 35S del virus del mosaico de la coliflor; *Tnos*: terminador del gen de la nopalina sintasa de *A. tumefaciens*. Se muestran los sitios de corte de enzimas de restricción únicas en el plásmido.

Las Figuras 4A-4D muestran las cepas mutantes identificadas mediante genética directa. Figura 4A: Comparación de niveles promedio de ALA ($\text{C}_{18:3n=9}$) en el tipo silvestre y en mutantes de T-ADN (RCM) seleccionados contra cerulenina. Figura 4B: Niveles relativos de ALA en un mutante RCM individual. Figura 4C: Comparación de la acumulación de lípidos en el tipo silvestre y en mutantes seleccionados contra violeta de tetrazolio. Figura 4D: Comparación de la acumulación de lípidos en el tipo silvestre y en mutantes seleccionados contra colorante fluorescente rojo Nilo.

Las Figuras 5A-5F muestran los estudios genéticos inversos de *ALD1* en *R. toruloides*. Figura 5A: Diagrama esquemático de *ALD1* y su estrategia de delección. Las secuencias homólogas usadas para la delección de *ALD1* eran de 730 pb y 829 pb de longitud, entre -185 a +535 y +1948 a +2776 desde el codón de inicio de la traducción. Figura 5B: Análisis de la transferencia Southern de Δald1 . La secuencia de ADN marcada con digoxigenina (*ALD1R* marcada en la Figura 5A) se usó como la sonda contra el ADN genómico digerido con HincII. Figura 5C: Rendimientos relativos de lípidos en el tipo

silvestre y en $\Delta ald1$ el día 3 y 4. Figura 5D: Niveles relativos de ácido α -linolénico (ALA) en el tipo silvestre y en $\Delta ald1$. ** representa una diferencia muy significativa mediante la prueba t estadística ($P < 0.01$). Figura 5E: Biomasa celular seca (Biomass) y contenidos de ALA (porcentaje de ácidos grasos totales, % de TFA) del tipo silvestre y de $\Delta ald1$. Figura 5F: Colores de células en agar de dextrosa de papa (PDA) y caldo YPD.

La Figura 6 muestra una alineación del aldehído de ácido graso deshidrogenasas de diversas especies de Basidiomycotous. Ab: *Agaricus bisporus* (EKM75339.1; SEQ ID NO:105); Pc: *Phanerochaete camosa* (EKM57674.1; SEQ ID NO:106); Mg: *Malassezia globosa* (XP_001730031.1; SEQ ID NO:107); Sr: *Sporisorium reilianum* (CBQ71609.1; SEQ ID NO:108); Um: *Ustilago maydis* (XP_762570.1; SEQ ID NO:109); Ml: *Melampsora larici-populina* (EGG04055.1; SEQ ID NO:110); Pg: *Puccinia graminis* f. sp. tritici (XP_003338710.1; SEQ ID NO:111); Pt: *Puccinia triticina* (XP_003338710.1; SEQ ID NO:112); Ps: *Puccinia striiformis* (locus genómico CQM_00777.1; SEQ ID NO:113); Mv: *Microbotryum violaceum* (locus genómico MVLG_02667.1; SEQ ID NO:114); Rg2: *R. glutinis* ATCC 204091 (SEQ ID NO:115); Rt3: *R. toruloides* NP11 (EMS18750.1; SEQ ID NO:116); Rt1: *R. toruloides* ATCC 10657 (SEQ ID NO:117). [Nota: el Listado de Secuencias incluye solo una secuencia parcial de las SEQ ID NO:112, 113, 115 y 117. Ver la Figura para las secuencias completas.]

Las Figuras 7A-7F muestran los efectos de *DGA1* y *MAE1* para mejorar la acumulación de lípidos. Figura 7A: Análisis de RT-PCR cuantitativo de *DGA1* y *MAE1* en *R. toruloides* ATCC 10657. Las expresiones génicas en los mutantes transformados por ingeniería genética (*DGA1* y *MAE1*) se normalizaron contra la expresión en la cepa de tipo silvestre. Figura 7B: Rendimientos relativos de lípidos en las cepas de tipo silvestre y en dos cepas transformadas por ingeniería genética (tipo silvestre, *DGA1* y *MAE1*). Las cantidades de lípidos se normalizaron contra las cantidades en el tipo silvestre. Figura 7C: Perfiles de ácidos grasos en las cepas anteriores. Los lípidos intracelulares se extrajeron de las cepas anteriores después de un bioproceso de 3 días. Figura 7D: Composición de ácidos grasos insaturados en las cepas anteriores. El % de TFA representa el porcentaje de ácidos grasos totales. Figura 7E: Diagrama esquemático de la delección de *DGA1*. Figura 7F: Análisis de la transferencia Southern de $\Delta dga1$. Los ADN genómicos del tipo silvestre y del mutante candidato se digirieron con *HincII* y se analizaron contra el fragmento de ADN de *DGA1L* marcado con DIG como se indica en la Figura 7E.

La Figura 8 muestra los perfiles de ácidos grasos en diferentes cepas de *R. toruloides* (*R. glutinis*). Rt1: *R. toruloides* ATCC 10657; Rt2: *R. toruloides* ATCC 10788; Rg1: *R. glutinis* ATCC 90781; Rg2: *R. glutinis* ATCC 204091. *R. toruloides* (*R. glutinis*) se cultivó en medio de acumulación de lípidos (MinLG) como se describió anteriormente [21] con algunas modificaciones. El medio MinLG contiene 30 g de glucosa, 1.5 g de extracto de levadura, 0.5 g de $(\text{NH}_4)_2\text{SO}_4$, 2.05 g de K_2HPO_4 , 1.45 g de KH_2PO_4 , 0.6 g de MgSO_4 , 0.3 g de NaCl , 10 mg de CaCl_2 , 1 mg de FeSO_4 , 0.5 mg de ZnSO_4 , 0.5 mg de CuSO_4 , 0.5 mg de H_3BO_3 , 0.5 mg de MnSO_4 , 0.5 mg de NaMoO_4 (por litro). El pH del medio se ajustó a 6.1. El cultivo celular se realizó a 28 °C durante 4 días con agitación constante (250 rpm).

Las Figuras 9A y 9B muestran los procesos de ingeniería genética para las cepas con alto contenido de ALA. Figura 9A: Contenido de ALA en cepas de tipo silvestre y de desaturasas transformadas por ingeniería genética. Figura 9B: Contenido de ALA en la cepa nula de *ald1* (*aldle*) y la cepa *aldle* que contiene tanto los casetes génicos RtGPD1:MaFAD2-2 y RtGPD1:LuFAD3-2.

La Figura 10 muestra el ensayo enzimático de ALD1 y de ALD1 truncada en el extremo C-terminal (ALD1n). La reacción se mezcló con tampón Tris-Cl 20 mM (pH8.0), NAD⁺ o NADP⁺ 1.5 mM, dodecanal 1.0 mM y 10 μl de enzima purificada. El ensayo se realizó a 25 °C durante 2 min.

Las Figuras 11A-11D muestran la caracterización del gen de la delta-12 desaturasa en *R. toruloides* ATCC 10657 (RtFAD2). Figura 11A: Análisis de la transferencia Southern de la cepa de tipo silvestre y la cepa con *FAD2* inactiva (*fad2Δ*). El ADN total se digirió con *PstI* y la transferencia se hibridó mediante el uso de un fragmento de PCR de homología derecha marcado con digoxigenina de *FAD2* (*FAD2R*, Figura 11C). Figura 11B: Perfil de contenido de ácidos grasos del tipo silvestre y de *fad2Δ* realizado en triplicados. Figura 11D: Perfil del contenido de ácidos grasos del tipo silvestre y de *fad2Δ* cultivados en medio YNB complementado con y sin LA (C18:2).

Las Figuras 12A-12D muestran la caracterización de los mutantes por delección de *ELO1* y *ELO2* en *R. toruloides* ATCC 10657 (*elo1Δ* y *elo2Δ*) (Figura 12A) y (Figura 12C). Análisis de la transferencia Southern de *elo1Δ* y *elo2Δ*. El ADN total se digirió con *PvuI* y se hibridó contra *ELO1R* y *ELO2L* marcados con digoxigenina, respectivamente (Figura 12B) y (Figura 12D). Perfiles del contenido de ácidos grasos de *elo1Δ* y *elo2Δ*.

Las Figuras 13A-13E muestran la ilustración esquemática de los vectores usados en la Figura 14. Figura 13A: Constructo de sobreexpresión de *FAD1* (delta-9-oleato desaturasa, *ELO1*). Figura 13B: Constructo de sobreexpresión de *FAD2* (delta-12 desaturasa). Figura 13C: Constructo de sobreexpresión de omega-3 desaturasa (delta-15 desaturasa). Figura 13D: Vector de sobreexpresión del gen doble de *FAD1* y *FAD2*. Figura 13E: Vector de sobreexpresión del gen triple de *FAD1*, *FAD2* y omega-3 desaturasa. LB: borde izquierdo del T-ADN; RB: borde derecho del T-ADN; *RgGPD1*: promotor de *GPD1* de 685 pb derivado de *Rhodotorula grammis* WP1; *RtGPD1*: promotor de *GPD1* de 795 pb derivado de *R. toruloides* ATCC 10657; *hpt-3*: gen de resistencia a higromicina con codón optimizado en base al sesgo en el uso de codones en *R. toruloides*; TSV40: Terminador del gen del antígeno T largo del virus vacuolante de simio 40; T_{35S}: terminador del gen de 35S del virus del mosaico de la coliflor; Tnos: terminador del gen de la nopalina sintasa de A.

tumefaciens; RtENO1: promotor de la versión de 445 pb del gen *ENO1* de *R. toruloides* ATCC 10657; RtACC1: promotor de la versión de 805 pb del gen *ACC1* de *R. toruloides* ATCC 10657; MaFAD2-2 es la SEQ ID NO:4, y VfFAD3-2 es la SEQ ID NO:11.

Figura 14: Perfil del contenido de ácidos grasos de cepas transformadas. Se usaron tres transformantes independientes para cada análisis excepto para Δ ald1-OMA2, para el cual se analizaron 18 transformantes independientes totales. Todos los transformantes se fermentaron en medio RL2 durante 5 días. WT: cepa de *R. toruloides* ATCC 10657 de tipo silvestre; Δ ald1-: los constructos se transformaron en mutantes inactivos de *ADL1*; los constructos se muestran en la Figura 13. OEL1: Figura 13A; MA: Figura 13B; AF3: Figura 13C; OM2: Figura 13D; OMA2: Figura 13E.

Las Figuras 15A-15D muestran la caracterización del gen de la ATP-citrato liasa putativa en *R. toruloides* ATCC 10657 (RtACL1). Figura 15A: Diagrama esquemático del gen *ACL1* y su estrategia de delección. Figura 15B: Análisis de la transferencia Southern de la cepa inactiva *acl1Δ*. El ADN total se digirió con *PvuI* y se hibridó contra *ACL1L* marcado con digoxigenina. M: Marcador de peso molecular de ADN marcado con digoxigenina VII (Roche Diagnosis, Estados Unidos). Figura 15C: Efectos de la alteración de *ACL1* en la biomasa, los rendimientos de lípidos, el contenido de lípidos y la glucosa residual. Figura 15D: Perfil del contenido de ácidos grasos de las cepas de tipo silvestre y *acl1Δ*.

Descripción detallada de la invención

La presente invención se refiere al campo de la biotecnología fúngica, más en particular a métodos de ingeniería genética para la producción de ácidos grasos poliinsaturados (PUFA) en huéspedes fúngicos seleccionados del género *Rhodospordium* y *Rhodotorula*.

A menos que se defina de cualquier otra manera, todos los términos técnicos y científicos usados en la presente descripción tienen el mismo significado que el que se conoce comúnmente por el experto en la técnica a la que pertenece la invención.

Como se usa en la presente descripción, "ALD1" es un aldehído deshidrogenasa que usa aldehídos de ácidos grasos como un sustrato.

Como se usa en la presente descripción, "alelo" se refiere a cualquiera de una o más formas alternativas de un locus génico, todos de cuyos alelos se relacionan con un rasgo o una característica. En una célula u organismo diploide, los dos alelos de un gen dado ocupan los loci correspondientes de cromosomas homólogos.

Un "dsARN" o "molécula de ARNi," como se usa en la presente descripción en el contexto de ARNi, se refiere a un compuesto, que es capaz de regular negativamente o reducir la expresión de un gen o la actividad del producto de tal gen en una medida suficiente para lograr un efecto biológico o fisiológico deseado. El término "dsARN" o "molécula de ARNi," como se usa en la presente descripción, se refiere a uno o más de un dsARN, siARN, shARN, ihpARN, shARN sintético, miARN.

El término "regulado negativamente," cuando se refiere a genes inhibidos por el método del ARNi sujeto, se refiere a una disminución del nivel de expresión de un(os) gene(s) en la presencia de uno o más constructos de ARNi cuando se compara con el nivel en la ausencia de tales constructos de ARNi. El término "regulado negativamente" se usa en la presente descripción para indicar que la expresión del gen diana disminuye en 1-100 %. Por ejemplo, la expresión puede reducirse en alrededor del 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 %.

El término "expresión" con respecto a una secuencia génica se refiere a la transcripción del gen y, según sea apropiado, la traducción del transcrito de ARNm resultante a una proteína. Por lo tanto, como será evidente a partir del contexto, la expresión de una secuencia codificante de proteína resulta de la transcripción y traducción de la secuencia codificante.

Como se usa en la presente descripción, "gen" se refiere a una secuencia de ácidos nucleicos que abarca una región promotora 5' asociada con la expresión del producto génico, cualquiera de las regiones de intrones y exones y regiones no traducidas 3' o 5' asociadas con la expresión del producto génico.

El término "silenciamiento génico" se refiere a la supresión de la expresión génica, por ejemplo, expresión de transgenes, de genes heterólogos y/o de genes endógenos. El silenciamiento génico puede estar mediado a través de procesos que afectan la transcripción y/o a través de procesos que afectan los mecanismos postraduccionales. El silenciamiento génico puede ser específico de alelo en donde se produce el silenciamiento específico de un alelo de un gen.

Como se usa en la presente descripción, "genotipo" se refiere a la constitución genética de una célula u organismo.

El término "heterólogo" o "exógeno" cuando se usa con referencia a las porciones de un ácido nucleico indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la naturaleza en la misma relación entre sí. Por ejemplo, el ácido nucleico típicamente se produce por vía recombinante, con dos o más secuencias a partir de genes no relacionados dispuestos para preparar un nuevo ácido nucleico funcional, por ejemplo, un promotor a partir de una fuente

y una región de codificación a partir de otra fuente. De la misma manera, una proteína heteróloga o exógena indica que la proteína comprende dos o más subsecuencias que no se encuentran en la naturaleza en la misma relación entre sí (por ejemplo, una proteína de fusión).

5 El término "homólogo" como se usa en la presente descripción se refiere a un gen relacionado con un segundo gen por descendencia de una secuencia de ADN ancestral común. El término, homólogo, puede aplicarse a la relación entre genes separados por un evento de especiación (ortólogo) o a la relación entre genes separados por el evento de la duplicación genética (parálogo). El término homólogo se usa genéricamente para referirse a todas las especies.

10 Como se usa en la presente descripción, "fenotipo" se refiere a las características detectables de una célula u organismo, cuyas características son la manifestación de la expresión génica.

15 Los términos "polinucleótido," "ácido nucleico" y "molécula de ácido nucleico" se usan indistintamente en la presente descripción para referirse a un polímero de nucleótidos que pueden ser una disposición lineal y secuencial natural o sintética de nucleótidos y/o nucleósidos, que incluyen ácido desoxirribonucleótico, ácido ribonucleótico y derivados de estos. Esto incluye ADN cromosómico, plásmidos de autorreplicación, polímeros infecciosos de ADN o ARN y ADN o ARN que desempeña un papel principalmente estructural. A menos que se indique lo contrario, los ácidos nucleicos o polinucleótidos se escriben de izquierda a derecha en orientación 5' a 3', los nucleótidos se mencionan mediante sus códigos de una sola letra comúnmente aceptados. Los intervalos numéricos incluyen los números que definen el intervalo.

20 Los términos "polipéptido," "péptido," y "proteína" se usan indistintamente en la presente descripción para referirse a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o más desechos de aminoácidos es un análogo químico artificial de un aminoácido de origen natural correspondiente, así como también a polímeros de aminoácidos de origen natural. Puede hacerse referencia a los aminoácidos mediante sus símbolos de tres letras o de una letra comúnmente conocidos. Las secuencias de aminoácidos se escriben de izquierda a derecha en orientación amino a carboxi, respectivamente. Los intervalos numéricos incluyen los números que definen el intervalo.

30 Se entiende que "unión operable" o "unido operablemente" o "unido operativamente" como se usa en la presente descripción significa, por ejemplo, la disposición secuencial de un promotor y el ácido nucleico a expresar y, si es apropiado, los elementos reguladores adicionales tales como, por ejemplo, un terminador, de una manera en que cada uno de los elementos reguladores puede completar su función en la expresión recombinante del ácido nucleico para producir dsARN. Esto no requiere necesariamente la unión directa en el sentido químico. Las secuencias de control genético tales como, por ejemplo, las secuencias potenciadoras, también pueden ejercer su función sobre la secuencia diana desde posiciones que son algo distantes, o de hecho desde otras moléculas de ADN (localización cis o trans). Las disposiciones preferidas son aquellas en las cuales la secuencia de ácido nucleico a expresar de manera recombinante se coloca aguas abajo de la secuencia que actúa como promotor, de manera que las dos secuencias se unen covalentemente entre sí. Las secuencias reguladoras o de control pueden colocarse en el lado 5' de la secuencia de nucleótidos o en el lado 3' de la secuencia de nucleótidos como se conoce bien en la técnica.

40 El término "alrededor de" o "aproximadamente" significa dentro de un intervalo estadísticamente significativo de un valor. Tal rango puede estar dentro de un orden de magnitud, preferentemente dentro del 50 %, con mayor preferencia dentro del 20 %, aún con mayor preferencia dentro del 10 %, e incluso con mayor preferencia dentro del 5 % de un valor o intervalo dado. La variación permisible abarcada por el término "alrededor de" o "aproximadamente" depende del sistema particular en estudio, y puede ser apreciada fácilmente por un experto en la técnica.

50 Como se usa en la presente descripción, el término "identidad de secuencia," "similitud de secuencia" u "homología" se usa para describir las relaciones de secuencia entre dos o más secuencias de nucleótidos. El porcentaje de "identidad de secuencia" entre dos secuencia se determina mediante la comparación de dos secuencias óptimamente alineadas sobre una ventana de comparación tal como la longitud completa de una SEQ ID NO: de referencia, en donde la porción de la secuencia en la ventana de comparación puede comprender adiciones o deleciones (es decir, interrupciones) en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las cuales la base de ácidos nucleicos o el residuo de aminoácido es idéntico en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia. Se dice que una secuencia que es idéntica en cada posición en comparación con una secuencia de referencia es idéntica a la secuencia de referencia y viceversa. Se dice que una primera secuencia de nucleótidos cuando se observa en la dirección 5' a 3' es un "complemento" de, o es complementaria a, una segunda secuencia o secuencia de nucleótidos de referencia observada en la dirección 3' a 5' si la primera secuencia de nucleótidos exhibe complementariedad completa con la segunda secuencia o secuencia de referencia. Tal como se usa en la presente descripción, se dice que las moléculas de la secuencia de ácido nucleico exhiben "complementariedad completa" cuando cada nucleótido de una de las secuencias leídas en la dirección 5' a 3' es complementario a cada nucleótido de la otra secuencia cuando se lee en la dirección 3' a 5'. Una secuencia de nucleótidos que es complementaria a una secuencia de nucleótidos de referencia exhibirá una secuencia idéntica a la secuencia complementaria inversa de la secuencia de nucleótidos de referencia. Estos términos y descripciones se definen bien en la técnica y se entienden fácilmente por los expertos en la técnica.

Como se usa en la presente descripción, una "ventana de comparación" se refiere a un segmento conceptual de al menos 6 posiciones contiguas, usualmente alrededor de 50 a alrededor 100, más usualmente alrededor de 100 a alrededor de 150, en el cual una secuencia puede compararse con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alinean de manera óptima. La ventana de comparación puede comprender adiciones o deleciones (es decir, interrupciones) de alrededor de 20 % o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. Los expertos en la técnica deben referirse a los métodos detallados usados para la alineación de secuencias, tal como en el Paquete de Software Release 7.0 de Wisconsin Genetics (Genetics Computer Group, 575 Science Drive Madison, Wise., Estados Unidos).

En un primer aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido α -linolénico (ALA) presente en una cantidad de al menos 9 % de ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa que usa un aldehído de ácido graso como un sustrato. La ALD nativa es codificada por un gen de ALD nativa. En una modalidad el gen de ALD nativa codifica una aldehído deshidrogenasa (ALD) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:3. En otra modalidad, el gen de ALD nativa codifica una ALD que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de aminoácidos expuesta en la SEQ ID NO:3. En una modalidad, el gen de ALD nativa tiene la secuencia de nucleótidos genómica expuesta en la SEQ ID NO:1. En otra modalidad, el gen de ALD nativa tiene la secuencia de nucleótidos de ADNc expuesta en la SEQ ID NO:2. En una modalidad adicional, el gen de ALD nativa tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos expuesta en la SEQ ID NO:1 o SEQ ID NO:2.

En algunas modalidades, la actividad reducida de la ALD nativa es provocada por la expresión reducida del gen de ALD nativa. La expresión reducida puede ser provocada por cualquier mecanismo genético o epigenético. En una modalidad, la expresión reducida es provocada por un mecanismo del ARNi, tales como siARN, shARN, miARN y similares. En otra modalidad, la expresión reducida es provocada por un represor de la transcripción artificial. En una modalidad adicional, la expresión reducida es provocada por un mecanismo antisentido. En una modalidad, la expresión reducida es provocada por la supresión con sentido. En una modalidad adicional, la actividad reducida es provocada por una mutación del gen nativo. En una modalidad, la mutación puede ser una de sustituciones, delección, inserciones, adición, o inversión y similares lo que da como resultado una actividad reducida. En otra modalidad, la mutación puede ser provocada mediante recombinación homóloga. En una modalidad adicional, la mutación puede ser provocada mediante la inserción del T-ADN o transposones.

En un segundo aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido α -linolénico (ALA) presente en una cantidad de al menos 49 % de ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa como se describe en la presente descripción. En otras modalidades, el genoma del huésped fúngico se ha modificado para incluir establemente dos o más genes que codifican proteínas que están implicadas en la biosíntesis de ácidos grasos. Los ejemplos de tales proteínas son una acil-CoA delta-12 desaturasa, una estearoil-CoA-delta-9-desaturasa, una omega-3 desaturasa, un ácido graso elongasa, una acetil-CoA carboxilasa (ACC), una ATP: citrato liasa (ACL), una diacilglicerol aciltransferasa (DGA) o una enzima málica (MAE). En algunas modalidades, las secuencias codificantes de tales genes se han modificado para que contengan al menos 55 % de contenido de G y C, preferentemente 60 %-70 % de contenido de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición.

En una modalidad, una ATP:citrato liasa (ACL) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:88. En otra modalidad, esta ATP:citrato liasa (ACL1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:86. En otra modalidad, esta ATP:citrato liasa (ACL1) es codificada por un ácido nucleico de ADNc que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:87. En otra modalidad, el gen de la ATP:citrato liasa (ACZ1) tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una diacilglicerol aciltransferasa (DGA1) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:82. En otra modalidad, esta diacilglicerol aciltransferasa (DGA1) es codificada por un ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:80. En una modalidad adicional, esta diacilglicerol aciltransferasa (DGA1) es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:81. En otra modalidad, el gen de la diacilglicerol aciltransferasa (DGA1) tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una enzima málica (MAE1) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:85. En otra modalidad, esta enzima málica (MAE1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:83. En una modalidad adicional, esta enzima málica (MAE1) es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:84. En otra modalidad, un gen de

la enzima málica (*MAE1*) tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En otro aspecto, la presente invención proporciona un huésped fúngico que tiene un ácido gamma-linolénico (GLA) presente en una cantidad de al menos el 30 % de los ácidos grasos totales en células del huésped fúngico. En una modalidad, el huésped fúngico es una especie del género *Rhodospordium*. En otra modalidad, el huésped fúngico es una especie del género *Rhodotorula*. En algunas modalidades, el huésped fúngico tiene actividad reducida de un aldehído deshidrogenasa (ALD) nativa como se describe en la presente descripción. En otras modalidades, el genoma del huésped fúngico se ha modificado para expresar establemente dos o más genes adicionales que codifican proteínas que están implicadas en la biosíntesis de ácidos grasos. Los ejemplos de tales proteínas son una acil-CoA delta-12 desaturasa, una estearoil-CoA-delta-9-desaturasa, acil-CoA delta-6 desaturasa, un ácido graso elongasa, una acetil-CoA carboxilasa (ACC), una ATP: citrato liasa (ACL), una diacilglicerol aciltransferasa (DGA) o una enzima málica (MAE). En algunas modalidades, las secuencias codificantes de tales genes contienen al menos 55 % de contenido de G y C, preferentemente 60 %-70 % de contenido de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición.

En una modalidad, una acetil-CoA carboxilasa (ACC) tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:91. En otra modalidad, esta acetil-CoA carboxilasa (ACC1) es codificada por un ácido nucleico de ADN genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:89. En una modalidad adicional, esta acetil-CoA carboxilasa (ACC1) es codificada por un ácido nucleico de ADNc que tiene una secuencia de nucleótidos expuesta en la SEQ ID N:90. En otra modalidad, el gen de la acetil-CoA carboxilasa (ACC1) tiene la secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una acil-CoA delta-12 desaturasa tiene una secuencia aminoacídica expuesta en las SEQ ID NO:5 y 94. En otra modalidad, esta acil-CoA delta-12 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en las SEQ ID NO:4, 92 y NO:93. En otra modalidad, los genes de la acil-CoA delta-12 desaturasa tienen una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de las secuencias de nucleótidos descritas en la presente descripción.

En una modalidad, una estearoil-CoA-delta-9-desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:8. En una modalidad, esta estearoil-CoA-delta-9-desaturasa es codificada por un ácido nucleico genómico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:6. En otra modalidad, esta estearoil-CoA-delta-9-desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:7. En otra modalidad, el gen de la estearoil-CoA-delta-9-desaturasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una omega-3 desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:10. En una modalidad, esta omega-3 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:9. En una modalidad adicional, una omega-3 desaturasa tiene una secuencia aminoacídica expuesta en la SEQ ID NO:12. En una modalidad, esta última omega-3 desaturasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en la SEQ ID NO:11. En otra modalidad, el gen de la omega-3 desaturasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En una modalidad, una acil-CoA delta-6 desaturasa tiene una secuencia aminoacídica expuesta en las SEQ ID NO:96 y 98 en donde las secuencias son codificadas por un ADN que contiene al menos G y C, preferentemente 60 %-70 % de G y C. En otras modalidades, al menos el 70 % de los codones tienen una C o G en la tercera posición. En una modalidad, la acil-CoA delta-6 desaturasa es codificada por un ácido nucleico expuesto en las SEQ ID NO:95 y 97.

En una modalidad, un ácido graso elongasa tiene la secuencia de aminoácidos expuesta en las SEQ ID NO:101 y 104. En otra modalidad, este ácido graso elongasa es codificada por un ácido nucleico que tiene una secuencia de nucleótidos expuesta en las SEQ ID NO:99, 100, 102 y 103. En otra modalidad, el gen del ácido graso elongasa tiene una secuencia de nucleótidos que tiene 75 %, 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la secuencia de nucleótidos descrita en la presente descripción.

En algunas modalidades, los genes descritos en la presente descripción que se han incorporado establemente en el genoma fúngico se unen operativamente a un promotor que permite la expresión eficiente en especies del género *Rhodospordium* y el género *Rhodotorula*. Los promotores para cada gen incorporado pueden ser iguales o diferentes. En algunas modalidades, los promotores son promotores encontrados en especies del género *Rhodospordium* y el género *Rhodotorula*. Los ejemplos de promotores adecuados incluyen, pero no se limitan a, promotores de los siguientes genes que codifican las siguientes proteínas: gliceraldehído 3-fosfato deshidrogenasa (GPD), proteína portadora de acil-CoA (ACP), ácido graso desaturasa, factor de alargamiento de la traducción (TEF), piruvato decarboxilasa (PDC), enolasa (2-fosfoglicerato deshidratasa) (ENO), peptidilprolil isomerasa (PPI), acetil-CoA carboxilasa (ACC) o transaldolasa. En otras modalidades, los genes descritos en la presente descripción incluyen además un terminador transcripcional de ARNm que puede encontrarse en cualquier especie eucariota y los virus de ADN.

En otro aspecto, la presente invención proporciona un método para producir ácidos grasos poliinsaturados (PUFA) omega-3 y omega-6 que comprende cultivar una célula huésped fúngica descrita en la presente descripción en condiciones adecuadas para producir los PUFA. Puede usarse cualquier medio con al menos una fuente de carbono al 5 %, como glucosa, manosa, glicerol, sacarosa. Un ejemplo del medio es Medio MinLG que contiene 30-100 g de glucosa, 1.5 g de extracto de levadura, 0.5 g de $(\text{NH}_4)_2\text{SO}_4$, 2.05 g de K_2HPO_4 , 1.45 g de KH_2PO_4 , 0.6 g de MgSO_4 , 0.3 g de NaCl , 10 mg de CaCl_2 , 1 mg de FeSO_4 , 0.5 mg de ZnSO_4 , 0.5 mg de CuSO_4 , 0.5 mg de H_3BO_4 , 0.5 mg de MnSO_4 , 0.5 mg de NaMoO_4 (por litro). El pH del medio se ajusta a 6-7. El cultivo celular se realiza a 25 °C - 32 °C.

En algunas modalidades, los genes descritos en la presente descripción que se han incorporado establemente en el genoma fúngico se unen operativamente a un promotor que permite la expresión eficiente en especies del género *Rhodospordium* y el género *Rhodotorula*. Los promotores para cada gen incorporado pueden ser iguales o diferentes. En algunas modalidades, los promotores son promotores encontrados en especies del género *Rhodospordium* y el género *Rhodotorula*. En otras modalidades, los promotores son promotores encontrados en otras especies fúngicas. Los ejemplos de promotores adecuados incluyen, pero no se limitan a, promotores de los siguientes genes que codifican las siguientes proteínas: gliceraldehído 3-fosfato deshidrogenasa (GPD), proteína portadora de acil-CoA (ACP), ácido graso desaturasa, factor de alargamiento de la traducción (TEF), piruvato decarboxilasa (PDC), enolasa (2-fosfoglicerato deshidratasa) (ENO), peptidilprolil isomerasa (PPI), acetil-CoA carboxilasa (ACC) o transaldolasa. En otras modalidades, los genes descritos en la presente descripción incluyen además un terminador transcripcional de ARNm que puede encontrarse en cualquier especie eucariota y los virus de ADN.

En algunas modalidades, un promotor adecuado es uno descrito en la publicación de solicitud de patente internacional núm. WO 2012/169969. Esta solicitud publicada describe varias secuencias polinucleotídicas derivadas de la región aguas arriba del gen de la *gliceraldehído fosfato deshidrogenasa (GPD1)*, *gen del factor de inicio de la traducción (TEF1)*, y el gen de la *estearoil-CoA-delta 9-desaturasa (FAD1)* que funcionan como promotores en hongos. Los promotores descritos en esta solicitud publicada se exponen en las SEQ ID NO:55-62. En otras modalidades, los promotores adicionales se describen en la solicitud de patente internacional núm. PCT/SG2014/000114 presentada el 10 de marzo de 2014. En una modalidad, las secuencias promotoras comprenden la secuencia expuesta en cualquiera de las SEQ ID NO:63-79. En otra modalidad, las secuencias promotoras polinucleotídicas comprenden la secuencia promotora de cualquiera de las SEQ ID NO:63-79, es decir, la secuencia sin los sitios de clonación.

Además, los fragmentos operables de las secuencias promotoras descritas en la presente descripción pueden aislarse mediante el uso de ensayos convencionales de tamizaje de promotores y pueden tamizarse para la selección eficiente de las células fúngicas transformadas mediante el uso de las técnicas descritas en la presente descripción. En una modalidad, un fragmento operable, también denominado una porción promotora en la presente descripción, tiene alrededor de 400 pares de bases hasta alrededor de 1100 pares de bases de longitud comenzando desde la posición -1 del codón ATG. Como se usa en la presente descripción "hasta" se refiere a la longitud de la porción promotora de los promotores establecidos en las SEQ ID NO descritas. Por lo tanto, "hasta" se refiere a la longitud máxima de la secuencia promotora si hay menos de 1100 nucleótidos de los promotores de las SEQ ID NO descritas.

En una modalidad, se proporciona una secuencia promotora que tiene al menos 60 % de identidad con cualquiera de estas secuencias promotoras. En otra modalidad, se proporciona una secuencia promotora que tiene al menos 70 % de identidad con cualquiera de estas secuencias promotoras. En una modalidad adicional, se proporciona una secuencia promotora que tiene al menos 80 % de identidad con cualquiera de estas secuencias promotoras. En una modalidad adicional, se proporciona una secuencia promotora que tiene al menos 90 % de identidad con cualquiera de estas secuencias promotoras. En otra modalidad, se proporciona una secuencia promotora que tiene al menos 95 % de identidad con cualquiera de estas secuencias promotoras. En otra modalidad, se proporciona una secuencia promotora que tiene al menos 98 % de identidad con cualquiera de estas secuencias promotoras.

Los genes a incorporar establemente en el genoma fúngico están típicamente en la forma de un constructo de ADN o polinucleótidos que comprende las secuencias promotoras descritas en la presente descripción, una secuencia codificante de polipéptido unida operablemente descrita en la presente descripción y una secuencia terminadora transcripcional de ARN unida operablemente. En una modalidad, puede usarse cualquier terminador transcripcional operable en las especies de los hongos. Los terminadores se ubican típicamente aguas abajo (3') del gen, después del codón de parada (TGA, TAG o TAA). Los terminadores desempeñan un papel importante en el procesamiento y la estabilidad del ARN, así como también en la traducción. La mayoría, pero no todos los terminadores, contienen una secuencia de poliadenilación o sitio de escisión. Ejemplos de secuencias específicas de poliadenilación son AAUAAA o AAUAAU. Estas secuencias se conocen como los elementos próximos aguas arriba (NUE) (Nagaya y otros, 2010). Los NUE usualmente residen aproximadamente 30 pb de una región rica en GU (Mogen y otros, 1990; Mogen y otros, 1992; Rothnie y otros 1994), conocida con elementos lejanos aguas arriba (FUE). Los FUE potencian el procesamiento en la secuencia de poliadenilación o el sitio de escisión, que usualmente es un CA o UA en una región rica en U (Bassett, 2007). Dentro del terminador, existen elementos que aumentan la estabilidad del ARN transcrito (Ohme-Takagi y otros, 1993; Newman y otros, 1993; Gutiérrez y otros, 1999) y que también pueden controlar la expresión génica (Ingelbrecht, 1989; An y otros, 1989).

Un constructo de ADN o ácidos nucleicos que comprende un promotor operable de hongos, secuencia de ADN codificante de proteína y un terminador operable de hongos puede denominarse también en la presente descripción como un casete

de expresión. El casete de expresión puede incluir otras regiones reguladoras transcripcionales como las que se conocen bien en la técnica. En otras modalidades, el constructo de ADN o ácidos nucleicos o casete de expresión comprende además un marcador seleccionable. Los marcadores seleccionables son bien conocidos por el experto ya que son casetes de expresión que incorporan tales marcadores seleccionables y promotores para conducir su expresión, tales como los descritos en la publicación de la solicitud de patente internacional núm. WO 2012/169969. Cualquier promotor adecuado unido operablemente a cualquier marcador seleccionable adecuado puede usarse en la presente invención.

En una modalidad, la secuencia codificante para el marcador seleccionable es una que existe ya sea de manera natural o se crea artificialmente y contiene al menos alrededor del 60 % de GC. En una segunda modalidad, la secuencia codificante para el marcador seleccionable es una que existe ya sea de manera natural o se crea artificialmente y contiene alrededor del 70 % de GC. En una tercera modalidad, la secuencia codificante para el marcador seleccionable es una que existe ya sea de manera natural o se crea artificialmente y contiene alrededor del 75 % de GC. En una modalidad, al menos alrededor del 70 % de los tripletes de codones de tales secuencias codificantes terminan en C o G. En otra modalidad, más de alrededor del 80 % de los tripletes de codones de tales secuencias codificantes terminan en C o G. En una modalidad, la secuencia codificante para un marcador seleccionable es al menos el 60 % de GC, preferentemente alrededor del 70 % de GC y con la máxima preferencia alrededor del 75 % de GC en la cual al menos el 70 % de los tripletes de codones terminan en C o G, preferentemente más del 80 % de los tripletes de codones terminan en C o G. En una modalidad, tales secuencias codificantes se componen de codones UCG en al menos alrededor del 40 % de los residuos de serina (Ser) totales.

En algunas modalidades, el marcador seleccionable es parte de un sistema libre de marcadores de la recombinación. En una modalidad, el sistema libre de marcadores de la recombinación es un sistema libre de marcadores de la recombinación Cre-lox, tal como el descrito por Zuo y otros [29]. Un sistema tal es útil para producir plantas transgénicas libres de marcadores de selección, que incluyen plantas de *Jatropha* transgénicas. En algunas modalidades, el sistema libre de marcadores de la recombinación se ubica entre el promotor operable de plantas y los uno o más fragmentos de ácidos nucleicos. En esta modalidad, la eliminación del gen marcador mediante el evento de recombinación coloca el promotor operable de plantas en unión operable con el uno o más fragmentos de ácidos nucleicos como se describe en la presente descripción.

En la preparación del constructo de ácido nucleico o un casete de expresión, los diversos fragmentos de ADN pueden manipularse, para proporcionar las secuencias de ADN en la orientación apropiada y, según sea apropiado, en el marco de lectura apropiado. Hacia este fin, los adaptadores o enlazadores pueden emplearse para unir los fragmentos de ADN o pueden estar implicadas otras manipulaciones para proporcionar sitios de restricción convenientes, eliminar ADN superfluo, eliminar sitios de restricción, o similares. Para este propósito, pueden estar implicadas la mutagénesis *in vitro*, reparación con cebadores, restricción, apareamiento, resustituciones, por ejemplo, transiciones y transversiones.

Los ácidos nucleicos de la presente invención también pueden sintetizarse, ya sea completamente o en parte, especialmente cuando sea deseable proporcionar secuencias preferidas de plantas, mediante métodos conocidos en la técnica. Por lo tanto, todo o una porción de los ácidos nucleicos de la presente invención pueden sintetizarse mediante el uso de codones preferidos por un huésped seleccionado. Los codones preferidos de especies pueden determinarse, por ejemplo, a partir de los codones usados más frecuentemente en las proteínas expresadas en una especie de huésped particular. Otras modificaciones de las secuencias de nucleótidos pueden dar como resultado mutantes que tienen actividad ligeramente alterada.

Puede ser útil generar un número de hongos individuales transformados con cualquier constructo recombinante para recuperar hongos libres de cualquier efecto posicional. También puede ser preferible seleccionar los hongos que contienen más de una copia del constructo de polinucleótido introducido de manera que se obtengan niveles altos de expresión de la molécula recombinante.

Puede ser conveniente producir líneas fúngicas que sean homocigóticas para un gen particular si es posible en las especies particulares. En algunas especies esto se logra mediante el uso de cultivos monospóricos. Al usar estas técnicas, es posible producir una línea haploide que porta el gen insertado y después duplicar el número de cromosomas de manera espontánea o mediante el uso de colchicina. Esto da lugar a un hongo que es homocigótico para el gen insertado, que puede evaluarse fácilmente si el gen insertado porta un gen de marcador de selección adecuado para la detección de los hongos que portan el gen. Alternativamente, los hongos pueden autofertilizarse, lo que conduce a la producción de una mezcla de esporas que consiste en, en el caso más simple, tres tipos, homocigotos (25 %), heterocigotos (50 %) y nulos (25 %) para el gen insertado. Aunque es relativamente fácil clasificar los hongos nulos de aquellos que contienen el gen, es posible en la práctica clasificar los hongos homocigotos de los heterocigotos mediante análisis de la transferencia Southern en el que se presta atención cuidadosa a cargar las cantidades exactamente equivalentes de ADN de la población mezclada, y se clasifican los heterocigotos por la intensidad de la señal de una sonda específica para el gen insertado. Es recomendable verificar los resultados del análisis de la transferencia Southern permitiendo que cada transformante independiente se autofertilice, dado que puede obtenerse evidencia adicional para homocigosis por el simple hecho que si los hongos eran homocigotos para el gen insertado, todas las líneas fúngicas subsecuentes del propio individuo contendrán el gen, mientras que si el hongo era heterocigoto para el gen, la generación que crece de la propia semilla contendrá líneas fúngicas nulas. Por lo tanto, con la simple autofertilización se puede seleccionar líneas fúngicas homocigotas que también pueden confirmarse por el análisis de la transferencia Southern.

La creación de líneas parentales homocigotas hace posible la producción de hongos y esporas híbridos que contendrán un componente proteico modificado. Las líneas parentales homocigotas transgénicas se mantienen con cada parental que contiene la primera o segunda secuencia de ADN recombinante unida operablemente a un promotor. También se incorpora en este esquema las ventajas de cultivar un cultivo híbrido, que incluye la combinación de los rasgos más valiosos y vigor híbrido.

La práctica de la presente invención emplea, a menos que se indique lo contrario, técnicas convencionales de química, biología molecular, microbiología, ADN recombinante, genética, inmunología biología celular, cultivo celular y biología transgénica, que están dentro del estado de la técnica. Ver, por ejemplo, Maniatis y otros, 1982, Molecular Cloning (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York); Sambrook y otros, 1989, Molecular Cloning, 2da Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York); Sambrook y Russell, 2001, Molecular Cloning, 3ra Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York); Green y Sambrook, 2012, Molecular Cloning, 4ta Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York); Ausubel y otros, 1992, Current Protocols in Molecular Biology (John Wiley & Sons, incluidas las actualizaciones periódicas); Glover, 1985, DNA Cloning (IRL Press, Oxford); Russell, 1984, Molecular biology of plants: a laboratory course manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.); Anand, Techniques for the Analysis of Complex Genomes, (Academic Press, Nueva York, 1992); Guthrie y Fink, Guide to Yeast Genetics and Molecular Biology (Academic Press, Nueva York, 1991); Harlow y Lane, 1988, Antibodies, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York); Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); Transcription And Translation (B. D. Hames & S. J. Higgins eds. 1984); Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells And Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the treatise, Methods In Enzymology (Academic Press, Inc., N.Y.); Methods In Enzymology, Vols. 154 y 155 (Wu y otros eds.), Immunochemical Methods In Cell And Molecular Biology (Mayer y Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volúmenes I-IV (D. M. Weir y C. C. Blackwell, eds., 1986); Riott, Essential Immunology, 6ta Edición, Blackwell Scientific Publications, Oxford, 1988; Fire y otros, ARN Interference Technology: From Basic Science to Drug Development, Cambridge University Press, Cambridge, 2005; Schepers, RNA Interference in Practice, Wiley-VCH, 2005; Engelke, RNA Interference (RNAi): The Nuts & Bolts of siRNA Technology, DNA Press, 2003; Gott, RNA Interference, Editing, and Modification: Methods and Protocols (Methods in Molecular Biology), Human Press, Totowa, NJ, 2004; Sohail, Gene Silencing by RNA Interference: Technology and Application, CRC, 2004.

EJEMPLOS

La presente invención se describe como referencia a los siguientes Ejemplos, que se ofrecen a modo de ilustración y no pretenden limitar la invención de ninguna manera. Se utilizaron técnicas estándar bien conocidas en la técnica o técnicas específicamente descritas a continuación.

Ejemplo 1

Cepas, químicos, medios y condiciones de cultivo

Las cepas de *R. toruloides* ATCC 10657 y ATCC 10788; la cepa de *R. glutinis* ATCC 90781 y la cepa de *R. glutinis* ATCC 204091 se adquirieron de ATCC (Estados Unidos). La cepa de *R. graminis* WP1 y la *Sporobolomyces roseus* FGSC 10293 (LAM13481) se obtuvieron de Fungal Genetics Stock Center (Universidad de Missouri, Estados Unidos). La cepa de *A. tumefaciens* AGL1 [30] se usó para la transformación mediada por *Agrobacterium tumefaciens* (ATMT). La higromicina B se compró de Roche Diagnostics (Estados Unidos). Las membranas de nailon N y N+ (Φ 82 mm, 0.45 μm) se obtuvieron de GE Healthcare (Uppsala, Suecia). La cerulenina (Sigma-Aldrich, Estados Unidos) se produjo como solución base de 5 mg/ml en DMSO. Otros químicos se adquirieron de Sigma-Aldrich a menos que se indique lo contrario.

Las cepas de *Rhodospiridium* se mantuvieron a 28 °C en caldo YPD (extracto de levadura al 1 %, peptona al 2 %, glucosa al 2 %) o en agar de patata-dextrosa sólido (PDA). *A. tumefaciens* se cultivó a 28 °C en medio 2YT ya sea líquido o sólido (triptona al 1.6 %, extracto de levadura al 1 %, NaCl al 0.5 %). *R. toruloides* se cultivó en medio de acumulación de lípidos a 30 °C con agitación constante (200 rpm) como se describió anteriormente [21] con algunas modificaciones. El medio MinLG contiene (por litro) 30 g de glucosa, 1.5 g extracto de levadura, 0.5 g de (NH₄)₂SO₄, 2.05 g de K₂HPO₄, 1.45 g de KH₂PO₄, 0.6 g de MgSO₄, 0.3 g de NaCl, 10 mg de CaCl₂, 1 mg de FeSO₄, 0.5 mg de ZnSO₄, 0.5 mg de CuSO₄, 0.5 mg de H₃BO₃, 0.5 mg de MnSO₄, 0.5 mg de NaMoO₄ (pH6). Para el análisis de la expresión génica, se usó una variante con limitación de nitrógeno, MinLG-N que se modifica a partir de MinLG con las concentraciones de extracto de levadura y de sulfato de amonio reducidas a 0.3 y 0.1 g/l respectivamente. El proceso de acumulación de lípidos se realizó a 30 °C con agitación constante (200 rpm).

Ejemplo 2

Constructos de ADN

Los oligonucleótidos usados se enumeran en la Tabla 1. Todas las enzimas de restricción y modificación se adquirieron de New England Biolabs (NEB, Massachusetts, Estados Unidos). El vector binario pEX2 es un derivado de pPZP200 usado para la selección dominante mediante el uso de higromicina B [22].

5 Tabla 1

Secuencias de oligonucleótidos			
Nombre	Secuencia (5'-3') (SEQ ID NO:)	Sitio de restricción	Propósito
Tnos-Sf	TTTCCGCGGTCGAATTTCCCGCATCGTTCA (13)	<i>Sma</i> I	Terminador Tnos
Tnos-Pmr	GAGTCGCTCACCTACTGCATC (14)	<i>Pme</i> I	
35T-Pmr	TTTGTTAAACATGCTAAATTCGGGGGATCTG (15)	<i>Pme</i> I	
Rt025	CACGTCGACTGAAACGCAG (16)		
Rt026	GACGAGGTCATCCGCGAG (17)		
Rt027	GTGCGGGTCGTGATGGAC (18)		
Rt028	CTGGAAGGCGGTACGAGGAC (19)		
Rt033	GTCGAAGCCGCCAGGCTGTC (20)		
Rt034	GGATCCGCCAAGTCGCGCAG (21)		
HptRU	GAGTCGCTCACCTACTGCATC (22)		PCR de colonias de <i>hpt-3</i>
HptRSL	GAGAACTCGCCGATGTCGAG (23)		
HptRSL 2	AGCGACTGGTAGAGCTGGTC (24)		
Natf	CTCTTGACGACACGGCTTAC (25)		PCR de colonias del gen de la nourseotricina acetiltransferasa
Natr	CAGGCCGCAGAGGGTGAAC (26)		
Rsp1	GATGAGTTTGGACAAACCAACTAG (27)		HiTAIL PCR
Rsp2	GGTTCAGGGGAGATGTGGGAG (28)		
Rsp3	GTACCGGCGCGCCACCTG (29)		
HRSP1	GAATCCTGTTGCCGGTCTTGCGATG (30)		
HRSP2	TTATGATTAGAGTCCCGCAATTATACA (31)		
HRSP3	CTAGCTTAGCTTGAGCTTGATC (32)		
HRRSP1	GTGCTGACGCGGGCATAGCCCCAG (33)		
HRRSP2	ATGCGACTAAACACACGCGACAAGA (34)		

HRRSP3	AGCAGCGGAGGGGTTGGATC (35)		
LAD1-1	ACGATGGACTCCAGAGCGGCCGC(GCA)N(GCA)NNNGGAA (36)		
LAD1-4	ACGATGGACTCCAGAGCGGCCGC(GCT)(GAT)N(GCT)NNCG GT (37)		
AC1	ACGATGGACTCCAGAG (38)		
Gas	CATACACCGGGCAAAGCAG (39)		
Rt055N	TITCCATGGCCAGCAGCGCACGC (40)	NcoI	
Rt056Ev	TTTGATATCAGGCGATAATATTGAGCTCC (41)	EcoRV	
Rt057N	TITCCATGGCGGGAGTGAAGGGTTGCC (42)	NcoI	
Rt058Ev	TTTGAGATCCTACTGCGCCTGCTGCTCC (43)	EcoRV	
PgpdR-Sf	TTTACTAGTCTCTTCAGACGGCTTGTCTC (44)	SpeI	
AID1Lf	CACCCGTCCTCTCCGCTTC (45)		
ALD1Rr	CCTCGCTCTTCGCTGGTTC (46)		
Rt113	CGCCCAATAACCTCACCTCAG (47)		
Rt114	GGCGATGGAGCGTAGAATAC (48)		
Rt148	CAGGTTTCATCGCAACTACATTGA (49)		
Rt149	AACAGAGCGAGTTGAAGAGTAGCC (50)		
Rt113	CGCCCAATAACCTCACCTCAG (51)		
Dga1-1	GAAGAAGACGCCGAGTAGGATG (52)		

Se han descrito anteriormente diversos promotores, tales como *gpd1* de *U. maydis* (*Pgpd*, 595 pb de longitud) [31, 32], *gpdA* de *Aspergillus nidulans* (*PgpdA*, 884 pb) [33], gen del factor de alargamiento traduccional 1 α de *Ashbya gossypii* (*Ptef*, 348 pb) [34] y *RtGPD1* (1429 pb) [22]. Los fragmentos de ADN promotor se obtuvieron mediante PCR mediante el uso de ADN plasmídico como plantilla y el par de cebadores *Pgpd*-Sf/*Pgpd*-Nr, *PgpdA*-Sf/*PgpdA*-Nr, *Ptef*-Sf/*Ptef*-Nr y *Rt011S*/*Rt012N* para *Pgpd*, *PgpdA*, *Ptef* y *PRtGPD1* respectivamente. Los fragmentos de PCR resultantes se digirieron con *SpeI* y *NcoI* y se usaron individualmente para la ligación de 3 fragmentos con el fragmento de *hpt-3* sintético digerido con *BspHI*/*SmaI* de 1030 pb [22] y el vector pEC3GPD-GUS digerido con *SpeI*/*SacI* (con extremos romos) de 8855 pb (Figura 1), para generar pEC3GPD-HPT3, pEC3GPDA-HPT3, pEC3TEF-HPT3 y pEC3GPDR-HPT3, respectivamente (Figura 1).

Para generar mutantes inactivos de *ALD1* y *DGA1*, se amplificaron secuencias codificantes completas y parciales (3 kb y 2.8 kb para *ALD1* y *DGA1*, respectivamente) mediante el uso de ADN total de *R. toruloides* ATCC 10657 como la plantilla y los pares de oligos *ALD1Lf*/*ALD1Rr* y *Rt113*/*Rt114* como cebadores respectivamente. Los productos de PCR con extremos romos se ligaron al vector pEX2 digerido doblemente con *PmeI*/*SacI* después del tratamiento con T4 ADN polimerasa en la presencia de dNTP para generar los plásmidos intermedios pEX2ALD1 y pEX2DGA1, a los cuales se insertó el casete de resistencia a higromicina con extremos romos *P_{GPD1}::hpt-3::Tnos* amplificados a partir del plásmido pRH2031 en el sitio de *XhoI*/*BspHI* y *SmaI*/*SpeI* respectivamente para generar los plásmidos de transformación de genes pKOALD1 y pKODGA1.

El gen de la diacil glicerol acil-transferasa *DGA1* (número de acceso de GenBank AB453835) y el gen de la enzima málica mitocondrial *MAE1* (etiqueta del locus RTG_03106 en el andamio genómico GL989657 de *Rhodotorula glutinis* ATCC 204091) se amplificaron mediante el uso de la plantilla de ADNc de *R. toruloides* ATCC 10657 y *R. glutinis* ATCC 204091, respectivamente. Se usaron los pares de cebadores *Rt055N*/*Rt056Ev* y *Rt057N*/*Rt058Ev* para la amplificación de *DGA1* y *MAE1*, respectivamente. Ambos productos de PCR se digirieron con *NcoI* y *EcoRV*, se ligaron con pRH2034 digerido doblemente con *NcoI*/*EcoRV*, el cual contenía un casete de expresión de proteínas que contenía el promotor *RtGPD1* de 795 pb y el terminador del gen de *35S* del virus del mosaico de coliflor y un casete de selección con higromicina *Umgpd::HPT-3:nos* específico de Cre-recombinasa [22] para generar pRHGA1 y pRHMAE1 (Figura 1).

Para estudios de transformación por ingeniería genética en ácido α -linolénico, los genes optimizados con codones que codifican la $\Delta 12$ desaturasa MaFAD2 de *Mortierella alpine* (SEQ ID NO:5), omega-3 desaturasa LuFAD3 de *Linum usitatissimum* (número de acceso de GenBank ABA02173.1; SEQ ID NO:10) y omega-3 desaturasa VfFAD3 de *Vernicia fordii* (conocido además como *Aleurites fordii*) (SEQ ID NO:12) se sintetizaron de acuerdo con la preferencia de codones de *R. toruloides*, generando genes sintéticos, MaFAD2-2 (SEQ ID NO:4), LuFAD3-2 (SEQ ID NO:9) y VfFAD3-2 (SEQ ID NO:11), que se insertaron a pRH2034 bajo la regulación del promotor *RtGPD1* para generar pRHE001, pRHE002 y pRHE003, respectivamente (Figura 2). Para producir el casete de sobreexpresión de genes dobles *FAD2-FAD3*, los casetes VfFAD3-2 y LuFAD3-2 se amplificaron a partir de los plásmidos pRHE002 y pRHE003 mediante el uso de los oligos *Rt012Sf*/*35T-Pmr*, se digirieron con *SpeI* (con extremos romos) y *PmeI* y se ligaron con pRHE001 cortado con *PmeI* para generar los plásmidos pRHE004 y pRHE005, respectivamente (Figura 3). De manera similar, el casete de *MAE1* se amplificó a partir de pRHMAE1, se digirió con (*SpeI*-*PmeI*), con extremos romos y posteriormente se ligó con pRHGA1 cortado con *PmeI* para generar el plásmido pRHE006 (Figura 3).

Ejemplo 3

Transformación mediada por *Agrobacterium tumefaciens*

Los vectores binarios se electroporaron en AGL1 de *A. tumefaciens* (2.5 kV, 25 μ F, 400 Ω) y posteriormente se seleccionaron con medio de agar 2YT complementado con estreptomycin (100 μ g/ml). La transformación de hongos a través de ATMT se realizó como se describió anteriormente a menos que se indique lo contrario [22].

Ejemplo 4

Análisis de la transferencia Southern

El ADN genómico de *R. toruloides* se extrajo como se describió anteriormente [22]. El ADN genómico se digirió con *PstI* y se separó mediante electroforesis en geles de agarosa al 0.8 % y la sonda marcada con DIG del fragmento del gen *hpt-3* parcial (de 375 a 1036 nt) se amplificó mediante el uso de los oligos *HptRU* y *HptRSL2*. Para el análisis de delección de genes, los ADN genómicos se digirieron con *HincII*, *PstI* y *HincII* para los mutantes inactivos putativos Δ ald1 y Δ dga1, respectivamente. Las sondas marcadas con DIG de aproximadamente 0.6 kb de la secuencia flanqueante aguas arriba de *ALD1* y *DGA1* se amplificaron mediante el uso de los oligos *Rt148*/*Rt149* y *Rt113*/*Dga1-1* respectivamente. La hibridación Southern se llevó a cabo de acuerdo con las instrucciones del fabricante (Kit II iniciador del marcaje y detección de ADN con sondas con alto contenido de DIG, Roche Diagnostics).

Ejemplo 5

PCR de transcripción inversa cuantitativa (q-RT-PCR)

El ARN total de *R. toruloides* se extrajo como se describió anteriormente [22]. Para eliminar la traza de ADN contaminante, las muestras de ARN se trataron con DNasa I (Roche Diagnostics, Estados Unidos) seguido de la precipitación con etanol. El ADNc se sintetizó mediante el uso del sistema de transcripción inversa Improm-II (Promega, Estados Unidos) y la PCR en tiempo real se realizó en la máquina de PCR en tiempo real iCycler™ (Bio-Rad, Estados Unidos) mediante el uso de la Supermezcla de qPCR Platinum SYBR-Green (Invitrogen, Estados Unidos). Las condiciones en tiempo real fueron las siguientes: una etapa de desnaturalización inicial a 95 °C durante 2 min seguido de 35 ciclos de desnaturalización a 95 °C durante 15 s, apareamiento a 58 °C durante 15 s y etapa de extensión a 72 °C durante 15 s. Los datos se adquirieron mediante el uso del software iCycler™ (Bio-Rad). El nivel de expresión del ARNm de *RtGPD1* se usó como la referencia para la normalización de la expresión del gen diana.

Ejemplo 6

Identificación de las posiciones de etiquetado del T-ADN

Las posiciones de las etiquetas del T-ADN en el genoma se identificaron mediante el uso de la PCR térmica de entrelazado asimétrico de alta eficiencia (hiTAIL-PCR) [35, 36]. Los cebadores específicos (HRSP1, HRSP2 y HRSP3) y el cebador arbitrario LAD1-4 se usaron para las secuencias flanqueantes del borde izquierdo (BI) del T-ADN mientras que los cebadores específicos (HRRSP1, HRRSP2 y HRRSP3) y el cebador arbitrario LAD1-4 se usaron para las secuencias flanqueantes del borde derecho (BD). Las reacciones PCR se llevaron a cabo con ADN polimerasa i-Taq (i-DNA, Singapore) en un Controlador Térmico Programable PTC-200™ (Bio-Rad, Estados Unidos). Los productos de PCR se purificaron mediante el uso del kit de extracción de gel (Qiagen) y se secuenciaron directamente mediante el uso del kit terminador BigDye (Applied Biosystems, Estados Unidos) con el oligo HRRSP3 (para BD) o HRSP3 (BI). En algunos casos, los productos de PCR se clonaron en el vector pGTM-T easy (Promega, Estados Unidos) y se secuenciaron mediante el uso de los oligos M13FP y M13RP como cebadores.

Ejemplo 7

Tamizaje de mutantes de acumulación de lípidos

El genoma de *R. toruloides* ATCC 90781 se sometió a mutagénesis mediante la inserción aleatoria del T-ADN de pRH201 (Figura 1E). Los transformantes se seleccionaron en medio de agar YPD complementado con 300 µg/ml de cefotaxima, 150 µg/ml de higromicina y ya sea 50 µg/ml de cerulenina, 10 µg/ml de violeta de tetrazolio o 0.5 µg/ml de rojo nilo (Sigma, Estados Unidos). Después de incubarse a 28 °C durante 5 días, los transformantes que muestran tamaños más grandes (por la selección contra cerulenina), pigmentación de color púrpura más oscuro (por la selección contra violeta de tetrazolio) o mayor intensidad de fluorescencia (por el tamizaje en rojo Nilo) se transfirieron a medio YPD líquido (300 µg/ml de cefotaxima, 150 µg/ml de higromicina) para la propagación. Después de frotar en placas de PDA complementadas con los antibióticos anteriores, se usaron colonias únicas para el tamizaje secundario mediante el uso de 50 ml de cultivos para verificar los fenotipos esperados.

Ejemplo 8

Comparación de los niveles de acumulación de lípidos mediante tinción con rojo Nilo

La tinción con rojo Nilo para la estimación rápida del contenido de lípidos se realizó como se describió anteriormente [37] con algunas modificaciones. En resumen, 10 µl de cultivo celular y 2 µl de solución base de rojo Nilo (50 mM en acetona) se mezclaron con 200 µl de tampón PBS (pH7.4) en un pocillo de una placa FluoroNunc (Thermo Fisher Scientific, Langenselbold, Alemania). Cada muestra se acompañó con un pocillo libre de rojo Nilo como el control de fondo. Otra fracción del cultivo celular (10 µl) se cargó a 90 µl de tampón PBS (pH7.4) en una placa transparente de fondo plano de 96 pocillos (Nunc, Roskilde, Dinamarca) para medir la densidad óptica celular. Los datos se adquirieron y se analizaron mediante el uso del lector de placas Infinite M200 (Tecan, Salzburg, Austria) mediante el uso del software iControl™ versión 3.0 (Tecan, Salzburg, Austria). La densidad óptica celular se leyó a 600 nm después de deducir el control de fondo mientras que la intensidad de fluorescencia se midió con longitudes de onda de excitación y emisión a 488 nm y 508 nm, respectivamente. El contenido relativo de lípidos se calcula mediante la normalización contra la absorción a 600 nm después de restar el control de fondo. En todas las pruebas, se incluyeron tanto los triplicados biológicos como estadísticos.

Ejemplo 9

Determinación del perfil de ácidos grasos mediante GCMS

El lípido total se extrajo como se describió anteriormente [38] con algunas modificaciones. Los cultivos celulares (1 ml) se sedimentaron y resuspendieron con 500 µl de solvente de extracción de lípidos (cloroformo:metanol = 2:1). Después de añadir 100 µg de perlas de vidrio (1 mm de diámetro, Sigma-Aldrich, Missouri, Estados Unidos), se aplicó agitación vigorosa a la mezcla durante 10 min y la fase de solvente se eliminó con una pipeta. La preparación de ésteres metílicos de ácidos grasos (FAME) y los análisis de cromatografía de gas-líquido (GC) se realizaron como se describió anteriormente [39] con algunas modificaciones. Los lípidos se evaporaron por rotación casi hasta la sequedad

(Concentrator, Eppendorf, Estados Unidos), se disolvieron en 1 ml de metanol con H₂SO₄ al 5 % (vol/vol) y se incubaron en un vial de vidrio sellado a 90 °C durante 2 hr. Los ésteres metílicos de ácidos grasos se extrajeron con 300 µl de n-hexano después de la adición de 1 ml de PBS en agua. Se inyectó 1 µl de la extracción con hexano en una columna capilar de sílice fundida DB-WAX (30-m de longitud, 0.25-µm de diámetro, y 0.25-mm de espesor de película) (Agilent J&W Scientific, Folsom, CA, Estados Unidos) en un espectrómetro de masas acoplado a cromatógrafo de gases (GCMS QP2010, Shimadzu, Japón). Las condiciones de corrida fueron típicamente 42.3 ml/min de flujo de nitrógeno, 180 °C para temperatura de inicio (3 min), una rampa de 15-min a 240 °C, y mantenimiento a 240 °C durante 7 min. Los picos de ésteres metílicos de ácidos grasos se identificaron mediante la búsqueda contra la biblioteca de compuestos NIST08 de Shimadzu y se cuantificaron como porcentajes de ácidos grasos totales (% de TFA).

Ejemplo 10

Identificación de *ALD1* mediante Tamizaje Directo

de bibliotecas de inserción del T-ADN de *R. toruloides*

Se conoce que el T-ADN se integra en los genomas nucleares predominantemente como copias únicas y esta característica se ha explotado ampliamente como una herramienta de mutagénesis en plantas y hongos [32, 40-43]. Para investigar si los nuevos genes que regulan el rendimiento o la calidad del aceite pueden identificarse mediante el tamizaje directo de bibliotecas de mutantes de T-ADN, se diseñaron tres estrategias de tamizaje independientes con el objetivo de identificar cambios en los perfiles o contenidos de ácidos grasos en los mutantes de T-ADN mediante el uso de la ayuda de fármacos o de colorante fluorescente.

La cerulenina, (2S)(3R)2,3-epoxi-4-oxo-7,10-dodecadienoilamida, es un fármaco aislado del caldo de cultivo de *Cephalosporium caeruleum* [44, 45] y se ha usado exitosamente para mejorar la acumulación de ácidos grasos poliinsaturados intracelulares [46] o el contenido de lípidos en microorganismos oleaginosos [47, 48]. Cuando se usa como un fungicida debido a su capacidad de bloquear la biosíntesis de ácidos grasos [49], se espera que los mutantes que sobrevivan a este tratamiento tengan un mayor nivel de lípidos o ácidos grasos poliinsaturados. Se tamizaron ~10,000 transformantes contra 50 µg/ml de cerulenina en medio de agar YPD y se encontraron 12 mutantes que parecían ser más resistentes a cerulenina. Estos *Mutantes de Cerulenina* de *Rhodospiridium* putativos se nombraron RCM1 a RCM12, respectivamente. Aunque se encontró que los contenidos de lípidos eran poco diferentes del tipo silvestre en cultivos líquidos a pequeña escala, los RCM exhibieron niveles de ácido α-linolénico (ALA) significativamente superiores (Figura 4). Notablemente, RCM6 produjo un nivel de ALA 3 veces superior al del tipo silvestre (Figura 4B).

En segundo lugar, el rojo nilo (NR) se ha usado ampliamente como un rastreador fluorescente de lípidos [50]. Mediante el tamizaje de ~10,000 mutantes de T-ADN, se identificaron cuatro candidatos que parecen mostrar una fluorescencia roja más fuerte, que se nombraron RNM1-4 (para *Mutante de rojo Nilo* de *Rhodospiridium*). La cuantificación de los rendimientos de lípidos reveló una mejora significativa en los mutantes RNM en comparación con el tipo silvestre (Figura 4D). Sin embargo, podrían observarse pequeñas diferencias en las composiciones de ácidos grasos (datos no mostrados).

De manera similar, violeta de tetrazolio, que se usa ya sea como un indicador colorante para la acumulación de lípidos [51], o como un indicador redox para el crecimiento microbiano [52] se usó como un indicador para tamizar ~ 3,000 transformantes, lo que condujo a la identificación de 6 mutantes pigmentados más profundos (Figura 4C). Sin embargo, el análisis repetido de los perfiles de ácido graso no permitió verificar el cambio de la acumulación de ácidos grasos (datos no mostrados).

Las posiciones de etiquetado del T-ADN en los mutantes anteriores se identificaron mediante la técnica de PCR Hi-TAIL, y los resultados demostraron que se obtuvieron y se secuenciaron exitosamente 11 de 12 RCM, 2 de 4 RNM y 6 de 6 RTM (Tabla 2). Los genes afectados se implicaron de manera dominante en el mantenimiento de la integración de la pared celular, el metabolismo de los lípidos, la transducción de señales, el plegamiento y el tráfico de proteínas, los metabolismos de los metabolitos secundarios, los aminoácidos, las vitaminas, los cofactores, etc. (Tabla 2).

Tabla 2

Posiciones de etiquetado del T-ADN de <u>Mutantes de Cerulenina de R. turoloides</u> (RCM), <u>Mutantes de violeta de tetrazolio de R. turoloides</u> (RTM), <u>Mutantes de Rojo de Nilo de R. toruloides</u> (RNM) y <u>Mutantes Albinos de R. toruloides</u> (RAM)					
Número de secuencia ^a	Sitio génico ^{b,c}	Mejor coincidencia ^d	Anotación ^e	Organismo ^f	Identidad ^g
Secuencias del BD					
RCM1	Aguas arriba-0.5 kb	XP_001549261	como mesotelina	<i>Botryotinia fuckeliana</i>	26 %
RCM2	Aguas arriba-1.0 kb	5545759 ID: Kpol_534p16	manosiltransferasa	<i>Vanderwaltozyma polyspora</i>	77 %
RCM3	Aguas arriba-0.5 kb	XP_001789963	regulador de los transcritos sin sentido; isocitrato deshidrogenasa dependiente de NADP	<i>Bos taurus</i>	40 %
RCM4	Secuencia génica	XP_001607008	pipsqueak	<i>Nasonia vitripennis</i>	29 %
RCM5	Aguas arriba-0.5 kb	XP_003174510	proteína que contiene dominios de dedos de zinc C6	<i>Arthroderma gypseum</i>	32 %
RCM6	Secuencia génica	XP_001629556	Ácido aldehído deshidrogenasa	<i>Nematostella vectensis</i>	51 %
RCM7	Secuencia génica	XP_571856	Proteína relacionada con el transporte de hexosas	<i>Cryptococcus neoformans</i>	53 %
RCM8	Secuencia génica	XP_001731990	Subunidad 2 del factor de la iniciación de la transcripción TFIID	<i>Malassezia globosa</i>	40 %
RCM9	Secuencia génica	NP_001125572	proteína de transferencia de lípidos relacionada con stAR 3	<i>Pongo abelii</i>	29 %
RCM10	Aguas arriba-1.0 kb	XP_001645395	GPI manosiltransferasa 3	<i>Vanderwaltozyma polyspora</i>	77 %
RCM11	Aguas arriba-0.5 kb	XP_662119	Proteína de la pared celular que funciona en la transferencia de quitina a beta (1-6) glucano	<i>Aspergillus nidulans</i>	36 %
RCM12	NA ^h	-	-	-	-
RTM1	Secuencia génica	EGU12290	Proteofosfoglicano ppg4	<i>Rhodotorula glutinis</i> ATCC 204091	93 %
RTM2	Aguas arriba 0.5 kb	EGU13095.1	salicilato hidroxilasa	<i>Rhodotorula glutinis</i> ATCC 204091	73 %
RTM3	Génica	XP_501740.1	YALI0C11858p	<i>Yarrowia lipolytica</i>	71 %
RTM4	Aguas arriba 0.5 kb	ZP_08453184.1	oxidorreductasa de unión a zinc putativa	<i>Streptomyces sp.</i>	47 %
RTM5	Génica	ZP_07628725.1	lipoproteína putativa	<i>Prevotella amnii</i>	45 %
RTM6	Génica	YP_001220603.1	recombinasa específica de sitio resolvasa	<i>Aeromonas bestiarum</i>	94 %
RNM1	Aguas abajo 0.3 kb	AAC98967.2	ácido graso omega-3 desaturasa	<i>Vemicia fordii</i>	41 %
RNM2	NA				

5	RNM3	Aguas arriba 0.5 kb	ZP_03104366	aminoácido permeasa	<i>Bacillus cereus W</i>	87 %
	RNM4	NA				
10	RAM1	Secuencia génica	XP_003032296	Transportador de riboflavina MCH5	<i>Schizophyllum commune</i>	52 %
	RAM2	Aguas arriba-0.5 kb	YP_001220603	resolvasa	<i>Aeromonas bestiarum</i>	95 %
15	RAM3	Secuencia génica	XP_571856	proteína relacionada con el transporte de hexosas	<i>Cryptococcus neoformans</i>	36 %
	RAM4	Secuencia génica	XP_758766	factor Taf2 asociado a proteína de unión a TATA (TBP) (MTCC 457 contig458_1:18376-18377+)	<i>Ustilago maydis</i>	35 %
20	RAM5	Secuencia génica	KF601426.1	fitoeno sintasa	<i>Rhodosporidium diobovatum</i>	98 %
	RAM6	Aguas abajo-0.2 kb	EGU11996	Serina/treonina quinasa	<i>Rhodotorula glutinis ATCC20409I</i>	56 %
25	RAM7	Secuencia génica	XP_001629556	Ácido aldehído deshidrogenasa	<i>Nematostella vectensis</i>	51 %
	^a Secuencias flanqueantes del BI ^b Los genes con etiqueta de T-ADN se determinaron de acuerdo con los resultados de BLASTx ^c Aguas arriba 1.0 kb, Aguas arriba 0.5 kb y aguas abajo 0.3 kb indica las inserciones del T-ADN dentro de 501-1000 pb, 500 pb aguas arriba y 300 pb aguas abajo del gen etiquetado correspondiente, respectivamente ^d La mejor coincidencia indica el resultado de BLASTx con la puntuación E más alta ^e Las anotaciones se determinaron de acuerdo con los resultados de BLASTx ^f El microorganismo indica el huésped de mejor coincidencia ^g Los valores de identidad fueron de los resultados de BLASTx ^h No disponible debido al mal resultado de la secuenciación					

Ejemplo 11

Caracterización del mutante de etiquetado del T-ADN RCM6

Para investigar más aún los mutantes tamizados mediante la genética directa, y como una prueba de principio, se abordó la genética inversa para estudios del posible efecto regulador en la acumulación de lípidos y la biosíntesis de carotenoides en los mutantes RCM6 y RAM5, respectivamente.

El análisis homólogo reveló que el T-ADN en RCM6 se integró dentro de 72542-72543 nt de 400^a contig (núm. de acceso de GenBank AEVR01000400). El BLASTx de secuencias adyacentes a la posición de etiquetado del T-ADN exhibió una proteína que contiene un dominio del aldehído deshidrogenasa putativa que se alteró mediante la integración del T-ADN en RCM6 (Figura 6A). La proteína diana se localizó entre una enzima de biosíntesis de tiazol putativa (EGU11956) y una dipeptidil aminopeptidasa candidata (EGU11957), que muestran gran homología con el aldehído hidrogenasas de otras especies fúngicas tales como *Streptomyces svceus* (EDY60340.1, valor E = 2E-68) y *Mycobacterium* sp. (YP_936108, valor E = 3E-66), etc. Por lo tanto, el aldehído hidrogenasa putativa (denominada ALD1) que codifica el gen alterado por T-ADN podría dar como resultado el fenotipo de RCM6. El análisis adicional de RT-PCR y la amplificación rápida de los extremos de ADNc (RACE) revelaron que el gen *ALD1* abarca 2461 pb en contig#400, que contiene 11 exones separados por 10 intrones (Figura 5A). Su 5' UTR es de 18 nt de longitud, seguido de un tramo corto de motivo rico en CT (caja CT, datos no mostrados). El empalme del ARNm se ajusta estrictamente a la regla canónica GU-AG, que produce un ARNm de 1506 nt de longitud, que codifica una proteína de 501-aa con el motivo de huella de unión a NAD estrictamente conservado Gly-X-Gly-XX-Gly (SEQ ID NO:53) (GSGTVG, 193-198 aa; SEQ ID NO:54), aminoácido de unión a adenosina ribosa (NAD) (E₁₄₈) y centro catalítico de Cys (C₂₄₉) (Figura 6). En RCM6, el T-ADN se integró en la posición entre el nucleótido 2097mo y 2098^{vo} desde el codón de inicio de *ALD1*, alterando el 9^{no} exón lo que resulta en la delección de los 58 aa del extremo C-terminal que forman parte del motivo RYPP (Figura 6).

Para demostrar más aún la función en la acumulación de lípidos, *ALD1*se delecionó mediante la recombinación homóloga con la ayuda de ATMT. La secuencia de nucleótidos en *ALD1* en el intervalo de +536 a +1947 se reemplazó con el casete de resistencia a higromicina (*P_{GPD1}::hpt-3::Tnos*, Figura 5A) y los mutantes nulos de *ald1* correctos se verificaron mediante el análisis de la transferencia Southern (Figura 5B). Cuando se cultiva ya sea en caldo líquido o en medio de agar, *Δald1* mostró un color pigmentado naranja al contrario del color pigmentado rosa en el tipo silvestre (Figura 5E). *Δald1* creció un poco más lento que el tipo silvestre en el medio de acumulación de lípidos, pero rindió una biomasa similar cuando la

fueron fuente de carbono en los medios se agotó el 4^{to} día (Figura 5E). Antes de que se agotara la glucosa (3^{er} día), los mutantes *Δald1* mostraron pocas diferencias con el tipo silvestre en cuanto a los rendimientos de lípidos. Cuando se agotó la glucosa, los niveles lipídicos disminuyeron en ambas cepas, sin embargo, el contenido de lípidos fue significativamente superior en el mutante *Δald1* (Figura 5C y 5D). Sorprendentemente, casi la mitad de ALA se degradó en el tipo silvestre mientras que se encontró poca degradación en el mutante de *Δald1* (Figura 5F). Estos resultados son consistentes con lo del mutante con etiqueta de T-ADN, RCM6 (Figura 4B).

Ejemplo 12

Mejora de la acumulación de TAG total

Niveles en *R. toruloides* mediante diseño racional

El triacilglicerol (TAG) es el principal lípido neutro que se produce en la mayoría de las células eucariotas y las vías de biosíntesis son altamente conservadas [53]. La ingeniería metabólica mediante diseños racionales ha sido muy exitosa en mejorar el contenido y la productividad de lípidos [54-56]. En *R. toruloides*, está disponible la información de secuencia para la diacilglicerol acetiltransferasa (*Dga1*) y la enzima málica (MAE1) (ver por ejemplo las SEQ ID NO:81 y SEQ ID NO:84, respectivamente). Los casetes de sobreexpresión para ambos genes se construyeron para que sean conducidos por el promotor de *RtGPD1* de 795-pb, *P_{GPD1}::DGA1* y *P_{GPD1}::MAE1*, Figura 1D, que se integran en el cromosoma de *R. toruloides* ATCC 10657 mediante la transformación mediada por *Agrobacterium tumefaciens*. El análisis de RT-PCR cuantitativo reveló que los niveles de transcrito de ARNm de ambos genes se potenciaron dramáticamente a través del bioproceso de 3 días (Figura 7A), dio como resultado una mejora en 2.3 y 1.8 veces del pico del rendimiento de lípidos sobre la cepa de tipo silvestre con la cepa de sobreexpresión de *DGA1* y *MAE1* respectivamente (Figura 7B). El perfil de ácidos grasos no cambió significativamente en ambas cepas (Figuras 7C y 7F). Como se esperaba, un mutante nulo de *dga1* construido (Figura 7D) tuvo una disminución dramática en la acumulación de lípidos (Figura 5E).

Ejemplo 13

Mejora de ácidos grasos poliinsaturados (PUFA)

Producción en *R. toruloides* mediante diseño racional

En cepas de *R. toruloides*, el ácido oleico (C18:1) es el componente de ácido graso predominante (~ 50 %) mientras que el ácido palmitoleico (C16:1) y el ácido linoleico (C18:2) consiste en ~ 20 % de FA total. El ácido α-linolénico (C18:3n=9, ALA), un ácido graso omega-3 poliinsaturado, es un componente menor presente en el 3~4 % de los ácidos grasos totales (% de TFA) (Figura 8). Para producir ALA a partir de ácido oleico, se requiere una delta-12 desaturasa (*Fad2*) y una omega-3 desaturasa (*Fad3*) [57, 58]. Los genes de longitud completa de lino (*L. usitatissimum*) *FAD3*, *tung* (*V. fordii*) *FAD3* y *M. alpine* *FAD2* se diseñaron y se sintetizaron comercialmente de acuerdo con la preferencia en el uso de codones de *R. toruloides* [22], en los cuales se reemplazaron todos los codones raros por los que se usan frecuentemente en *R. toruloides*. El contenido de GC de los genes varió de 65.3 % en *LuFAD3-2* (SEQ ID NO:9), 64.8 % en *MaFAD2-2* (SEQ ID NO:4) y 63.3 % en *VfFAD3-2* (SEQ ID NO:11).

El alto nivel de expresión de los tres genes sintéticos se logró mediante la unión operativa al promotor *RtGPD1* y al terminador 35S del virus del mosaico de la coliflor, seguido por la determinación del perfil de las composiciones de ácidos grasos de las cepas de ATMT seleccionadas. Las cepas élites que sobreexpresan *LuFAD3-2* (RHE001), *VfFAD3-2* (RHE002) o *MaFAD2-3* (RHE003) mostraron una mejora del contenido de ALA en 1.8, 2.2 y 1.6 veces respectivamente (Figura 9A). Posteriormente, los casetes *RtGPD1::afFAD3-2:35S* y *RtGPD1::MaFAD2-2:35S* se apilaron en un único vector de T-ADN, pRHE004, el cual se transformó mediante ATMT a un derivado (RT1CE6, que contiene un gen de Cre inducible con 17β-estradiol integrado establemente en el genoma) de *R. toruloides* ATCC 10657. La cepa élite seleccionada (RHE004) se modificó más aún mediante la delección del gen *ALD1* mediante el uso del vector pKOALD1. La cepa resultante (denominada *aldleAM*) produjo 3.74 veces más ALA que el mutante nulo de *ald1* (*aldle*) (Figura 9B), con el contenido de ALA que alcanza ~ 49 % de los ácidos grasos totales.

Ejemplo 14

Análisis bioquímico de Ald1

Para caracterizar Ald1 y confirmar que la delección de 58 residuos del extremo C-terminal como resultado de la inserción de T-ADN en el mutante RCM6 comprometía su actividad enzimática, tanto la versión de longitud completa como la truncada de las proteínas Ald1 se expresaron en *E coli* BL21(DE3) como una proteína de fusión con la etiqueta de 6X histidina en el extremo C-terminal. Las Ald1 y Ald1n recombinantes se purificaron con la columna HisTrap (GE healthcare, Estados Unidos) y se analizaron mediante el uso del método informado anteriormente [59] con algunas modificaciones. En resumen, la mezcla de reacción estaba compuesta de 40 μl de tampón Tris-Cl (pH8.0) 100 mM, 30 μl de NAD⁺ o NADP⁺ 10 mM (Sigma-Aldrich, Estados Unidos), 10 μl de dodecanal 20 mM (dodecil aldehído, C12-aldehído, Sigma-Aldrich, Estados Unidos), 110 μl de agua y 10 μl de enzima purificada. La reacción se realizó a temperatura ambiente (25 °C) y se inició mediante la adición de enzima. El curso temporal del valor de densidad óptica a 340 nm se leyó a través

del lector de placas Infinite M200 (Tecan, Salzburg, Austria) mediante el uso del software iControl™ versión 3.0 (Tecan, Salzburg, Austria) como se describió anteriormente [40]. Como se muestra en la Figura 10, tanto Ald1 como Ald1n demostraron una clara actividad deshidrogenasa, con una ligera preferencia a NAD⁺. Notablemente, la proteína mutante con la delección de 58 aa del extremo C-terminal mostró actividad enzimática significativamente menor.

Ejemplo 15

Caracterización de los genes de la biosíntesis de ácidos grasos en *R. toruloides*

Los homólogos de *R. toruloides* de diversas ácidos grasos desaturasa, elongasa y ATP-citrato liasa se identificaron mediante la búsqueda de BLAST contra las secuencias de andamio 204091 del genoma de *R. toruloides* ATCC 204091 (denominadas anteriormente *Rhodotorula glutinis*, núm. de acceso de GenBank AEVR02000000, proyecto de secuenciación en disparo del genoma completo PRJNA59971, Mississippi State University, Estados Unidos.) mediante el uso de secuencias enzimáticas de *Yarrowia lipolytica* y *Ustilago maydis* conocidas según las consultas. La manipulación genética y la caracterización de secuencias de ADN se realizaron con la cepa de *R. toruloides* ATCC 10657 o su derivado Rt1ck, un mutante deficiente de *KU70* que exhibe una eficiencia extremadamente alta en la recombinación homóloga [60]. Los oligonucleótidos usados se enumeran en la Tabla 3.

Tabla 3

Oligonucleótidos usados para:		
Delección del gen de la estearoil-CoA-delta-9- desaturasa (RtFAD1)		
DS9L-Sf	GCGAGGGATGGCAGTAAGACG (118) ^a	SacI
DS9L-Br	AAAGGATCCAACCTTGCTCGCCAGTACC (119)	BamHI
DS9R-Hf	TITAAGCTTCACGTACAGCCTGTGGTAGCC (120)	HindIII
DS9R-Str	TTTAGGCCCTGGAGGAGTCGAGCGTGAGAGT (121)	StuI
Rt227Nf	TITccatggCTGCCTCGTCGGCACTCGAG (122)	NcoI
Rt228Evr	TITgatatcCATTACGCCTTGACCGTCAG(123)	EcoRV
Para la delección del gen de la delta-12 desaturasa (RtFAD2)		
DS12L-Sf2	AAAGAGCTCGGTGACTGCATGCTCCGTTAC (124)	SacI
DS12L-Br2	AAAGGATCCTGATGGAGTAGTTGGGCACGA(125)	BamHI
DS12R-Hf	TTTAAGCTTCCTCCTCCTTGATCTTTCGCCG (126)	HindIII
DS12R-Str	TTTAGGCCCTGACCTTTGCGTCCTCCCTTCA (127)	StuI
OEDS12f	TCAGAACAACACCAGATCACTCACAATGGCCGCTACCCTCCGCC (128)	
Rt229Ndf	TTTCATATGGCCGCTACCCTCCGCCA (129)	NdeI
Rt230Evr	TTTGATATCTCTAGGGCATCGTCTAGAGTC (130)	EcoRV
Para la delección del gen 1 de la elongasa (RtELO1)		
ELO1L-Sf	AAAGAGCTCTATTGTTCTGACTAGACTGCGCCAC (131)	SacI
ELO1L-Br	AAAGGATCCAAGGAGGATATTGTGCACGAGGA(132)	BamHI
ELO1R-Hf	CGACTCCAAATCACCCAGTTCCTC (133)	HindIII
ELO1R-Str	TTTAGGCCCTGACCGACTTTGACGACGAC (134)	StuI
Para la delección del gen de la ATP-citrato liasa (RtACL1)		
ACL1L-Sf2	AAAGAGCTCGCGGCACTGTACTTCACTACG (135)	SacI
ACL1L-Br2	AAAGGATCCACGAGACCTATCCAAACGC (136)	BamHI
ACL1R-Hf2	TTTAAGCTTAAGGTCAAGTCCAAGGCCAAC (137)	HindIII
ACL1R-Str2	TTTAGGCCCTGGCTGCTGGAGAAACGAAACT)138)	StuI

Para la delección del gen 2 de la elongasa (RtELO2)		
ELO2L-Stf	AAAAGGCCCTCATTCCTCGACTCGACGCAT (139)	Stul
ELO2L-Hr	AGAGGAGGAAGTTGTGCAGCA (140)	HindIII
ELO2R-Bf	TTTGGATCCGTTGCGGCGAGTCCTGTCTATC (141)	BamHI
ELO2R-Sr	TTTGAGCTCCGGAGCGAGTAAGACGAGG (142)	SacI
Para la expresión del ADNc de ELO1		
Rt236Nf	TTTCCATGGCCTCGTACGCCGCCCATCC (143)	NcoI
Rt237Evr	TTTGATATCGGAAGGGACCGCGCTAGTIG (144)	coRV
Para la expresión del ADNc de ELO2		
Rt259Nf	TTTCCATGGTTCGCACCGTCCCCCG (145)	NcoI
Rt260Evr	TTTGATATCGAGAAGAAGGTGGGGTGTITAG (146)	EcoRV
^a SEQ ID NO:		

Para la delección del homólogo FAD1 (o OLE1) del gen de la delta-9-oleato desaturasa, el fragmento flanqueante de homología izquierdo y derecho (~0.9 kb cada uno) se amplificaron mediante el uso del ADN genómico de *R. toruloides* ATCC 10657 con los pares de oligos DS9L-Sf/DS9L-Br y DS9R-Hf/DS9R-Str, respectivamente. Se realizó una ligación de cuatro fragmentos con el vector binario pEX2 digerido con SacI/PmeI, el fragmento flanqueante izquierdo digerido con SacI/BamHI, el casete de selección con higromicina optimizado con codones digerido con BamHI/HindIII de pDXP795hptR (P_{GPD1}::hpt-3::T_{nos} [60]) y el fragmento flanqueante derecho digerido con HindIII/Stul para generar la delección génica del plásmido pKOLE1. Se aplicó una estrategia similar para construir tanto pKOFAD2 como pKOELO1, para la inactivación del gen de la delta-12 desaturasa putativa y el gen 1 de la elongasa, respectivamente. Los pares de oligo DS12L-Sf2/DS12L-Br2 y DS12R-Hf/DS12R-Str se usaron para amplificar el fragmento flanqueante de homología izquierdo (0.6 kb) y derecho (0.9 kb) para pKOFAD2, y ELO1L-Sf/ELO1L-Br y ELO1R-Hf/ELO1R-Str se usaron para pKOELO1 (~0.9 kb cada uno). Para ELO2, los pares de oligos ELO2L-Stf/ELO2L-Hr y ELO2R-Bf/ELO2R-Sr se usaron (-0.8 kb cada uno) y se digirieron con Stul/HindIII y BamHI/SacI para el fragmento flanqueante de homología izquierdo y derecho, respectivamente. Ambos fragmentos se usaron de manera similar en la ligación de cuatro fragmentos para generar pKOELO2.

Para la delección del gen de la ATP-citrato liasa putativa (RtACL1), los pares de oligos ACL1L-Sf2/ACL1L-Br2 y ACL1R-Hf2/ACL1R-Str2 se usaron para amplificar los fragmentos flanqueantes de homología izquierdo y derecho (0.9 kb cada uno) para generar pKOAACL1 mediante el uso de una estrategia similar descrita anteriormente.

Las secuencias de ADNc de los genes de interés se obtuvieron mediante RT-PCR, con la RACE 5' y 3' realizada mediante el uso del Kit de amplificación de ADNc BD SMARTer™ RACE (Clontech, California, Estados Unidos) de acuerdo con las instrucciones del fabricante. Los pares de oligos OLE1U1/OLE1L1, FAD2U1/FAD2L1 se usaron como el cebador específico para RACE 5'/3' de FAD1 (OLE1) y FAD2, respectivamente.

El ORF predicho de Fad1 (Ole1) y Fad2 codifica la proteína de 545 y 451 aa de longitud, respectivamente. Ambos Fad comparten un dominio conservado común del ácido graso desaturasa de membrana (familia de proteínas núm. pfam00487, EMBL-EBI). Sin embargo, Fad2 carece del dominio de unión a hemo/esteroide como citocromo β5 (pfam00173). Las búsquedas de BLAST revelaron que Ole1 y Fad2 exhiben la identidad más alta con respecto a la estearoil-CoA desaturasa de *Puccinia graminis* (XP_003326562.1, 70 % de identidad), Δ12-ácido graso desaturasa de *Ustilago maydis* (XP_757193.1, 57 % de identidad), respectivamente.

Se identificaron dos elongasas. ELO1 (Seq ID. No. 99 y 100) y ELO2 (Seq ID. No. 102 y 103) codifica la proteína de 329 (Seq ID 101) y 293 aa (Seq ID 104) de longitud, respectivamente. Ambas ácido graso elongasas putativas comparten un dominio conservado común de la familia GNS1/SUR4 que está implicado en el sistema de alargamiento de ácidos grasos de cadena larga (pfam01151). Elo1 y Elo2 exhibieron la identidad más alta con respecto al ácido graso elongasa de *Puccinia graminis* (PGTG06945, XP_003325743.2, 43 % de identidad) y *Melampsora larici-populina* (MELLADRAFT_42723, XP_007407925.1, 65 % de identidad), respectivamente. El análisis de los perfiles de ácidos grasos de ELO1 y ELO2 inactivos reveló que ELO1 inactivo conduce a un pequeño cambio en el perfil excepto por una disminución moderada de C18:0 y un aumento pequeño de C16:0 y C18:1. Por el contrario, el ELO2 inactivo conduce a la pérdida completa de la síntesis de ácidos graso de cadena larga (>C18) (Figura 12). Estos resultados sugieren fuertemente que OLE1 es un ácido graso de cadena corta-CoA elongasa mientras que ELO2 es capaz de alargar tanto los ácidos grasos de cadena larga y de cadena corta-CoA.

Para los estudios de sobreexpresión, el ADNc de OLE1, FAD2 y ELO2 se amplificó mediante el uso de la plantilla de ADNc de *R. toruloides* sintetizada mediante la transcripción inversa con los pares de cebadores Rt227Nf/Rt228Evr,

Rt229Ndf/Rt230Evr y Rt259Nf/Rt260Evr, respectivamente. Los productos de PCR digeridos con NcoI/EcoRV se ligaron con el vector pKC1 digerido con NcoI/EcoRV para generar los pKC1OLE1, pKC1FAD2 y pKC1ELO2 y dio como resultado la sobreexpresión del gen debido al promotor RtGPD1 fuerte usado en el vector pKC1.

5 Ejemplo 16

Análisis de delección génica

Para verificar las funciones de cada gen, los mutantes inactivos se generaron mediante la transformación mediada por *Agrobacterium* de los constructos inactivos respectivos; el tamizaje mediante PCR de colonias y el análisis de la transferencia Southern. La inactivación de *FAD1* no fue exitosa en varios intentos. La delección de *FAD2* fue exitosa después de la suplementación de ácido linoleico en los medios de transformación y propagación. El ácido linoleico (C18:2, LA) y el ácido α -linolénico (C18:3, ALA) estuvieron ausentes en el mutante nulo de *FAD2* mientras que el contenido de C18:1 aumentó a casi el 70 % de los ácidos grasos totales (Figura 11B). Esto confirma que el gen *FAD2* (SEQ ID NO:92) codifica una $\Delta 12$ -ácido graso desaturasa que cataliza la conversión de ácido oleico (C18:1) para formar LA. La carencia de ALA en el mutante inactivo de *FAD2* sugiere que Fad2 es un ácido graso desaturasa bifuncional $\Delta 12$ y $\Delta 15$. Esto fue apoyado por la incapacidad del mutante *fad2* Δ de producir la producción de ALA incluso cuando se complementa con el precursor de LA (C18:2) (Figura 11D).

La función de *FAD1* (*OLE1*) puede demostrarse mediante los estudios de sobreexpresión. La transformación del casete RtGDP1::*OLE1* (Figura 13A) en la cepa de tipo silvestre y la cepa inactiva de *ALD1* (Δ ald1e) dio como resultado el aumento significativo en el contenido de ácido oleico (Figura 14). En consecuencia, la apilación de RtGDP1::*OLE1* a Ma*FAD2*-2 (Figura 13D), o ambos a Ma*FAD2*-2 y el casete de sobreexpresión Vf*FAD3*-2 (Figura 13E) dio como resultado un aumento de LA y ALA respectivamente (Figura 14). Entre los 18 transformantes que expresaban el casete de gen triple, tres mostraron >20 % del contenido de ALA con uno que contenía ~24 % del fondo de aldle. El contenido inferior de ALA en esta serie se atribuye probablemente al promotor de *ACC1* más débil usado para Ma*FAD2*-2 a medida que el contenido de LA fue muy bajo (no mostrado).

Ejemplo 17

Caracterización de los genes de la ATP-citrato liasa (*ACL1*) en *R. toruloides*

En base a los estudios en mamíferos, se piensa que la ATP-citrato liasa (ACL) fúngica es un factor importante para la acumulación de aceites. Un gen de la ATP-citrato liasa/sintasa putativa *ACL1* (seq ID No. 86 y 87) se identificó mediante la búsqueda de BLAST de las secuencias genómicas parciales de *Rhodotorula glutinis* ATCC 204091. La secuencia de la proteína *ACL1* putativa se expone en la Seq ID. No. 88. Se generó un mutante inactivo del gen *ACL1* y mostró una reducción significativa de la acumulación de aceites y del crecimiento de la biomasa (Figura 15). Esto sugiere fuertemente que la expresión de *ACL1* facilita la acumulación de aceites y la producción de biomasa en *R. toruloides*.

40 Bibliografía

- An, G., y otros, 1989. Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. *Plant Cell*, 1: 115-122.
1. Venegas-Calderón, M., O. Sayanova, y J.A. Napier, An alternative to fish oils: metabolic engineering of oil-seed crops to produce omega-3 long chain polyunsaturated fatty acids. *Progress in lipid research*, 2010. 49(2): p. 108-119.
2. Horrobin, D., Nutritional and medical importance of gamma-linolenic acid. *Progress in lipid research*, 1992. 31(2): p. 163-194.
3. Simopoulos, A.P., The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomedicine & pharmacotherapy*, 2002. 56(8): p. 365-379.
4. Gong, Z., y otros, Efficient conversion of biomass into lipids by using the simultaneous saccharification and enhanced lipid production process. *Biotechnology for biofuels*, 2013. 6(1): p. 1-12.
5. Ratledge, C., Regulation of lipid accumulation in oleaginous micro-organisms. *Biochem Soc Trans*, 2002. 30(Pt 6): p. 1047-50.
6. Meng, X., y otros, Biodiesel production from oleaginous microorganisms. *Renewable Energy*, 2009. 34(1): p. 1-5.
7. Ward, O.P. y A. Singh, Omega-3/6 fatty acids: alternative sources of production. *Process Biochemistry*, 2005. 40(12): p. 3627-3652.
8. Beopoulos, A., y otros, *Yarrowia lipolytica* as a model for bio-oil production. *Prog Lipid Res*, 2009. 48(6): p. 375-87.
9. Katre, G., y otros, Evaluation of single cell oil (SCO) from a tropical marine yeast *Yarrowia lipolytica* NCIM 3589 as a potential feedstock for biodiesel. *AMB Express*, 2012. 2(1): p. 36.
10. Neuveglise, C., y otros, A shuttle mutagenesis system for tagging genes in the yeast *Yarrowia lipolytica* Gene, 1998. 213(1-2): p. 37-46.
11. Sabirova, J.S., y otros, The 'LipoYeasts' project: using the oleaginous yeast *Yarrowia lipolytica* in combination with specific bacterial genes for the bioconversion of lipids, fats and oils into high-value products. *Microb Biotechnol*, 2011. 4(1): p. 47-54.
12. Tai, M. y G. Stephanopoulos, Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metab Eng*, 2013. 15: p. 1-9.

13. Li, Y., Z.K. Zhao, y F. Bai, High-density cultivation of oleaginous yeast *Rhodospiridium toruloides* Y4 in fed-batch culture. *Enzyme and Microbial Technology*, 2007. 41(3): p. 312-317.
14. Zhao, X., y otros, Lipid production by *Rhodospiridium toruloides* Y4 using different substrate feeding strategies. *J Ind Microbiol Biotechnol*, 2010.
- 5 15. Pan, J.G., M.Y. Kwak, y J.S. Rhee, High density cell culture of *Rhodotorula glutinis* using oxygen-enriched air. *Biotechnology letters*, 1986. 8(10): p. 715-718.
16. Frengova, G.I. y D.M. Beshkova, Carotenoids from *Rhodotorula* and *Phaffia*: yeasts of biotechnological importance. *Journal of industrial microbiology & biotechnology*, 2009. 36(2): p. 163-180.
17. Kirk, M.P., y otros, en *Dictionary of the Fungi*. 2008, CABI: Wallingford. p. 716.
- 10 18. Hu, C., y otros, Effects of biomass hydrolysis by-products on oleaginous yeast *Rhodospiridium toruloides*. *Bioresour Technol*, 2009. 100(20): p. 4843-7.
19. Zhao, X., y otros, Lipid production from Jerusalem artichoke by *Rhodospiridium toruloides* Y4. *J Ind Microbiol Biotechnol*, 2010. 37(6): p. 581-5.
- 20 20. Wu, S., y otros, Microbial lipid production by *Rhodospiridium toruloides* under sulfate-limited conditions. *Bioresour Technol*, 2010.
21. Wu, S., y otros, Phosphate-limitation mediated lipid production by *Rhodospiridium toruloides*. *Bioresour Technol*, 2010. 101(15): p. 6124-9.
22. Liu, Y., y otros, Characterization of glyceraldehyde-3-phosphate dehydrogenase gene *RtGPD1* and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodospiridium toruloides*. *Appl Microbiol Biotechnol*, 2013. 97(2): p. 719-29.
- 20 23. Ji, L., N. Peng, y H.-I. Cheng, Polynucleotide sequences from *Rhodospiridium* and *Rhodotorula* and use thereof, en *solicitud provisional de Estados Unidos núm. 61/782,832*. 2013.
24. Liu, Y., y otros, Genetic manipulation and expression systems for *Pucciniomycotina* and *Ustilaginomycotina* subphyla, en el documento de patente WO 2012169969. 2011.
- 25 25. Ye, V.M. y S.K. Bhatia, Metabolic engineering for the production of clinically important molecules: Omega-3 fatty acids, artemisinin, and taxol. *Biotechnology Journal*, 2012. 7(1): p. 20-33.
26. Nykiforuk, C., y otros, High level accumulation of gamma linolenic acid (C18:3Δ6,9,12 cis) in transgenic safflower (*Carthamus tinctorius*) seeds. *Transgenic Research*, 2012. 21(2): p. 367-381.
27. Wu, G., y otros, Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. *Nat Biotech*, 2005. 23(8): p. 1013-1017.
- 30 28. Lee, L.Y. y S.B. Gelvin, T-DNA binary vectors and systems. *Plant Physiol*, 2008. 146(2): p. 325-32.
29. Zuo, J., y otros, Chemical-regulated, site-specific DNA excision in transgenic plants. *Nat Biotechnol*, 2001. 19(2): p. 157-61.
30. Lazo, G.R., P.A. Stein, y R.A. Ludwig, A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology (N Y)*, 1991. 9(10): p. 963-7.
- 35 31. Smith, T.L. y S.A. Leong, Isolation and characterization of a *Ustilago maydis* glyceraldehyde-3-phosphate dehydrogenase-encoding gene. *Gene*, 1990.93(1): p. 111-7.
32. Ji, L., y otros, A simplified and efficient method for transformation and gene tagging of *Ustilago maydis* using frozen cells. *Fungal Genet Biol*, 2010. 47(4): p. 279-87.
- 40 33. Punt, P.J., y otros, Functional elements in the promoter region of the *Aspergillus nidulans* *gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene*, 1990. 93(1): p. 101-9.
34. Steiner, S. y P. Philippsen, Sequence and promoter analysis of the highly expressed *TEF* gene of the filamentous fungus *Ashbya gossypii*. *Mol Gen Genet*, 1994. 242(3): p. 263-71.
35. Liu, Y.G. y R.F. Whittier, Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, 1995. 25(3): p. 674-81.
- 45 36. Liu, Y.G. y Chen, High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques*, 2007. 43(5): p. 649-50, 652, 654 passim.
37. Kimura, K., M. Yamaoka, y. Kamisaka, Rapid estimation of lipids in oleaginous fungi and yeasts using Nile red fluorescence. *J Microbiol Methods*, 2004. 56(3): p. 331-8.
- 50 38. Elsey, D., y otros, Fluorescent measurement of microalgal neutral lipids. *J Microbiol Methods*, 2007. 68(3): p. 639-42.
39. Voelker, T.A. y H.M. Davies, Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. *J Bacteriol*, 1994. 176(23): p. 7320-7.
40. Liu, Y., y otros, Tartronate semialdehyde reductase defines a novel rate-limiting step in assimilation and bioconversion of glycerol in *Ustilago maydis*. *PLoS One*, 2011. 6(1): p. e16438.
- 55 41. An, S., y otros, Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiology*, 2003. 133(4): p. 2040-2047.
42. Rosso, M.G., y otros, An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant molecular biology*, 2003. 53(1-2): p. 247-259.
43. Choi, J., y otros, Genome -wide analysis of T-DNA integration into the chromosomes of *Magnaporthe oryzae*. *Molecular microbiology*, 2007. 66(2): p. 371-382.
- 60 44. Omura, S., The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriol Rev*, 1976. 40(3): p. 681-97.
45. Vance, D., y otros, Inhibition of fatty acid synthetases by the antibiotic cerulenin. *Biochem Biophys Res Commun*, 1972. 48(3): p. 649-56.
- 65 46. Zhong, Y., y otros, Application of T-DNA insertional mutagenesis for improving cellulase production in the filamentous fungus *Trichoderma reesei*. *Bioresour Technol*, 2012. 110:p. 572-7.

47. Morita, N., y otros, Enhancement of polyunsaturated fatty acid production by cerulenin treatment in polyunsaturated fatty acid-producing bacteria. *Biotechnol Lett*, 2005. 27(6): p. 389-93.
48. Wang, J., y otros, A quick isolation method for mutants with high lipid yield in oleaginous yeast. *World Journal of Microbiology and Biotechnology*, 2009. 25(5): p. 921-925.
- 5 49. Nguyen, L.N. y J.D. Nosanchuk, The inhibitory effect of cerulenin to yeasts is fungicidal *Commun Integr Biol*, 2011. 4(6): p. 631-2.
50. Greenspan, P., E.P. Mayer, y S.D. Fowler, Nile red: a selective fluorescent stain for intracellular lipid droplets. *The Journal of cell biology*, 1985. 100(3): p. 965-973.
51. Spiekermann, P., y otros, A sensitive, viable-colony staining method using Nile red for direct screening of bacteria that accumulate polyhydroxyalkanoic acids and other lipid storage compounds. *Arch Microbiol*, 1999. 171(2): p. 73-80.
- 10 52. Zhao, Y.-F. y Q.-Z. Kong, Tetrazolium violet inhibits cell growth and induces cell death in C127 mouse breast tumor cells. *Chemico-biological interactions*, 2008. 174(1): p. 19-26.
53. Beopoulos, A., J.M. Nicaud, y C. Gaillardin, An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Appl Microbiol Biotechnol*, 2011. 90(4): p. 1193-206.
- 15 54. Li, Z., y otros, Overexpression of malic enzyme (ME) of *Mucor circinelloides* improved lipid accumulation in engineered *Rhodotorula glutinis*. *Appl Microbiol Biotechnol*, 2012.
55. Zhang, Y., I.P. Adams, y C. Ratledge, Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiology*, 2007. 153(Pt 7): p. 2013-25.
- 20 56. Rani, S.H., S. Saha, y R. Rajasekharan, A soluble diacylglycerol acyltransferase is involved in triacylglycerol biosynthesis in the oleaginous yeast *Rhodotorula glutinis*. *Microbiology*, 2013. 159(Pt 1): p. 155-66.
57. Dyer, J., y otros, Production of linolenic acid in yeast cells expressing an omega-3 desaturase from tung (*Aleurites fordii*). *Journal of the American Oil Chemists' Society*, 2004. 81(7): p. 647-651.
58. Huang, Y.S., y otros, Cloning of delta12- and delta6-desaturases from *Mortierella alpina* and recombinant production of gamma-linolenic acid in *Saccharomyces cerevisiae*. *Lipids*, 1999. 34(7): p. 649-59.
- 25 59. Ueda, M. y A. Tanaka, Long-chain aldehyde dehydrogenase of *Candida* yeast. *Methods in enzymology*, 1990. 188: p. 176-178.
60. Koh, C.M., y otros, Molecular characterization of KU70 and KU80 homologues and exploitation of a KU70-deficient mutant for improving gene deletion frequency in *Rhodospiridium toruloides*. *BMC Microbiology*, 2014. 14(1): p. 50.
- 30 61. Gong, W.F., y otros, [Effect of silencing lycB gene on the carotenoid synthesis in *Haematococcus pluvialis*]. *Yi Chuan*, 2013. 35(2): p. 233-40.
62. Chan, S.T., y otros, Quercetin supplementation suppresses the secretion of pro-inflammatory cytokines in the lungs of Mongolian gerbils and in A549 cells exposed to benzo[a]pyrene alone or in combination with beta-carotene: in vivo and ex vivo studies. *J Nutr Biochem*, 2012. 23(2): p. 179-85.
- 35 Zuo, J. y otros (2001). Chemical-regulated, site-specific DNA excision in transgenic plants. *Nat Biotechnol* 19:157-161.

Listado de secuencias

5	<110> Temasek Life Sciences Laboratory Limited	
	<120> Métodos para la producción eficiente de ácidos grasos poliinsaturados (PUFA) en especies de Rhodosporidium y Rhodotorula.	
	<130> 2577-237	
	<150> US 60/047,300	
	<151> 2014-09-08	
10	<160> 146	
	<170> PatentIn versión 3.5	
	<210> 1	
	<211> 2461	
	<212> ADN	
	<213> Rhodosporidium toruloides	
15	<400> 1	
	atgcaggaca ctcccatcga cagcatcccc caggtacatc ccacgctgac ccgcgcttct	60
	ctcgctcgct tgcgtgtcgt cctcgcttct ggagcttcga aacaggcggg acgaggggat	120
20	cggtggtgcag cgcgcgcctg gggaggattc gctggacatc ggggatattg ctggaggata	180
	ctcaggactt tctggattgg tcccgccctt cccgcgcccg tcgtcgctct ccagcatcca	240
	ttccccactc gacactccgg caaactcgcc gctgactctc cccttcgctc attccgcagg	300
25	cttacgacac cgtcaccaag gcgttcttgt cgggaaagac taggcccacg gcctggcgca	360
	aggcgcagat caaaaagctc gggttccttg tcgtgagttc gcggttcttt cgcgcggtct	420
	ttgcgacggc tcgtccagtg caggtcgcgc tggcgcggt cgtcaatgta gttgcgatga	480
30	aacctgacct agcaactccc cctccacag caagacaacg aagacgcctt cgtccgcgct	540
	cttgagcaag actttggccg cccagccttc gagacaatca cgcgcgagat caaccccgtc	600
	aaggctgaga tcaacgaggt ctacgaccac ctcgagaagt gggccaagcc gagcgcgctc	660
35	aagacttcgg cgacgtggtg cgtaccaag ccgacggtct actcggagcc taagggtgtc	720
	acgctcgta tcgggacgtg gaactgtgct tgcacttgtt ttgtacgag ttgagtgtgt	780
	gtactgacgg aagatgtcgc cgcagaccg atcacgctcc tcctcgctcc gcttctcggc	840
40	gccatctctg ccggctgcac cgcgctcgtc aagggtgcct gccgttcgaa gtccgtcgta	900
	ctgcatgtct cactgacact cgtcgctccc acagcccgt gagcaagccc ctcacgtcgc	960
	cgcgctcgtc gcgacctcc tgcccaagta cctcgacccc accgccttca tctgcatcaa	1020
45	cggcgccatc cctcaagcga ccgctctcct caaactcaag ttcgatcaca tcttctacac	1080
	cggttcggga acggtcggca agatcgttgc gcgtgcggcg gcggagcacc tttgcccggt	1140
50	tacgcttgag ttggggggaa agagtcgggc ggttgtgtgt gatgatgcgg atattgaggt	1200
	tgtggcgagg aggatcgttt gggccaagtt taccaacgct gggcaggtgc gtcgcgagaa	1260
55		
60		
65		

ES 2 724 998 T3

	ccgggttggtg tcgttggtct cgccgaagcg ggccgagatg cttagtcac cgccttggtgt	1320
	tgcgcacaga tctgcatctc gacagactac gtccctacga ccccgagac cgagcccaag	1380
5	ctcctcgaag cctcaagcg cgctctcgcc gccttctccg ccaaccccg cgcctcctcc	1440
	tcctcggaaa agtcgtcaac ctgcgtcgtg cacaaccoga actactcgcg catcatcaac	1500
10	cagaaccact acaaccgcgt ttogaagttg cttgatgcga ctaagggcga ggtggttgtt	1560
	ggcggcgggg gggacgagaa ggagcgcaag atcgaggtca cgatcgtgag gggcgtaag	1620
	ccggatgact cgctcatgtc gggcgccac tgcggctctc cccctgaaga acgaatgtgg	1680
15	ctgacgaatg cgaccgagca gaggagattt tcggccccgt ccccccgatc gtgaccctcc	1740
	cgacgctcga cgacatggtc aagttcattc agtcgcgcga cagcctctc gcgctttacg	1800
	tcttcacgca gaggaggaag aaccgcgact ttagtgcggt ccccgctct ctctcgctcg	1860
20	ctgaccttcg actgactcgg tggtagatgc agtttttgag cgcactcgtc cgggcggatt	1920
	cggttcagaac gatgtgctcg ttcagttcat gatccctggg ttgcccgttc gcggtacggg	1980
25	cgccggcgggc tacggaaact accacggcag gcggtgcgtc ccgcttcttc ggcaccgtgc	2040
	ttccccgagt ccggctgacc tgctcgcacg cagcaccttc gacacgttct cgcacgagcg	2100
	cgcgctggcc aatgtcccca cctggatgga catgatcatg gcgtcgcggt accctcccta	2160
30	cacccggttcg tccgggtcgt ccttctctct ctgtgctcgc taacacactc gtcccacacg	2220
	tgcagagaag aagctgaaga tgctcctggt cgcgaccaag gcggtgatca agaagcccag	2280
	caagtttggc tcgatctcgc gcttgctcaa ggtgattgcc gcgatggtcg ctctcttggc	2340
35	tgtcagggcc aggtctcgac tgaccgcgtc tcgtccacc cctcccttc tccaccactc	2400
	ctcttctctc cgggtcttgg gaatcgtgtg cgctgggcag aagttgacgg gacaggcgtg	2460
40	a	2461
	<210> 2	
	<211> 1515	
	<212> ADN	
	<213> Rhodosporidium toruloides	
45	<220>	
	<221> CDS	
	<222> (1)..(1515)	
	<400> 2	
50	atg gcc gcc atg cag gac act ccc atc gac agc atc ccc cag gct tac	48
	Met Ala Ala Met Gln Asp Thr Pro Ile Asp Ser Ile Pro Gln Ala Tyr	
	1 5 10 15	
	gac acc gtc acc aag gcg ttc ttg tcg gga aag act agg ccc atc gcc	96
55	Asp Thr Val Thr Lys Ala Phe Leu Ser Gly Lys Thr Arg Pro Ile Ala	
	20 25 30	
60		
65		

ES 2 724 998 T3

	tgg cgc aag gcg cag atc aaa aag ctc ggg ttc ctt gtc caa gac aac	144
	Trp Arg Lys Ala Gln Ile Lys Lys Leu Gly Phe Leu Val Gln Asp Asn	
	35 40 45	
5	gaa gac gcc ttc gtc cgc gct ctt gag caa gac ttt ggc cgc cca gcc	192
	Glu Asp Ala Phe Val Arg Ala Leu Glu Gln Asp Phe Gly Arg Pro Ala	
	50 55 60	
10	ttc gag aca atc acc gcc gag atc aac ccc gtc aag gct gag atc aac	240
	Phe Glu Thr Ile Thr Ala Glu Ile Asn Pro Val Lys Ala Glu Ile Asn	
	65 70 75 80	
15	gag gtc tac gac cac ctc gag aag tgg gcc aag ccg agg cgc gtc aag	288
	Glu Val Tyr Asp His Leu Glu Lys Trp Ala Lys Pro Arg Arg Val Lys	
	85 90 95	
	act tcg gcg acg tgg tac gct acc aag ccg acg gtc tac tcg gag cct	336
	Thr Ser Ala Thr Trp Tyr Ala Thr Lys Pro Thr Val Tyr Ser Glu Pro	
	100 105 110	
20	aag ggt gtc acg ctc gtt atc ggg acg tgg aac tac ccg atc acg ctc	384
	Lys Gly Val Thr Leu Val Ile Gly Thr Trp Asn Tyr Pro Ile Thr Leu	
	115 120 125	
25	ctc ctc gtc ccg ctt ctc ggc gcc atc tct gcc ggc tgc acc gcg ctc	432
	Leu Leu Val Pro Leu Leu Gly Ala Ile Ser Ala Gly Cys Thr Ala Leu	
	130 135 140	
	gtc aag ccc gct gag caa gcc cct cac gtc gcc gcg ctc gtc gcc gac	480
	Val Lys Pro Ala Glu Gln Ala Pro His Val Ala Ala Leu Val Ala Asp	
	145 150 155 160	
30	ctc ctg ccc aag tac ctc gac ccc acc gcc ttc atc tgc atc aac ggc	528
	Leu Leu Pro Lys Tyr Leu Asp Pro Thr Ala Phe Ile Cys Ile Asn Gly	
	165 170 175	
35	gcc atc cct caa gcg acc gct ctc ctc aaa ctc aag ttc gat cac atc	576
	Ala Ile Pro Gln Ala Thr Ala Leu Leu Lys Leu Lys Phe Asp His Ile	
	180 185 190	
	ttc tac acc ggt tcg gga acg gtc ggc aag atc gtt gcg cgt gcg gcg	624
	Phe Tyr Thr Gly Ser Gly Thr Val Gly Lys Ile Val Ala Arg Ala Ala	
	195 200 205	
40	gcg gag cac ctt tgc ccg gtt acg ctt gag ttg ggg gga aag agt ccg	672
	Ala Glu His Leu Cys Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro	
	210 215 220	
45	gcg gtt gtg ctg gat gat gcg gat att gag gtt gtg gcg agg agg atc	720
	Ala Val Val Leu Asp Asp Ala Asp Ile Glu Val Val Ala Arg Arg Ile	
	225 230 235 240	
	gtt tgg gcc aag ttt acc aac gct ggg cag atc tgc atc tcg aca gac	768
	Val Trp Ala Lys Phe Thr Asn Ala Gly Gln Ile Cys Ile Ser Thr Asp	
	245 250 255	
50	tac gtc ctc acg acc ccg cag acc gag ccc aag ctc ctc gaa gcc ctc	816
	Tyr Val Leu Thr Thr Pro Gln Thr Glu Pro Lys Leu Leu Glu Ala Leu	
	260 265 270	
55	aag cgc gct ctc gcc gcc ttc tcc gcc aac ccc gcc gcc tcc tcc tcc	864
	Lys Arg Ala Leu Ala Ala Phe Ser Ala Asn Pro Ala Ala Ser Ser Ser	
	275 280 285	

5
10
15
20
25
30
35
40
45
50
55
60
65

```

tcg gaa aag tcg tca acc tcg ctc gtg cac aac ccg aac tac tcg cgc      912
Ser Glu Lys Ser Ser Thr Ser Leu Val His Asn Pro Asn Tyr Ser Arg
 290                               295                               300

atc atc aac cag aac cac tac aac cgc gtt tcg aag ttg ctt gat gcg      960
Ile Ile Asn Gln Asn His Tyr Asn Arg Val Ser Lys Leu Leu Asp Ala
305                               310                               315                               320

act aag ggc gag gtg gtt gtt ggc ggc ggg agg gac gag aag gag cgc      1008
Thr Lys Gly Glu Val Val Val Gly Gly Gly Arg Asp Glu Lys Glu Arg
                               325                               330                               335

aag atc gag gtc acg atc gtg agg ggc gtt aag ccg gat gac tcg ctc      1056
Lys Ile Glu Val Thr Ile Val Arg Gly Val Lys Pro Asp Ser Leu
                               340                               345                               350

atg tcg gag gag att ttc ggc ccc gtc ccc ccg atc gtg acc ctc ccg      1104
Met Ser Glu Glu Ile Phe Gly Pro Val Pro Pro Ile Val Thr Leu Pro
                               355                               360                               365

acg ctc gac gac atg gtc aag ttc att cag tcg cgc gac acg cct ctc      1152
Thr Leu Asp Asp Met Val Lys Phe Ile Gln Ser Arg Asp Thr Pro Leu
                               370                               375                               380

gcg ctt tac gtc ttc acg cag agc agg aag aac cgc gac ttt att ttt      1200
Ala Leu Tyr Val Phe Thr Gln Ser Arg Lys Asn Arg Asp Phe Ile Phe
385                               390                               395                               400

gag cgc act cgc tcg ggc gga ttc gtt cag aac gat gtg ctc gtt cag      1248
Glu Arg Thr Arg Ser Gly Gly Phe Val Gln Asn Asp Val Leu Val Gln
                               405                               410                               415

ttc atg atc cct ggc ttg ccg ttc ggc ggt acg ggc gcg gcg ggc tac      1296
Phe Met Ile Pro Gly Leu Pro Phe Gly Gly Thr Gly Ala Ala Gly Tyr
                               420                               425                               430

gga aac tac cac ggc agg cgc acc ttc gac acg ttc tcg cac gag cgc      1344
Gly Asn Tyr His Gly Arg Arg Thr Phe Asp Thr Phe Ser His Glu Arg
                               435                               440                               445

gcg tcg gcc aat gtc ccc acc tgg atg gac atg atc atg gcg tcg cgg      1392
Ala Ser Ala Asn Val Pro Thr Trp Met Asp Met Ile Met Ala Ser Arg
450                               455                               460

tac cct ccc tac acc cag aag aag ctg aag atg ctc ctg ttc gcg acc      1440
Tyr Pro Pro Tyr Thr Gln Lys Lys Leu Lys Met Leu Leu Phe Ala Thr
465                               470                               475                               480

aag gcg gtg atc aag aag ccc agc aag ttt ggc tcg atc tcg cgc ttg      1488
Lys Ala Val Ile Lys Lys Pro Ser Lys Phe Gly Ser Ile Ser Arg Leu
                               485                               490                               495

ctc aag aag ttg acg gga cag gcg tga      1515
Leu Lys Lys Leu Thr Gly Gln Ala
                               500

<210> 3
<211> 504
<212> PRT
<213> Rhodosporidium toruloides
<400> 3

```

ES 2 724 998 T3

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

	245	250	255
5	Tyr Val Leu Thr Thr Pro Gln Thr Glu Pro Lys Leu Leu Glu Ala Leu 260 265 270		
10	Lys Arg Ala Leu Ala Ala Phe Ser Ala Asn Pro Ala Ala Ser Ser Ser 275 280 285		
15	Ser Glu Lys Ser Ser Thr Ser Leu Val His Asn Pro Asn Tyr Ser Arg 290 295 300		
20	Ile Ile Asn Gln Asn His Tyr Asn Arg Val Ser Lys Leu Leu Asp Ala 305 310 315 320		
25	Thr Lys Gly Glu Val Val Val Gly Gly Gly Arg Asp Glu Lys Glu Arg 325 330 335		
30	Lys Ile Glu Val Thr Ile Val Arg Gly Val Lys Pro Asp Asp Ser Leu 340 345 350		
35	Met Ser Glu Glu Ile Phe Gly Pro Val Pro Pro Ile Val Thr Leu Pro 355 360 365		
40	Thr Leu Asp Asp Met Val Lys Phe Ile Gln Ser Arg Asp Thr Pro Leu 370 375 380		
45	Ala Leu Tyr Val Phe Thr Gln Ser Arg Lys Asn Arg Asp Phe Ile Phe 385 390 395 400		
50	Glu Arg Thr Arg Ser Gly Gly Phe Val Gln Asn Asp Val Leu Val Gln 405 410 415		
55	Phe Met Ile Pro Gly Leu Pro Phe Gly Gly Thr Gly Ala Ala Gly Tyr 420 425 430		
60	Gly Asn Tyr His Gly Arg Arg Thr Phe Asp Thr Phe Ser His Glu Arg 435 440 445		
65	Ala Ser Ala Asn Val Pro Thr Trp Met Asp Met Ile Met Ala Ser Arg 450 455 460		
	Tyr Pro Pro Tyr Thr Gln Lys Lys Leu Lys Met Leu Leu Phe Ala Thr 465 470 475 480		
	Lys Ala Val Ile Lys Lys Pro Ser Lys Phe Gly Ser Ile Ser Arg Leu 485 490 495		
	Leu Lys Lys Leu Thr Gly Gln Ala 500		
	<210> 4		
	<211> 1206		
	<212> ADN		
	<213> Mortierella alpine		
	<220>		
	<221> CDS		
	<222> (1)..(1206)		
	<400> 4		

	atg gcc ccg ccg aac acc atc gac gcc ggc ctc acc cag cgc cac atc	48
10	Met Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile 1 5 10 15	
	tog acc tog gcc gcc ccg aac tog gcc aag ccg acc ttc gag cgc aac	96
	Ser Thr Ser Ala Ala Pro Asn Ser Ala Lys Pro Thr Phe Glu Arg Asn 20 25 30	
15	tac cag ctc ccg gag ttc acc atc aag gag atc cgc gag tgc atc ccg	144
	Tyr Gln Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro 35 40 45	
20	gcc cac tgc ttc gag cgc tog ggc ctc cgc ggc ctc tgc cac gtc gcc	192
	Ala His Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala 50 55 60	
25	atc gac ctc acc tgg gcc tog ctc ctc ttc ctc gcc gcc acc cag atc	240
	Ile Asp Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile 65 70 75 80	
	gac aag ttc gag aac ccg ctc atc cgc tac ctc gcc tgg ccg gtc tac	288
	Asp Lys Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Val Tyr 85 90 95	
30	tgg atc atg cag ggc atc gtc tgc acc ggc atc tgg gtc ctc gcc cac	336
	Trp Ile Met Gln Gly Ile Val Cys Thr Gly Ile Trp Val Leu Ala His 100 105 110	
35	gag tgc ggc cac cag tog ttc tog acc tog aag acc ctc aac aac acc	384
	Glu Cys Gly His Gln Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr 115 120 125	
	gtc ggc tgg atc ctc cac tog atg ctc ctc gtc ccg tac cac tog tgg	432
	Val Gly Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Trp 130 135 140	
40	cgc atc tog cac tog aag cac cac aag gcc acc ggc cac atg acc aag	480
	Arg Ile Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys 145 150 155 160	
45	gac cag gtc ttc gtc ccg aag acc cgc acc cag gtc ggc ctc ccg ccg	528
	Asp Gln Val Phe Val Pro Lys Thr Arg Thr Gln Val Gly Leu Pro Pro 165 170 175	
	aag gag tog gcc gcc gcc acc gtc cag gag gag gag gac atg tog gtc	576
50	Lys Glu Ser Ala Ala Ala Thr Val Gln Glu Glu Glu Asp Met Ser Val 180 185 190	

5	cac ctc gac gag gag gcc ccg atc gtc acc ctc ttc tgg atg gtc atc His Leu Asp Glu Glu Ala Pro Ile Val Thr Leu Phe Trp Met Val Ile 195 200 205	624
10	cag ttc ctc ttc gcc tgg ccg gcc tac ctc atc atg aac gcc tcg gcc Gln Phe Leu Phe Gly Trp Pro Ala Tyr Leu Ile Met Asn Ala Ser Gly 210 215 220	672
15	cag gac tac gcc cgc tgg acc tcg cac ttc cac acc tac tcg ccg atc Gln Asp Tyr Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile 225 230 235 240	720
20	ttc gag ccg cgc aac ttc ttc gac atc atc ctc tcg gac ctc gcc gtc Phe Glu Pro Arg Asn Phe Phe Asp Ile Ile Leu Ser Asp Leu Gly Val 245 250 255	768
25	ctc gcc acc ctc gcc gcc ctc atc tac gcc tcg atg cag ctc tcg ctc Leu Ala Thr Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu 260 265 270	816
30	ctc acc gtc acc aag tac tac atc atc ccg tac ctc ttc gtc aac ttc Leu Thr Val Thr Lys Tyr Tyr Ile Ile Pro Tyr Leu Phe Val Asn Phe 275 280 285	864
35	tgg ctc gtc ctc atc acc ttc ctc cag cac acc gac ccg aag ctc ccg Trp Leu Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro 290 295 300	912
40	cac tac cgc gag gcc gcc tgg aac ttc cag cgc gcc gcc ctc tgc acc His Tyr Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr 305 310 315 320	960
45	gtc gac cgc tcg ttc gcc aag ttc ctc gac cac atg ttc cac gcc atc Val Asp Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile 325 330 335	1008
50	gtc cac acc cac gtc gcc cac cac ctc ttc tcg cag atg ccg ttc tac Val His Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr 340 345 350	1056
55	cac gcc gag gag gcc acc tac cac ctc aag aag ctc ctc gcc gag tac His Ala Glu Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr 355 360 365	1104
60	tac atc tac gac gcc tcg ccg atc gtc gtc gcc gtc tgg aag tcg ttc Tyr Ile Tyr Asp Ala Ser Pro Ile Val Val Ala Val Trp Lys Ser Phe 370 375 380	1152
65	cgc gag tgc cgc ttc gtc gag gac cac gcc gac gtc gtc ttc ttc aag Arg Glu Cys Arg Phe Val Glu Asp His Gly Asp Val Val Phe Phe Lys 385 390 395 400	1200
	aag tag Lys <210> 5 <211> 401 <212> PRT <213> Mortierella alpine <400> 5	1206

ES 2 724 998 T3

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

5 Leu Ala Thr Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu
 260 265 270
 Leu Thr Val Thr Lys Tyr Tyr Ile Ile Pro Tyr Leu Phe Val Asn Phe
 275 280 285
 10 Trp Leu Val Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro
 290 295 300
 15 His Tyr Arg Glu Gly Ala Trp Asn Phe Gln Arg Gly Ala Leu Cys Thr
 305 310 315 320
 Val Asp Arg Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile
 325 330 335
 20 Val His Thr His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr
 340 345 350
 25 His Ala Glu Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr
 355 360 365
 Tyr Ile Tyr Asp Ala Ser Pro Ile Val Val Ala Val Trp Lys Ser Phe
 370 375 380
 30 Arg Glu Cys Arg Phe Val Glu Asp His Gly Asp Val Val Phe Phe Lys
 385 390 395 400
 35 Lys
 <210> 6
 <211> 2760
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 6
 40 caccgcttcc cggagaactt tgctgtactc tgcttctccc ttcacactct cacacccact 60
 cacacaccct tccatccaca cacaagctat ccgcacacct ctcacaccog accccagctc 120
 45 gccccatcct cttcgcaccc ggctcatcgc cacacacgca atgactgcct cgtcggcact 180
 cgagacctcg ctcccgcaact ctgtcggggc cgagtctgcg actaccaccg caaagccgccc 240
 ccgtgcgccc ctcaggatgc gtcaccccga ctacaccag accgacgtcc tcgagtcgtc 300
 50 agactcggac gcagcgtcgg attccgaggg cgagacgacg gctgtcgacg acgggaccta 360
 cgaggacgat aactacgtcc gcaaggtcct cagcaaggag aagccgctcc cgcccatcac 420
 ctggaagaac atccaccgca acatccagtg gatctcgacc ctgcacctca ccatcgtgcc 480
 55
 60
 65

ES 2 724 998 T3

	cctcctctcg atctacggag cgttcacgac gcccctgaag tggcagacgg cggctctggag	540
5	tgtcgtctac tactacttca ccggtctcgg gtgagtcgcc gcacccttcc ctgcctctc	600
	ctccctcgct ctctacttcc ggccgctccg ttctctgttc atgacgtct gtaagacgtg	660
	tttgagcttg gagggcgctc agatggcccg gcggacgagt cgcctgtcga gaaaccgtcg	720
10	ctgaccttct ctctgccttt tcctcctcct ccatttgtcc cgatcgtccc gatcgtctg	780
	ttctcgttct acagcatcac cgcaggctac cacaggctgt acgtgtcatt tactctgac	840
	tcgaggcgtt gagcgccgaa aaacggactc tgaactgacg aaaactcgct tttctcgct	900
15	tcctcgcctg taggtgggac cacaggctct acaactgcct cctgcctctc cagtacttcc	960
	tggcacttgg cggaagcggc gcagtgagg gttctgtgaa atggtgggac cgaggacacc	1020
	ggcacacca ccgctacacc gacacggacc tcgacctgta ctacgcgacg aagggttct	1080
20	ggtgggcaca cctcggttg atgattgtca agccgcgcg tcgtcccggt gtcgccgatg	1140
	tctccgacct caacaacaac ccagtcgtca agtggcagca ccgcttctac ctcccgctca	1200
	tcctcggcat gggcttctac ttccctacca tcgtcgttg actcggttg ggcgacttcc	1260
25	ggggcggtt tttcttcgcc ggcgctgtc gcctcctctt tgtccaccac gtgcgttctc	1320
	gcctcctcgt tctctcctcc ttcttctgt gacgttttct tattgcacag tcgacgttct	1380
	gcgtcaactc gctcgcacac tggctcggcg agacgccgtt tgacgacaag cacacgccga	1440
30	aggaccactg gctcaccccg ctgcgcacgg tcggcgaggg ctaccacaac ttctgtgagtc	1500
	tcocgccgcg tctctcgttg atctacggcg tcttctgtga ctgctgccac agtcgcgacg	1560
35	tcgttcgagt cgttttgccat cgaactcgt ctctcgcacc tcccgctgac ctctccctt	1620
	cccctccagc accacgagtt cccctccgac taccgcaacg cgtcagatg gtggcagtat	1680
	gatccgacta agtggttctat ttacgcgatg tcgaaactcg gattggcgtc gcagctcaag	1740
40	acgttcccg acaacgagat caagaagggt cagtaacca tgacgtcaa ggctgtcgg	1800
	cgcgaggcgg agaacatcga gtggcccaag tcgtcgaacc acttgctgt gctcacctgg	1860
	gatgagtga cgttcagctt tcctcgtcgt cgcggcgac ctgcaccttc ctgaacttcc	1920
45	agcgtgact ttctcccg ctctctcgc tcgcagtcca ggaggcctgc aagactcgc	1980
	agctcctcgt tgtcgcgggt ttcatccacg atgtcagcac ctcatcgac cagcacctg	2040
	gcggtgcgg cttgatcaag acccgtctcg gcccgatgc gacgaccgc ttctacggtg	2100
50	gctactacga gtgcgtcttc acattctcct gcctcatctt cgtcaagtc gctgacgtc	2160
	tcgtcccg agccactcga acggcgacg caacttgctc gccagtagc gtgtcggcgt	2220
	catcgagggc ggctacgagg tcgagcacat gaagaagtag tctgaggtcg tcgagaacct	2280
55	caagaagcac ggcccgacg gcgtggccgg caagagcgcc gacctcgca agggtcgaa	2340

60

65

	gcagatgtcg gtcacaaagg ggcaccctca gctcaagggc ggcgcgctcg agacgctcgc	2400
5	caagccgcct accttcagcg agaccaacct tttgggcggt ctcagcctga cggtaagggc	2460
	gtaatggccc gtcaccgcca tcggttcaag ggagagtcgt cgaccgactc ttccgttcgt	2520
	cgcttcaggt tcacccgtct tcgcatgctt agagaccctc ttccctcgtt ttttgtgcat	2580
10	cagttgcccg cagtcgtccc ttaacgagcc cactactacc ctttccatgc ccagtattct	2640
	tgcccccgcc cccttcgttc tttgatacgt ccaaccacct cgctctttcg ctgtagctct	2700
15	acttcgttct ctcccttgcg ggttgccgcg ttactccgca tggaaacttg accagtgtgc	2760
	<210> 7	
	<211> 1638	
	<212> ADN	
	<213> Rhodosporidium toruloides	
	<220>	
20	<221> CDS	
	<222> (1)..(1638)	
	<400> 7	
25	atg act gcc tcg tcg gca ctc gag acc tcg ctc ccg cac tct gtc ggg	48
	Met Thr Ala Ser Ser Ala Leu Glu Thr Ser Leu Pro His Ser Val Gly	
	1 5 10 15	
	ccc gag tct gcg act acc acc gca aag ccg ccc cgt gcg ccg ctc agg	96
	Pro Glu Ser Ala Thr Thr Thr Ala Lys Pro Pro Arg Ala Pro Leu Arg	
	20 25 30	
30	atg cgt cac ccc gac tac acc cag acc gac gtc ctc gag tcg tca gac	144
	Met Arg His Pro Asp Tyr Thr Gln Thr Asp Val Leu Glu Ser Ser Asp	
	35 40 45	
35	tcg gac gca gcg tcg gat tcc gag gcc gag acg acg gct gtc gac gac	192
	Ser Asp Ala Ala Ser Asp Ser Glu Gly Glu Thr Thr Ala Val Asp Asp	
	50 55 60	
	ggg acc tac gag gac gac aac tac gtc cgc aag gtc ctc agc aag gag	240
	Gly Thr Tyr Glu Asp Asp Asn Tyr Val Arg Lys Val Leu Ser Lys Glu	
	65 70 75 80	
40	aag ccg ctc ccg ccc atc acc tgg aag aac atc cac cgc aac atc cag	288
	Lys Pro Leu Pro Pro Ile Thr Trp Lys Asn Ile His Arg Asn Ile Gln	
	85 90 95	
45	tgg atc tcg acc ctc gcc ctc acc atc gtg ccc ctc ctc tcg atc tac	336
	Trp Ile Ser Thr Leu Ala Leu Thr Ile Val Pro Leu Leu Ser Ile Tyr	
	100 105 110	
	gga gcg ttc acg acg ccc ctg aag tgg cag acg gcg gtc tgg agt gtc	384
	Gly Ala Phe Thr Thr Pro Leu Lys Trp Gln Thr Ala Val Trp Ser Val	
	115 120 125	
50	gtc tac tac tac ttc acc ggt ctc gcc atc acc gca gcc tac cac agg	432
	Val Tyr Tyr Tyr Phe Thr Gly Leu Gly Ile Thr Ala Gly Tyr His Arg	
	130 135 140	
55	ctg tgg gcc cac agg tcc tac act gcc tcc ctg cct ctc cag tac ttc	480

	Leu Trp Ala His Arg Ser Tyr Thr Ala Ser Leu Pro Leu Gln Tyr Phe	
	145 150 155 160	
5	ctg gca ctt ggc gga agc ggc gca gtc gag ggt tct gtg aaa tgg tgg	528
	Leu Ala Leu Gly Gly Ser Gly Ala Val Glu Gly Ser Val Lys Trp Trp	
	165 170 175	
10	gcc cga gga cac cgc gca cac cac cgc tac acc gac acg gac ctc gac	576
	Ala Arg Gly His Arg Ala His His Arg Tyr Thr Asp Thr Asp Leu Asp	
	180 185 190	
15	cgc tac tca gcg cag aag ggc ttc tgg tgg gca cac ctc ggc tgg atg	624
	Pro Tyr Ser Ala Gln Lys Gly Phe Trp Trp Ala His Leu Gly Trp Met	
	195 200 205	
20	att gtc aag ccg cgc cgt cgt ccc ggt gtc gcc gat gtc tcc gac ctc	672
	Ile Val Lys Pro Arg Arg Arg Pro Gly Val Ala Asp Val Ser Asp Leu	
	210 215 220	
25	aac aac aac cca gtc gtc aag tgg cag cac cgc ttc tac ctc ccg ctc	720
	Asn Asn Asn Pro Val Val Lys Trp Gln His Arg Phe Tyr Leu Pro Leu	
	225 230 235 240	
30	atc ctc ggc atg ggc ttc atc ttc cct acc atc gtc gct gga ctc ggc	768
	Ile Leu Gly Met Gly Phe Ile Phe Pro Thr Ile Val Ala Gly Leu Gly	
	245 250 255	
35	tgg ggc gac ttc cgc ggc gga ttt ttc ttc gcc ggc gct gct cgc ctc	816
	Trp Gly Asp Phe Arg Gly Gly Phe Phe Phe Ala Gly Ala Ala Arg Leu	
	260 265 270	
40	ctc ttt gtc cac cac tcg acg ttc tgc gtc aac tcg ctc gca cac tgg	864
	Leu Phe Val His His Ser Thr Phe Cys Val Asn Ser Leu Ala His Trp	
	275 280 285	
45	ctc ggc gag acg ccg ttt gac gac aag cac acg ccg aag gac cac tgg	912
	Leu Gly Glu Thr Pro Phe Asp Asp Lys His Thr Pro Lys Asp His Trp	
	290 295 300	
50	ctc acc gcg ctc gcg acg gtc ggc gag ggc tac cac aac ttc cac cac	960
	Leu Thr Ala Leu Ala Thr Val Gly Glu Gly Tyr His Asn Phe His His	
	305 310 315 320	
55	gag ttc ccc tcc gac tac cgc aac gcg ctc aga tgg tgg cag tat gat	1008
	Glu Phe Pro Ser Asp Tyr Arg Asn Ala Leu Arg Trp Trp Gln Tyr Asp	
	325 330 335	
60	ccg act aag tgt ttc att tac gcg atg tcg aaa ctc gga ttg gcg tcg	1056
	Pro Thr Lys Cys Phe Ile Tyr Ala Met Ser Lys Leu Gly Leu Ala Ser	
	340 345 350	
65	cag ctc aag acg ttc ccc gac aac gag atc aag aag ggt cag tac gcc	1104
	Gln Leu Lys Thr Phe Pro Asp Asn Glu Ile Lys Lys Gly Gln Tyr Ala	
	355 360 365	
70	atg acg ctc aag gct gtc gcg cgc gag gcg gag aac atc gag tgg ccc	1152
	Met Thr Leu Lys Ala Val Ala Arg Glu Ala Glu Asn Ile Glu Trp Pro	
	370 375 380	
75	aag tcg tcg aac cac ttg cct gtg ctc acc tgg gat gag ttc cag gag	1200
	Lys Ser Ser Asn His Leu Pro Val Leu Thr Trp Asp Glu Phe Gln Glu	
	385 390 395 400	

ES 2 724 998 T3

5	gcc tgc aag act cgc cag ctc ctc gtt gtc gcc ggt ttc atc cac gat Ala Cys Lys Thr Arg Gln Leu Leu Val Val Ala Gly Phe Ile His Asp 405 410 415	1248
10	gtc agc acc ttc atc gac cag cac cct ggc ggt gcc ggc ttg atc aag Val Ser Thr Phe Ile Asp Gln His Pro Gly Gly Ala Gly Leu Ile Lys 420 425 430	1296
15	acc cgt ctc ggc cgc gat gcg acg acc gcc ttc tac ggt ggc tac tac Thr Arg Leu Gly Arg Asp Ala Thr Thr Ala Phe Tyr Gly Gly Tyr Tyr 435 440 445	1344
20	gac cac tcg aac ggc gca gcc aac ttg ctc gcc cag tac cgt gtc ggc Asp His Ser Asn Gly Ala Ala Asn Leu Leu Ala Gln Tyr Arg Val Gly 450 455 460	1392
25	gtc atc gag ggc ggc tac gag gtc gag cac atg aag aag tac tct gag Val Ile Glu Gly Gly Tyr Glu Val Glu His Met Lys Lys Tyr Ser Glu 465 470 475 480	1440
30	gtc gtc gag aac ctc aag aag cac ggc gcc gac ggc gtg gcc ggc aag Val Val Glu Asn Leu Lys Lys His Gly Ala Asp Gly Val Ala Gly Lys 485 490 495	1488
35	agc gcc gac ctc gcc aag ggt ccg aag cag atg tcg gtc atc aag ggc Ser Ala Asp Leu Ala Lys Gly Pro Lys Gln Met Ser Val Ile Lys Gly 500 505 510	1536
40	gac cct cag ctc aag ggc gcg ccg ctc gag acg ctc gcc aag ccg cct Asp Pro Gln Leu Lys Gly Ala Pro Leu Glu Thr Leu Ala Lys Pro Pro 515 520 525	1584
45	acc ttc agc gag acc aac ctt ttg ggc ggt ctc agc ctg acg gtc aag Thr Phe Ser Glu Thr Asn Leu Leu Gly Gly Leu Ser Leu Thr Val Lys 530 535 540	1632
50	gcg taa Ala 545 <210> 8 <211> 545 <212> PRT <213> Rhodosporidium toruloides <400> 8 Met Thr Ala Ser Ser Ala Leu Glu Thr Ser Leu Pro His Ser Val Gly 1 5 10 15	1638
55	Pro Glu Ser Ala Thr Thr Thr Ala Lys Pro Pro Arg Ala Pro Leu Arg 20 25 30	
60	Met Arg His Pro Asp Tyr Thr Gln Thr Asp Val Leu Glu Ser Ser Asp 35 40 45	
65	Ser Asp Ala Ala Ser Asp Ser Glu Gly Glu Thr Thr Ala Val Asp Asp 50 55 60	

ES 2 724 998 T3

5 Gly Thr Tyr Glu Asp Asp Asn Tyr Val Arg Lys Val Leu Ser Lys Glu
65 70 75 80

Lys Pro Leu Pro Pro Ile Thr Trp Lys Asn Ile His Arg Asn Ile Gln
85 90 95

10 Trp Ile Ser Thr Leu Ala Leu Thr Ile Val Pro Leu Leu Ser Ile Tyr
100 105 110

15 Gly Ala Phe Thr Thr Pro Leu Lys Trp Gln Thr Ala Val Trp Ser Val
115 120 125

Val Tyr Tyr Tyr Phe Thr Gly Leu Gly Ile Thr Ala Gly Tyr His Arg
130 135 140

20 Leu Trp Ala His Arg Ser Tyr Thr Ala Ser Leu Pro Leu Gln Tyr Phe
145 150 155 160

25 Leu Ala Leu Gly Gly Ser Gly Ala Val Glu Gly Ser Val Lys Trp Trp
165 170 175

Ala Arg Gly His Arg Ala His His Arg Tyr Thr Asp Thr Asp Leu Asp
180 185 190

30 Pro Tyr Ser Ala Gln Lys Gly Phe Trp Trp Ala His Leu Gly Trp Met
195 200 205

35 Ile Val Lys Pro Arg Arg Arg Pro Gly Val Ala Asp Val Ser Asp Leu
210 215 220

Asn Asn Asn Pro Val Val Lys Trp Gln His Arg Phe Tyr Leu Pro Leu
225 230 235 240

40 Ile Leu Gly Met Gly Phe Ile Phe Pro Thr Ile Val Ala Gly Leu Gly
245 250 255

45 Trp Gly Asp Phe Arg Gly Gly Phe Phe Phe Ala Gly Ala Ala Arg Leu
260 265 270

Leu Phe Val His His Ser Thr Phe Cys Val Asn Ser Leu Ala His Trp
275 280 285

50 Leu Gly Glu Thr Pro Phe Asp Asp Lys His Thr Pro Lys Asp His Trp
290 295 300

55 Leu Thr Ala Leu Ala Thr Val Gly Glu Gly Tyr His Asn Phe His His
305 310 315 320

Glu Phe Pro Ser Asp Tyr Arg Asn Ala Leu Arg Trp Trp Gln Tyr Asp
325 330 335

Pro Thr Lys Cys Phe Ile Tyr Ala Met Ser Lys Leu Gly Leu Ala Ser
340 345 350

Gln Leu Lys Thr Phe Pro Asp Asn Glu Ile Lys Lys Gly Gln Tyr Ala
355 360 365

Met Thr Leu Lys Ala Val Ala Arg Glu Ala Glu Asn Ile Glu Trp Pro
370 375 380

Lys Ser Ser Asn His Leu Pro Val Leu Thr Trp Asp Glu Phe Gln Glu
385 390 395 400

Ala Cys Lys Thr Arg Gln Leu Leu Val Val Ala Gly Phe Ile His Asp
405 410 415

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

Thr Arg Leu Gly Arg Asp Ala Thr Thr Ala Phe Tyr Gly Gly Tyr Tyr
435 440 445

Asp His Ser Asn Gly Ala Ala Asn Leu Leu Ala Gln Tyr Arg Val Gly
450 455 460

Val Ile Glu Gly Gly Tyr Glu Val Glu His Met Lys Lys Tyr Ser Glu
465 470 475 480

Val Val Glu Asn Leu Lys Lys His Gly Ala Asp Gly Val Ala Gly Lys
485 490 495

Ser Ala Asp Leu Ala Lys Gly Pro Lys Gln Met Ser Val Ile Lys Gly
500 505 510

Asp Pro Gln Leu Lys Gly Ala Pro Leu Glu Thr Leu Ala Lys Pro Pro
515 520 525

Thr Phe Ser Glu Thr Asn Leu Leu Gly Gly Leu Ser Leu Thr Val Lys
530 535 540

Ala
545
<210> 9
<211> 1164
<212> ADN
<213> Linum usitatissimum
<220>
<221> CDS
<222> (1)..(1164)
<400> 9

ES 2 724 998 T3

5	atg aag cag cag cag tac aag gac acc ccg atc ctc aac ggc gtc aac Met Lys Gln Gln Gln Tyr Lys Asp Thr Pro Ile Leu Asn Gly Val Asn 1 5 10 15	48
10	ggc ttc cac gcc aag gag gag gag gag gag gac ttc gac ctc tcg Gly Phe His Ala Lys Glu Glu Glu Glu Glu Glu Asp Phe Asp Leu Ser 20 25 30	96
15	aac ccg ccg ccg ttc aac atc ggc cag atc cgc gcc gcc atc ccg aag Asn Pro Pro Pro Phe Asn Ile Gly Gln Ile Arg Ala Ala Ile Pro Lys 35 40 45	144
20	cac tgc tgg gtc aag aac ccg tgg cgc tcg ctc acc tac gtc ttc cgc His Cys Trp Val Lys Asn Pro Trp Arg Ser Leu Thr Tyr Val Phe Arg 50 55 60	192
25	gac gtc gtc gtc gtc ttc gcc ctc gcc gcc gcc gcc ttc tac ttc aac Asp Val Val Val Val Phe Ala Leu Ala Ala Ala Phe Tyr Phe Asn 65 70 75 80	240
30	tcg tgg ctc ttc tgg ccg ctc tac tgg ttc gcc cag ggc acc atg ttc Ser Trp Leu Phe Trp Pro Leu Tyr Trp Phe Ala Gln Gly Thr Met Phe 85 90 95	288
35	tgg gcc atc ttc gtc ctc gcc cac gac tgc gcc cac ggc tcg ttc tcg Trp Ala Ile Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser 100 105 110	336
40	aac aac tcg tcg ctc aac aac gtc gtc gcc cac ctc ctc cac tcg tcg Asn Asn Ser Ser Leu Asn Asn Val Val Gly His Leu Leu His Ser Ser 115 120 125	384
45	atc ctc gtc ccg tac cac gcc tgg cgc atc tcg cac cgc acc cac cac Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His 130 135 140	432
50	cag aac cac gcc aac gtc gag aag gac gag tcg tgg gtc ccg ctc ccg Gln Asn His Gly Asn Val Glu Lys Asp Glu Ser Trp Val Pro Leu Pro 145 150 155 160	480
55	gag aag atc tac aag gag atg gac ctc tcg acc cgc atc ctc cgc tac Glu Lys Ile Tyr Lys Glu Met Asp Leu Ser Thr Arg Ile Leu Arg Tyr 165 170 175	528
60	tcg gtc ccg ctc ccg atg ttc gcc ctc ccg ttc tac ctc tgg tgg cgc Ser Val Pro Leu Pro Met Phe Ala Leu Pro Phe Tyr Leu Trp Trp Arg 180 185 190	576
65	tcg ccg gcc aag gag gcc tcg cac ttc aac ccg aac tcg gac ttc ttc Ser Pro Gly Lys Glu Gly Ser His Phe Asn Pro Asn Ser Asp Phe Phe 195 200 205	624
70	gcc ccg cac gag cgc aag gcc gtc ctc acc tcg aac ttc tgc ttc tcg	672

5 Ala Pro His Glu Arg Lys Ala Val Leu Thr Ser Asn Phe Cys Phe Ser
210 215 220

atc atg gcc ctc ctc ctc ctc tac tgc tgc ttc gtc ttc gcc ccg gtc 720
Ile Met Ala Leu Leu Leu Leu Tyr Ser Cys Phe Val Phe Gly Pro Val
225 230 235 240

10 cag gtc ctc aag ttc tac gcc atc ccg tac ctc gtc ttc gtc atg tgg 768
Gln Val Leu Lys Phe Tyr Gly Ile Pro Tyr Leu Val Phe Val Met Trp
245 250 255

15 ctc gac ttc gtc acc tac atg cac cac cac gcc cac gag gag aag ctc 816
Leu Asp Phe Val Thr Tyr Met His His His Gly His Glu Glu Lys Leu
260 265 270

cgc tgg tac cgc gcc aag gag tgg tgc tac ctc cgc gcc gcc ctc acc 864
Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr
275 280 285

20 acc gtc gac cgc gac tac gcc tgg atc aac aac atc cac cac gac atc 912
Thr Val Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp Ile
290 295 300

25 gcc acc cac gtc atc cac cac ctc ttc ccg cag atc ccg cac tac cac 960
Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His
305 310 315 320

30 ctc atc gag gcc acc aag gcc gcc aag ccg gtc ctc gcc aag tac tac 1008
Leu Ile Glu Ala Thr Lys Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr
325 330 335

cgc gag ccg aag aag tgc gcc ccg ttc ccg ttc cac ctc ttc tgc aac 1056
Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Ser Asn
340 345 350

35 ctc gtc cgc tgc atg tgc gag gac cac tac gtc tgc gac atc gcc gac 1104
Leu Val Arg Ser Met Ser Glu Asp His Tyr Val Ser Asp Ile Gly Asp
355 360 365

40 atc gtc ttc tac cag acc gac ccg gac atc tac aag gtc gac aag tgc 1152
Ile Val Phe Tyr Gln Thr Asp Pro Asp Ile Tyr Lys Val Asp Lys Ser
370 375 380

aag ctc aac tag 1164
Lys Leu Asn
385

45 <210> 10
<211> 387
<212> PRT
<213> Linum usitatissimum
<400> 10

50 Met Lys Gln Gln Gln Tyr Lys Asp Thr Pro Ile Leu Asn Gly Val Asn
1 5 10 15

Gly Phe His Ala Lys Glu Glu Glu Glu Glu Asp Phe Asp Leu Ser
20 25 30

55

60

65

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

Thr Val Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp Ile
 290 295 300
 Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His
 305 310 315 320
 Leu Ile Glu Ala Thr Lys Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr
 325 330 335
 Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Ser Asn
 340 345 350
 Leu Val Arg Ser Met Ser Glu Asp His Tyr Val Ser Asp Ile Gly Asp
 355 360 365
 Ile Val Phe Tyr Gln Thr Asp Pro Asp Ile Tyr Lys Val Asp Lys Ser
 370 375 380
 Lys Leu Asn
 385
 <210> 11
 <211> 1164
 <212> ADN
 <213> Vernicia fordii
 <220>
 <221> CDS
 <222> (1)..(1164)
 <400> 11
 atg aag cag cag cag tac aag gac acc ccg atc ctc aac ggc gtc aac 48
 Met Lys Gln Gln Gln Tyr Lys Asp Thr Pro Ile Leu Asn Gly Val Asn
 1 5 10 15
 ggc ttc cac gcc aag gag gag gag gag gag gac ttc gac ctc tcg 96
 Gly Phe His Ala Lys Glu Glu Glu Glu Glu Glu Asp Phe Asp Leu Ser
 20 25 30
 aac ccg ccg ccg ttc aac atc ggc cag atc cgc gcc gcc atc ccg aag 144
 Asn Pro Pro Pro Phe Asn Ile Gly Gln Ile Arg Ala Ala Ile Pro Lys
 35 40 45
 cac tgc tgg gtc aag aac ccg tgg cgc tcg ctc acc tac gtc ttc cgc 192
 His Cys Trp Val Lys Asn Pro Trp Arg Ser Leu Thr Tyr Val Phe Arg
 50 55 60
 gac gtc gtc gtc gtc ttc gcc ctc gcc gcc gcc gcc ttc tac ttc aac 240
 Asp Val Val Val Val Phe Ala Leu Ala Ala Ala Ala Phe Tyr Phe Asn
 65 70 75 80
 tcg tgg ctc ttc tgg ccg ctc tac tgg ttc gcc cag ggc acc atg ttc 288
 Ser Trp Leu Phe Trp Pro Leu Tyr Trp Phe Ala Gln Gly Thr Met Phe
 85 90 95

5	tgg gcc atc ttc gtc ctc ggc cac gac tgc ggc cac ggc tgc ttc tgc Trp Ala Ile Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser 100 105 110	336
10	aac aac tgc tgc ctc aac aac gtc gtc ggc cac ctc ctc cac tgc tgc Asn Asn Ser Ser Leu Asn Asn Val Val Gly His Leu Leu His Ser Ser 115 120 125	384
15	atc ctc gtc ccg tac cac ggc tgg cgc atc tgc cac cgc acc cac cac Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His 130 135 140	432
20	cag aac cac ggc aac gtc gag aag gac gag tgc tgg gtc ccg ctc ccg Gln Asn His Gly Asn Val Glu Lys Asp Glu Ser Trp Val Pro Leu Pro 145 150 155 160	480
25	gag aag atc tac aag gag atg gac ctc tgc acc cgc atc ctc cgc tac Glu Lys Ile Tyr Lys Glu Met Asp Leu Ser Thr Arg Ile Leu Arg Tyr 165 170 175	528
30	tgc gtc ccg ctc ccg atg ttc gcc ctc ccg ttc tac ctc tgg tgg cgc Ser Val Pro Leu Pro Met Phe Ala Leu Pro Phe Tyr Leu Trp Trp Arg 180 185 190	576
35	tgc ccg ggc aag gag ggc tgc cac ttc aac ccg aac tgc gac ttc ttc Ser Pro Gly Lys Glu Gly Ser His Phe Asn Pro Asn Ser Asp Phe Phe 195 200 205	624
40	gcc ccg cac gag cgc aag gcc gtc ctc acc tgc aac ttc tgc ttc tgc Ala Pro His Glu Arg Lys Ala Val Leu Thr Ser Asn Phe Cys Phe Ser 210 215 220	672
45	atc atg gcc ctc ctc ctc ctc tac tgc tgc ttc gtc ttc gcc ccg gtc Ile Met Ala Leu Leu Leu Leu Tyr Ser Cys Phe Val Phe Gly Pro Val 225 230 235 240	720
50	cag gtc ctc aag ttc tac ggc atc ccg tac ctc gtc ttc gtc atg tgg Gln Val Leu Lys Phe Tyr Gly Ile Pro Tyr Leu Val Phe Val Met Trp 245 250 255	768
55	ctc gac ttc gtc acc tac atg cac cac cac gcc cac gag gag aag ctc Leu Asp Phe Val Thr Tyr Met His His His Gly His Glu Glu Lys Leu 260 265 270	816
60	ccg tgg tac cgc gcc aag gag tgg tgc tac ctc cgc gcc gcc ctc acc Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr 275 280 285	864
65	acc gtc gac cgc gac tac gcc tgg atc aac aac atc cac cac gac atc Thr Val Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp Ile 290 295 300	912
70	ggc acc cac gtc atc cac cac ctc ttc ccg cag atc ccg cac tac cac Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His 305 310 315 320	960
75	ctc atc gag gcc acc aag gcc gcc aag ccg gtc ctc gcc aag tac tac Leu Ile Glu Ala Thr Lys Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr 325 330 335	1008
80	cgc gag ccg aag aag tgc gcc ccg ttc ccg ttc cac ctc ttc tgc aac Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Ser Asn 340 345 350 355 360 365 370 375 380 385 390 395 400	1056

ES 2 724 998 T3

	340	345	350	
5	ctc gtc cgc tcg atg tcg gag gac cac tac gtc tcg gac atc ggc gac			1104
	Leu Val Arg Ser Met Ser Glu Asp His Tyr Val Ser Asp Ile Gly Asp			
	355	360	365	
10	atc gtc ttc tac cag acc gac ccg gac atc tac aag gtc gac aag tcg			1152
	Ile Val Phe Tyr Gln Thr Asp Pro Asp Ile Tyr Lys Val Asp Lys Ser			
	370	375	380	
15	aag ctc aac tag			1164
	Lys Leu Asn			
	385			
	<210> 12			
	<211> 387			
	<212> PRT			
	<213> Vermicia fordii			
	<400> 12			
20	Met Lys Gln Gln Gln Tyr Lys Asp Thr Pro Ile Leu Asn Gly Val Asn			
	1	5	10	15
25	Gly Phe His Ala Lys Glu Glu Glu Glu Glu Glu Asp Phe Asp Leu Ser			
	20	25	30	
30	Asn Pro Pro Pro Phe Asn Ile Gly Gln Ile Arg Ala Ala Ile Pro Lys			
	35	40	45	
35	His Cys Trp Val Lys Asn Pro Trp Arg Ser Leu Thr Tyr Val Phe Arg			
	50	55	60	
40	Asp Val Val Val Val Phe Ala Leu Ala Ala Ala Ala Phe Tyr Phe Asn			
	65	70	75	80
45	Ser Trp Leu Phe Trp Pro Leu Tyr Trp Phe Ala Gln Gly Thr Met Phe			
	85	90	95	
50	Trp Ala Ile Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser			
	100	105	110	
55	Asn Asn Ser Ser Leu Asn Asn Val Val Gly His Leu Leu His Ser Ser			
	115	120	125	
60	Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His			
	130	135	140	
65	Gln Asn His Gly Asn Val Glu Lys Asp Glu Ser Trp Val Pro Leu Pro			
	145	150	155	160
70	Glu Lys Ile Tyr Lys Glu Met Asp Leu Ser Thr Arg Ile Leu Arg Tyr			

ES 2 724 998 T3

	165	170	175
5	Ser Val Pro Leu Pro Met Phe Ala Leu Pro Phe Tyr Leu Trp Trp Arg 180 185 190		
10	Ser Pro Gly Lys Glu Gly Ser His Phe Asn Pro Asn Ser Asp Phe Phe 195 200 205		
15	Ala Pro His Glu Arg Lys Ala Val Leu Thr Ser Asn Phe Cys Phe Ser 210 215 220		
20	Ile Met Ala Leu Leu Leu Leu Tyr Ser Cys Phe Val Phe Gly Pro Val 225 230 235 240		
25	Gln Val Leu Lys Phe Tyr Gly Ile Pro Tyr Leu Val Phe Val Met Trp 245 250 255		
30	Leu Asp Phe Val Thr Tyr Met His His His Gly His Glu Glu Lys Leu 260 265 270		
35	Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr 275 280 285		
40	Thr Val Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp Ile 290 295 300		
45	Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His 305 310 315 320		
50	Leu Ile Glu Ala Thr Lys Ala Ala Lys Pro Val Leu Gly Lys Tyr Tyr 325 330 335		
55	Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Ser Asn 340 345 350		
60	Leu Val Arg Ser Met Ser Glu Asp His Tyr Val Ser Asp Ile Gly Asp 355 360 365		
65	Ile Val Phe Tyr Gln Thr Asp Pro Asp Ile Tyr Lys Val Asp Lys Ser 370 375 380		
70	Lys Leu Asn 385		
75	<210> 13		
80	<211> 30		
85	<212> ADN		
90	<213> Secuencia Artificial		
95	<220>		
100	<223> Oligonucleótido		
105	<400> 13		
110	ttccgcggt cgaattccc cgaatcggtca 30		
115	<210> 14		
120	<211> 21		
125	<212> ADN		
130	<213> Secuencia Artificial		
135	<220>		
140	<223> oligonucleótido		

5 <400> 14
 gagtcgctca cctactgcat c 21
 <210> 15
 <211> 31
 <212> ADN
 <213> Secuencia Artificial
 <220>
 10 <223> oligonucleótido
 <400> 15
 ttgtttaa catgctaatt cggggatct g 31
 <210> 16
 <211> 19
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 16
 cacgtcgact gaaacgcag 19
 20 <210> 17
 <211> 18
 <212> ADN
 <213> Secuencia Artificial
 <220>
 25 <223> oligonucleótido
 <400> 17
 gacgaggtca tccgcgag 8
 <210> 18
 <211> 18
 <212> ADN
 30 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 18
 gtgcgggtcg tgatggac 18
 35 <210> 19
 <211> 19
 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> oligonucleótido
 <400> 19
 ctggaaggcg tacgaggac 19
 <210> 20
 <211> 20
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 20
 gtcaagccgc ccaggctgtc 20
 50 <210> 21
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 55 <223> oligonucleótido
 <400> 21
 ggatccgcc agtcgcgcag 20
 <210> 22
 <211> 21
 <212> ADN
 60 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 22
 65 gagtcgctca cctactgcat c 21

	<210> 23	
	<211> 20	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido	
	<400> 23	
	gagaactcgc cgatgtcgag	20
10	<210> 24	
	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
15	<223> oligonucleótido	
	<400> 24	
	agcgactggt agagctggc	20
	<210> 25	
	<211> 20	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido	
	<400> 25	
25	ctctgacga cacggcttac	20
	<210> 26	
	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
30	<223> oligonucleótido	
	<400> 26	
	caggccgcag aggggaac	19
	<210> 27	
	<211> 26	
35	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido	
	<400> 27	
40	gatgagttg gacaaaccac aactag	26
	<210> 28	
	<211> 22	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
45	<223> oligonucleótido	
	<400> 28	
	ggttcagggg gagatgtggg ag	22
	<210> 29	
50	<211> 19	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> oligonucleótido	
	<400> 29	
55	gtaccggcgc gccacctg	19
	<210> 30	
	<211> 25	
	<212> ADN	
	<213> Secuencia Artificial	
60	<220>	
	<223> oligonucleótido	
	<400> 30	
	gaatcctgtt gccgtcttg cgatg	25
	<210> 31	
65	<211> 27	

5 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 31
 ttatgattag agtcccgcaa ttataca 27
 <210> 32
 <211> 23
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 32
 15 ctacgttagc ttgagcttg atc 23
 <210> 33
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 <220>
 20 <223> oligonucleótido
 <400> 33
 gtgctgacgc gggcatagcc cag 23
 <210> 34
 <211> 24
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 34
 30 atgcgactaa aacacgcgac aaga 24
 <210> 35
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 35 <223> oligonucleótido
 <400> 35
 agcagcggag ggggtggatc 20
 <210> 36
 <211> 37
 40 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <220>
 45 <221> misc_feature
 <222> (27)..(27)
 <223> n es a, c, g, o t
 <220>
 <221> misc_feature
 50 <222> (31)..(33)
 <223> n es a, c, g, o t
 <400> 36
 acgatggact ccagagcggc cgcgcangca nnnggaa 37
 <210> 37
 <211> 40
 55 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <220>
 60 <221> misc_feature
 <222> (30)..(30)
 <223> n es a, c, g, o t
 <220>
 65 <221> misc_feature

<222> (34)..(36)
 <223> n es a, c, g, o t
 <400> 37
 5 acgatggact ccagagcggc cgcgctgatn gctnnncggt 40
 <210> 38
 <211> 16
 <212> ADN
 <213> Secuencia Artificial
 <220>
 10 <223> oligonucleótido
 <400> 38
 acgatggact ccagag 16
 <210> 39
 <211> 19
 15 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 39
 20 catacacgg gcaaagcag 19
 <210> 40
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 <220>
 25 <223> oligonucleótido
 <400> 40
 ttccatggg ccagcaggcg acgc 24
 <210> 41
 <211> 29
 30 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 41
 35 ttgatatca ggcgataata ttgagctcc 29
 <210> 42
 <211> 28
 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> oligonucleótido
 <400> 42
 ttccatggg cgggagtga ggggtgcc 28
 <210> 43
 <211> 28
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 43
 50 ttgagatcc tactgcgcct gctgctcc 28
 <210> 44
 <211> 30
 <212> ADN
 <213> Secuencia Artificial
 <220>
 55 <223> oligonucleótido
 <400> 44
 ttactagtc tctcagacg gctgtctc 30
 <210> 45
 <211> 19
 60 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 65

5 <400> 45
 caccgctct ctccgcttc 19
 <210> 46
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 10 <400> 46
 cctcgctctt tcgctggtc 20
 <210> 47
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> oligonucleótido
 <400> 47
 ccgccaataa cctcacctca g 21
 <210> 48
 20 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 25 <400> 48
 ggcgatggga gcgtagaata c 21
 <210> 49
 <211> 24
 <212> ADN
 30 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 49
 caggtttcat cgcaactaca ttga 24
 <210> 50
 35 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 40 <400> 50
 aacagagcga gttgaagagt agcc 24
 <210> 51
 <211> 21
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 <400> 51
 ccgccaataa cctcacctca g 21
 <210> 52
 50 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> oligonucleótido
 55 <400> 52
 gaagaagacg ccgagtagga tg 22
 <210> 53
 <211> 6
 60 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> motivo peptídico
 <220>
 65 <221> misc_feature

<222> (2)..(2)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <220>
 5 <221> misc_feature
 <222> (4)..(5)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <400> 53
 Gly Xaa Gly Xaa Xaa Gly
 10 1 5
 <210> 54
 <211> 6
 <212> PRT
 <213> Secuencia Artificial
 15 <220>
 <223> motivo peptídico
 <400> 54
 Gly Ser Gly Thr Val Gly
 1 5
 20 <210> 55
 <211> 597
 <212> ADN
 <213> Ustilago maydis
 <400> 55
 25 gattagatct tgctgatagg caggtttgct tggagaatgg ggggaaaaga ctgaccgaag 60
 aaacagcgag atctagaagt gataagcgga aagaatctga cttgctgtga tcagcagcca 120
 attttttttt cgtttttttt ttttcactcc acatcgctcg gcgtgcacgg tctgcatgtg 180
 30 taaattgtat tcatcgaaag ccacagtga atacatcagc cogatgtgga tttcgaaaac 240
 caattaatct tggaattcac gcgctcagat cagtccatag agtcgacttc ggctgtttcc 300
 aagagcttct tctctgcgag gtggttgccc gtgtttctcg ctgggaaaaa aggatcgatt 360
 35 attattcgct tctacctcgc tcgcaccctt ggctgctga aggaaacagc gccgagactc 420
 ggtcacgggt gctgggctcc gtgttgatgc tgggacggcg caaagtgggg ccgcgcact 480
 cttcgagcca aggacctcac tcttcaagaa caagcgctgt cgcctcgtc ttcttctttc 540
 40 tgctccacca tcgaatcttt ctttctcggt tcgaaaccaa aacactcttc caccatg 597
 <210> 56
 <211> 1434
 <212> ADN
 <213> Rhodosporidium toruloides
 45 <400> 56

50

55

60

65

5	ctgcagaact acgccctctc acacccaact tccgactcga ccggcgggtac gagcacgacc	60
	tactttctact gcctgccatc gacatccggg cgggtcgctg cctaccctgt gcgttctgog	120
	ccctccctcg tctcgggagg cagtgtctga cagaagcttt ggcgcagta ccccgtaag	180
10	atgcaactct acgcaacgtt cggcacagaa gtcgccaagc tccgcgcac cccgcctcaa	240
	gctctcgccg tgcgcgacgg tgtcgtctat tacgaggcgg agaagctcga gttgccggct	300
	ttgccagcgg cggtaaggt tgaggtggag acggagaagg cgggagtagc gggggaggac	360
15	aatgaggcga agggtagat ggtgctggtg gagactctta cggtaggaca ggaggagatt	420
	gaattgggct cgggagtcgt gcagattgag gactcgttgc tcgtcaagct ggaggtcagc	480
	ggctgatcct tccgttcgtt gcaaggatcg tctgcatgtt tcgttctct caatgacaca	540
20	acctggagag cgctcccgtc agcgagaatc gaggacattc cgcagctcgt gagcaagcgg	600
	aggtgcgagg ctccctcgaa agctgcgcct cttcagacgg cttgttctct cctgctctgg	660
	tgggctggcc tgacatgtaa tgtgctccgc cgcaagtcgg tcgtcgggtc caattcgacg	720
25	ttgaaagggc atagcgcaag gaagaacct ctgcggacat gcagaattac tggctcgccct	780
	gctccttcgt ctactggaat aagtccgtgc tcgttaaagc cccaacgtcg tttttcgacg	840
30	tttgtaaggc gcaagaggtg ctatgggcta cgcaggaagc tgagaggaca tagaagtcgg	900
	gggagggaac gcgcagagcg gcagttcgcg aagcatgagg aaagcgagac ggtccagcat	960
	ctgcagcgcc aatccgcaat ctctcggttg agcctgcacc ggaagcgctg gaacagtatg	1020
35	cgcagagtcg aacgcaagta agaaagacgc accctcacac tcgttactt cgagccatac	1080
	aacggatcaa agctgcgcgt atctcggtt gtaagggccg gaaagcaacc tcggagatgg	1140
	acacgtcaca tcaccaactt atcgatctcg gccgtcgacg tcgcagagag ggcgagagaa	1200
40	gcggtgaagg agggaaacaa cccctcgaga gcatgatccg accgaatctg cagcgcagga	1260
	agccgttaca agccgcctc gagcgcaggt cgggtccagc cgggggacga aacgcgcgag	1320
	gctgattcgt gagcgaagga agccgcacg acaagttcgc tcccctttgc cctctttccc	1380
45	atcacccgtt ctgccttac ccgtcagaa caacaccaga tcaactcaca tgtc	1434
	<210> 57	
	<211> 888	
	<212> ADN	
	<213> Aspergillus nidulans	
50	<400> 57	
	ctgtacagtg accggtgact ctttctggca tgcggagaga cggacggacg cagagagaag	60
	ggctgagtaa taagcgccac tgcgccagac agctctggcg gctctgaggt gcagtggatg	120
55	attattaatc cgggaccggc cgcctcccg ccccgagtg gaaaggctgg tgtgccctc	180

60

65

	gttgaccaag aatctattgc atcatcggag aatatggagc ttcacgaat caccggcagt	240
	aagcgaagga gaatgtgaag ccaggggtgt atagccgtcg gcgaaatagc atgccattaa	300
5	cctagggtaca gaagtccaat tgcttccgat ctggtaaaag attcacgaga tagtaccttc	360
	tccgaagtag gtagagcgag taccggcgcc gtaagctccc taattggccc atccggcatc	420
10	tgtaggggcgt ccaaatatcg tgctctcct gctttgccc gtgtatgaaa ccggaaaggc	480
	cgtccaggag ctggccagcg gcgcagaccg ggaacacaag ctggcagtcg acccatccgg	540
	tgctctgcac tcgaactgct gaggtccctc agtcctggt aggcagcttt gcccctctg	600
15	tccgcccgt gtgtcggcgg ggttgacaag gtcgttcgt cagtccaaca tttgttgcca	660
	tattttctctg ctctccccc cagctgctct tttctttct cttcttttc ccatcttcag	720
	tatattcatc ttcccatcca agaacttta tttccctaa gtaagtaett tgcatactc	780
20	atactccatc ctcccatcc cttattcctt tgaaccttc agttcgagct ttcccacttc	840
	atgcagctt gactaacagc taccgcgtt gagcagacat caccatgg	888
	<210> 58	
	<211> 617	
	<212> ADN	
25	<213> <i>Aspergillus niger</i>	
	<400> 58	
	gtcgacgaga tcgtaggagt gagtaccggt cgtgatggag ggggagcacg ctcatggtc	60
	cgtacggcag ctgccgaggg ggagcaggag atccaaatat cgtgagctc ctgctttgcc	120
30	cgggtgatga aaccggaaag gactgctggg gaactgggga gcggcgcaag ccgggaatcc	180
	cagctgacaa ttgaccatc ctcatgccgt ggcagagctt gaggtagctt ttgcccctc	240
35	tgtctccccg gtgtgcgcac tcgactgggc gcggcatctg tgctctctcc aggagcggag	300
	gaccagtag taagtaggac tgacctggtc gttgcgtcag tccagaggtt cctccctca	360
	ccctttttct acttccctc cccgcgcgt caactttct ttcctttta cttctctct	420
40	ctcttctct tcatccatcc tctcttcatc acttccctc tcccttcatc caattcatct	480
	tccaagttag tcttctctcc catctgtccc tccatcttc ccatcatcat ctcccctcc	540
	agctctctcc ctctctctgt ctctcaaga agcttgacta accattaccc cgcacatag	600
45	acacatctaa accatgg	617
	<210> 59	
	<211> 388	
	<212> ADN	
	<213> <i>Ashibia gossipii</i>	
50	<400> 59	
	cactatacgt gcctcgtccc cgcgggtca cccggccagc gacatggagg ccagaatac	60
	cctccttgac agtcttgacg tgcgcagctc aggggcataa tgtgactgtc gccgtacat	120
55	ttagcccata catcccatg tataatcatt tgcatacata ctttttgatg gcgcacggc	180
	gcgaagcaaa aattacggct cctcgtgca gacctgcag cagggaacg ctcccctcac	240
	agacgcgttg aattgtcccc acgcgcgcc cctgtagaga aatataaaag gttaggattt	300
60	gccactgagg ttcttcttcc atatacttcc ttttaaaatc ttgctaggat acagttctca	360
	catcacatcc gaacataaac aaccatgg	388
	<210> 60	
	<211> 963	
65	<212> ADN	

<213> Rhodoturula grammis
 <400> 60
 5 ggcgcgccta cgtctacgtc aagggcaatg ccgacgtgac caaggccatc ggccaggacc 60
 tcgccttctt ctcggtccct gtcgagctcg gcgtgcgtcc cgcctctctc tctctctctt 120
 tctctcgcc gcgcctcacg tgatccacga cgtcgtactg acccttgcca atgtgcgcgc 180
 10 ccgcagccca acggcgtcga gaagggtcac ccgctcggcg acctgacggc gtctgagaag 240
 gagctcctcg aggcgtgcct cggcgagctg cccgggtcca tctccaaggc cgagtcgttc 300
 atccagggct ccaagctctg actcgccggc gcctcgacgg gcgcgagcca caaggcgagg 360
 15 atgtgagagg aggcgtttcc tccaccttg accccatctg ccgcctccct ttctctctct 420
 ttctttccct tctctctctc ctctctctct ctctgtctcc tcttctggg cctctcgac 480
 20 ctcttctctg ccgtcgactc gtgaaaatgc agtgcgcgtt tctgtacctt gtctgcgag 540
 agagatctgg ttctgcgagg gtgagtcgtt gccttgcccg tggcacgcct cgcgcgagcg 600
 agagagaaga ggccacggtc caggacgacg acgacgagga ggaagcgcaa aaggcgagac 660
 25 accgagtgcc atcgattccc cgctcgaacc tgctcacggc tgtcgaaggc ggtgcgccac 720
 ggtgcttgcg ggagcgaaag caagctggcg tcgtcctctt gaactggttc gagtccgtga 780
 gggcggcgac gagaactcag gcgaggtgct cgcgtcggaa caagccgggc ttgtggtcga 840
 30 gggagcgaga gcgaggcagc gccgtcgtcg ccgaggcaag agcggcatcg acaagttggc 900
 ccgtcgctc tcgtccctc ttctcctcct ccaccacca cctttctcca gctcgaacca 960
 35 tgg 963
 <210> 61
 <211> 1009
 <212> ADN
 <213> Sporobolomyces roseus
 <400> 61
 40 ggcgcgccga agttatacct cagaggtctc aaaaacgaaa aagtcattga agaattcct 60
 ttgacgtgag ggttatttct ctctctctag tagtctacga gaatcgcaaa gatcgaaaa 120
 45
 50
 55
 60
 65

	ctgatgcac	tttgtgttca	cgggttagcg	atttgatctt	ttcgattccc	aaaatcgtat	180					
5	cgttcctgtc	gcagggaaact	acgctcaaag	cgggcaactc	gatcatcacg	ggagtgaatt	240					
	ttgagctctc	cctctatgag	agtgcagggt	tcgtcgtga	tgggtgaatc	cgctcatgcc	300					
	ttccctctca	ccttctcctt	tgtccattct	ctctactacg	gttgtcacat	cttccttctc	360					
10	cgacagaccc	cgcacggaat	tggagcgtac	tgaatcctc	cggaattctt	caaggacgga	420					
	gacgtcttca	gggtcgagat	ctcgggaggc	atcgggagtt	tggtaacaa	gatcgaatat	480					
	gaaaagtaga	taatccgtta	ctcagggtcaa	tggatatggct	tgaagatgc	tggaatcagc	540					
15	cggaaagcaa	agctgggagag	aaaaatcgag	attgcgaaac	gtgcgatgtc	atttcgtttc	600					
	gagctcgcaa	ccatctcgta	tccctctgag	ctacatacaa	acgtcactac	ggcctcggag	660					
	tgactccctg	cgagcggatt	gaaggagatc	acggtcgaat	cagctagacc	ttcgcaacgt	720					
20	tttcgcgctc	gcacgttctt	atcgatctac	tgagattgac	tcgaaaaagt	cttctctcac	780					
	ggtcgattga	actttgaatg	aactctcagg	ttgcgcgaga	gccaatcga	gccgaccaga	840					
	ggcaattcgg	agcttcccg	aacgttccaa	ggagagggat	tttcgagag	attacgattg	900					
25	cgagatagaa	aaaaggctag	ctttcgattt	cgagagagat	tactttcaag	ttcgtctgctt	960					
	ccaactcttg	ctccaacccc	ctccactcct	tctctacaaa	acaccatgg		1009					
30	<210> 62	<211> 2148	<212> ADN	<213> Rhodotorula glutinis ATCC 204091	<400> 62	ggcgcgccac	catctcctcg	tgccttcttc	cctctccttc	ggcgcgccaca	ccgcttcgca	60
35	gggctcacgg	actgctcaca	tcgtttgtgt	gcgtcgtgt	gcatgtccac	gcaccactcc						120
	cagcccccac	gagcgcctca	aaagacggcg	acgcagacgc	ccgcgaacg	acggcacgcc						180
	cctcttctca	ctagcgcgac	gaaccagctg	cgacgattcg	tgcgcttatg	ttagccggac						240
40	ttctggcttg	ctttgcgctg	ctgcgtccgt	cttggtgtgc	ggatcggctc	gatgggggtt						300
	tgctcgtttg	ctgggagacg	gtcgctctc	cctcctcctc	ttcactcctc	gttagctttc						360
45	tacgctcatt	ggttctgcga	accatctaca	tcacgctcgc	tcgtcatgct	cgtactacga						420
	tcaacacccc	tgtcgtcgt	gctttccctc	ctctccgtcc	tctcggccgc	gtccagcgac						480
	ttgccagcc	aacttcccc	gcacgcgggt	gagctctcca	cacttccttg	cgaccccaac						540
50	ccagcatctg	acatccgcat	cacgcagccc	tccgccttc	ccactcctcc	ctcttcaccg						600
	actcctcctc	ctcctccct	gattcctcgt	ccctcaaagc	cccgcagcct	cttccttca						660
	aatcaagcg	ccccgcctcg	ctcgaacaag	tcgacgagaa	cctcgggaag	aggctggcga						720
55	agcgcggcga	ggaggggagt	aagacggaga	gggtgccgtt	tggtcagagg	agtgcgacgg						780
60												
65												

5 cggcgagtgc ggggtggacaa ggtggagcgg ggacggggag ggcgacgcag cgcgttacgg 840
 ggcgaggaag cagaggtgca ggaggaggcg gagggagtggt cgcgctgct cagcctgtcc 900
 cttcgactac ccagacggtc gagacaggct ctaagatcgt ctcgactggt cttctgaccg 960
 tagcgtcgcc gtcgacggca gatggaggag gcgggacggt caccaggtc gagacggcct 1020
 10 cctcaggggt attgatcacc agcacggcgg gagcggcgag ttcagcggcg gcgtcggacg 1080
 tcgctagcgc acaggcagcg gaggcgacgt cgagtacgag catgatcagc ggaggagcgg 1140
 cggtcggcgg gagtttaagc aggatgctgg cgggaggagt tgcgggtgca gccctgatcc 1200
 15 tcctcgtcgc gtgagcaggc gaagcgagga gctcatgtag atacagcata gacagtatat 1260
 atcgccagga tagcttgcaa cagccgccgg tcggtttatt ccattgtcct cgaccccatg 1320
 cgaaggcgag ctctgctcgt cagctggcca agctggccag cagacgagcg ttgggggtggc 1380
 20 ggaacgccaa cggcatggag taaagcagcc gtgaggatga cggaggagct cgggcgaggt 1440
 gatggggatt ctacgaggaa cagcagagcg gcgaggagga gaggaaccgg aagcacagtc 1500
 tcgtggccgc ttgttgacga tccagtgtc gctagagtgc tcgtcgtcat cagagcagtc 1560
 25 gaacaaagcg atgccctgaa gaacgatgag cgaatgagtc gaagcggcgt ctaccggtga 1620
 actcggggtg tggcaaatga gcgagacgag gagtgccgc cagagttgcc acgtcgaccc 1680
 cacgtcgaa tcgacgttga tagagtgaac gaagccattg cagaccccag aaggtggcca 1740
 30 tgttgtgaa gcgagggcag gagcgagggg agaaggcgag gaggaggagg ggctggggaa 1800
 gcccgctccg gaatggcgca gctgggtgcc ggggatgtgc gcgagtggcg gaggagtcga 1860
 gcgtgagagt tctggaacac ggggcgcgca caagggtcga gggccgtgac gagttcgccg 1920
 35 ggcggtggtc gggctgaggg cgagcgcgcg ttggggacga cgacccccga cgccctcgtc 1980
 cttcgtcctc accgcttccc ggagaacttt gctgtactct gcttctccct tcacactctc 2040
 40 acaccactc acacaccctt ccatccacac acaagctatc cgcacacctc tcacaccga 2100
 cccagctcg ccccatcctc ttgcacccg gctcatcgaa aaccatgg 2148
 <210> 63
 <211> 1027
 <212> ADN
 <213> Rhodotorula glutinis ATCC 204091
 <220>
 <221> misc_feature
 <222> (1)..(6)
 50 <223> sitio de clonación de KpnI
 <220>
 <221> misc_feature
 <222> (1022)..(1027)
 <223> sitio de clonación de NcoI
 55
 60
 65

ES 2 724 998 T3

<400> 63
 ggtaccgtgc gagaagaagc gaggcacgcg aagcggtaga agcaatgaag cgaggcgaga 60
 5 gcgagagagg cagggcctca gccatgtcca gctgatcggc tgtaacgtcg cgccgggcca 120
 gtctgttgaa tttgttgctg cgcctgagcg taatagaagt gcagtagtct actccgcatg 180
 ccgagaacgt cgaagagcgc gaagtaggga gtcgagggaa gcgagggtag caaacacagc 240
 10 aacgacaagc ggttcgctt cgcctaaaag ctggttgacg ttgttttgac gttttgaaga 300
 cagtacaaca gcagcaagag gcgtgcgaag cgttggtggc gagagcagcg acaaggaggg 360
 aggaatgagg gagtggtagc gagggctcgc aaacgggctg acgcctcgaa tggagacgtg 420
 15 cgagtcgttc ttcgacgtcc gagggatgcc gagcgccgag acggagcacg caacgagcga 480
 gaggagagca gccgcgaag gtgattcgag tggcgcaagc ggaggacgac gaggagacgg 540
 acgagggagg aggagggatg gcgagcgagc atcggaaggc ggggcgcgag agacggcgtg 600
 20 aggagccggg tgtggagagt ttgaggaggc gcgggatgag aagtggctgg gtgtgcggag 660
 tgagcggtag caaagagcgc acttagagtc tagagcgagg cagtagtagt agagctgtat 720
 gaatgaatac aaagtgtgaa tacaacagtt tgtaatgcca ttctgagctt ggacgtgtgc 780
 25 gcgcgagagg gcgacttgca agccagcgcc cgcctcgtct tcttcctctt gcacctcgcg 840
 tcaaccctcg catctcacac ctacactcgc attcaaagtg cgtacactct cccacgacac 900
 acggggacgg cgcacaccac cgcgcgtcgc ttgaacggcg tcgccacttc gagccgtcac 960
 30 tgacttcgtc ctgctcctcc ctctctact ctctgtact gtactgtgta ctggggggga 1020
 accatgg 1027
 <210> 64
 <211> 956
 35 <212> ADN
 <213> Rhodotorula glutinis ATCC 204091
 <220>
 <221> misc_feature
 <222> (1)..(6)
 40 <223> sitio de clonación de KpnI
 <220>
 <221> misc_feature
 <222> (951)..(956)
 <223> sitio de clonación de NcoI
 <400> 64
 45 ggtaccatg ctgctgctgc ccctcaaagg tctctcgtc cacgtccgac gagtctggac 60
 agctttcaca gtcccagag tgcaagagc aggcgggctc acggttccgc aaaggagcgc 120
 gaggtccgac cgcgggcccg tctccttgcc cgcctcgcct cacctcctct tgcagcaggt 180
 50 tcacctcttc gaggtcactc gctcgtcgc agcgatgcgc aggtacaagt acgctaggcg 240

5 agagcgtcga aagcggggtt ctgcgagga ctggacgctg cagagcgcgg tcgagagagg 300

ctcgagtggtc gctttgaccg ctgcacgcaa ggcattgcgt cctccgtttg agctcgcaga 360

10 tactgccgtg cgaagacgag cataggctgt ggctgcggta gcaaggagcc ggcgagagaa 420

agctgtgctc gagcaggacg agagacggtc cgcgcgcttg agaaggtcga ggtgaggcgt 480

cgcaaccggg ttggatctcg attctcggcg aactacggct tcggcgaggg ccaaagcgac 540

ggcaggccgc gcaagctggc caggcgagag cgcgagagtc gcgagctgaa gcgggcgcg 600

ggtagagcaa gctggggaag cgagagaggg agagagagag agtgaggggg tggcgagggtg 660

15 gagacgaggg gagcgggttg cttgcgcgcg cgcgagaggg atcgagggca gaggcgagcc 720

ccgagagtgg aaggaaggac gaggaacctt gcgtgcggag gcgcgcgcg cgcgtgccac 780

ctggctgagc acggggccga gcttgaggga gctgggggag cgcgagcgag acgagggcag 840

20 ggcgagcccc cgcgtggcgg ccgcctcgca acccaaggct cgccttgcc gccgctcttg 900

ctctctttcc tcacacctcg cgtctacca ctccaatctc acttcatcca ccatgg 956

<210> 65

<211> 1474

25 <212> ADN

<213> cepa de Rhodotorula graminis WP1

<220>

<221> misc_feature

<222> (1)..(6)

30 <223> sitio de clonación de KpnI

<220>

<221> misc_feature

<222> (1469)..(1474)

<223> sitio de clonación de NcoI

35 <400> 65

ggtacccaaa gggagaggag cgggaggagg atggtggtgc cggacagggc gagagggaaag 60

gtcgagggag agatgggagc gcgaaaggtc gaggcgggga gaggaggggt ggggtcggga 120

40 ccaagggggc agagaggctc acaaggacgg aggagcttac tccgccttga ccttgcggtt 180

ggcggtggtg ccctcgcgga agctgcggcg cgaggaaggt cgtcagggcc gggcccaggg 240

gaggaacgag gacggcgacg acgacgcacc cgttctttaa gcggcgcgag acgcccttca 300

45 ggtgctgcat gcgccagtg ccagtgggtg ggcgcgcgtt ggccttctgt cccactcgt 360

ctgcgcggcg agggagaagg cgaggtagc acgacggcgc gcgagggccg gacgaggctg 420

agagggggac gcacacgagc ggagcttggc cgcggggtag ccgcacgaag cgcacgactt 480

50 gtgctgcttg tggaacgagc ggcggccgca gcgacggcac agactgtggg accacgaggg 540

tcaaccgggt gctcgcgaga caggagcgcg gcttgtctcg aagcacgggc aaagagagcg 600

55

60

65

	ttggacgcac gtgtgactct tgggtgtggcg gagaccgaac gaggtggtac ccttcgtctg	660
5	tggggcgcaa ggaggagtgg gtcagcgtcg ggcctcgagg cgcctgggtc gtcgacctcg	720
	cccgtcccg atcctcgcg cgtcctgctc ctccctcttc caaccctgcg acgtgttgcg	780
	gcagcagcag cttgctggga catgtgggga gggcggaag gcgaggggag gtcgaggtgc	840
10	gaatgtgggt ggtcgcgctt ggcggggcag catgtcgtcg cggcctcgag ccgggcgggc	900
	gacctggtgg ccgggtcgag cgagaggcgt gggagggagt ggcgcaaag gcgtgcgctc	960
	agaggcgggt tgtcgaggcg tcgaggcgga cgaggtcgag gaggtcgagg tgggaagctg	1020
15	ctgctgctgc tcgggcgtcg tcgcccgtc ccgagtgccc cgtgcgcgcc cctgctgccg	1080
	ctccttgggc cgtcctggtc ccacctgcc gtgcgcctt ccacgagagc gcgagtgagg	1140
	ctgtgcgccg ggttgcgctc caactttgag agagagcgag gacgggggca tggctcgctc	1200
20	gccggcctcg ggtcgttcga ggggtcgggg gcgggttgag ggaggggtgt gcgaggtggc	1260
	gggcttacca ttgtcgcgtc ggagaggggg gtttggcctg cgagaagagc aggagacgag	1320
	aggccggggg aggcgaggcg gcgagcgggc gagacggctc ggaccaagcg cgcgcccca	1380
25	aagtctgctt cgcgcgtcgc gctcgccctc ctcttgctct ccacctctc ctaggaccac	1440
	aaaggcacc ttgtgtaggc gtaggtcacc atgg	1474
30	<210> 66	
	<211> 1327	
	<212> ADN	
	<213> Rhodotorula glutinis ATCC 204091	
	<220>	
	<221> misc_feature	
35	<222> (1)..(6)	
	<223> sitio de clonación de KpnI	
	<220>	
	<221> misc_feature	
	<222> (1322)..(1327)	
	<223> sitio de clonación de BspHI	
40	<400> 66	
	ggtacctgag cgggcgagcc gcgagcgagg agcgttgagg aggaaggag ggagggagga	60
	gggagagggg gggaggggac ccccatcttt ctattcatc acaaagacga cgggtcggag	120
45	gggtccctcg agtgtttggg ctgggcttcg gagtctcgta gcgagcaagt agtgtttctc	180
	tccgtttcga cagctcgat tgtcatttct tgttcattgt cgtttccggc gactgcagg	240
	acgctgattt tcggcggaga cgacaagcac gtgggttggtg agcagcgagt tgagcaagaa	300
50	aaagcggacg aaggccctcg tcgggggctt caagtcaaga ttctgcggag attctgcgag	360
	agactgcaag cgttgaacct gttgagatct cgtcggacga cagcacagtg tccgtctcgc	420
	tcaatgcgat aggaagcgag agagaggagg aggatatcgg aggaaggcgt gtttgcgttc	480
55		
60		
65		

gctccaggcg tcgcaagatc cggcgtagag cacaatcgtc gttggttcga cgtttgtagt 540
5 tgcgtcacgag tgaggggcga gccctggcaag caaagaaggg gacgagcgac tcggcagcta 600
tcgctggagg agggcgactt tgtggcccggt ttccgtcgag ctcgacgcga gtgagcgag 660
ggtcgggtccg aaccgatgcc atggacgcag tgagcgaggc cggatgtgcg atgctgtttc 720
10 aagcgagcga aggaaggag aaagcgagcg agaggctctc ctctgtctt cctcacgcct 780
tccgaaggcc gacaagaggc gtagacgtcg acgagtcaac ggtttgacgt cgctcaggcc 840
tgtagcgggt cgtcggaagc tgggaaagag aggaaccaac gagtaacaag cgcgagagtc 900
15 tcctcaaggc ggacaattgc ctgccttcgg tcccggtcga gctcttcag taccagcgag 960
ggcgaaagtc gtcgatgcgt gcgcattcaa ggccaagcgt cgcagtcgag aagagcgaga 1020
gtgaagcgag tgaagcgga gagtgagagc gggaatccg cgtacttacg agtgggttgt 1080
20 attccttctt gtaatggcag attacctoga ttggccacgt cacgttcgg gagtgcccg 1140
gcgtgggcaa aagggcgagc gcggcgctc tctctcttc ttcctcagca gacgagctct 1200
25 cccctcgagt acgtcgacgg gctcactaca gctagcaaca gcaaggctac cagccagct 1260
acacgccagc tcacccaact cacaccgctc gttgtgcgcg cgcgcgcgag gaaaactttg 1320
ttcagtc 1327
30 <210> 67
<211> 1509
<212> ADN
<213> Rhodotorula glutinis ATCC 204091
<220>
35 <221> misc_feature
<222> (1)..(6)
<223> sitio de clonación de KpnI
<220>
<221> misc_feature
40 <222> (1504)..(1509)
<223> sitio de clonación de NcoI
<400> 67
ggtaccgctc atcatcgagc gagggcagca gcgacctgcg gactggccga agaacttccc 60
cgcccgagcg ttcgaggaca aggacattcg cagcctcgc tctcagtggc ggtacatcaa 120
45 gctggcgacc acagacgacc tctcgccgac tgaggagaac acgacgtcct gcgcggtccg 180
gtacggcgag gactcgcagc tcgccatctt ccacgttccc ggcaaggggt acttctgcac 240
gcagcagatg tgtcctcgtg cgtcgcgctt gctctctctt tctttatttg ctggctgact 300
50 cgtgctgact cgcgcgaaac ctacagacaag cgcgccttca tctcagagca tggcatcgtc 360
tcggacgacg ggagtggcca cctctacgtc tctgcccgc tccacaagcg caacttccgc 420

	ctcgacaacg gcgactgcct caacgacgag gagtacaaga tcctcgcgtt tgatgtcaag	480
5	gaggagaacg gcgacttget cgttcagggt cctccgcctg acgagctcga cgttttgatt	540
	gggtgogtct cgttagccc tctctcaaag acctgagctg acccttctga ttgtccgcag	600
	ctcgtcgaag tggatggtgc gcaaagcgac cgcgaagcc ttcggtcgca acgcagcgac	660
10	agccatcgag tgcgtccct ccaagcttct gttttccgcg cgcacactag gctgacgaca	720
	agtctctgca ggatcgtcgg acctcaggc gaggttgacg aggacaagaa ggcagcggga	780
	acagagtgcg gcgaagcgga taagtcttgc gggacgcaca agctcgagtg gtgattcttg	840
15	cgggtccgtc acagccaatg tatctatctc tagatgtcct tctcgggtat atcagttggt	900
	cgtgcacgt agacgtcgtt tagcagctct cgttcagcca cttgcgaagg cccgcttctt	960
	cgacgacaag gacggcttcg cttcctttac ctcgtcgtct gagcgttctc aagggaacct	1020
20	cctacgccct tcttcgcaca ggagcggccg acgaggcagc cttgctggct tatcgtcgtc	1080
	tccgcctttc atgctcgagc aagtcctcct gcgagtgtcc cgacgtcggc ccgccttgcc	1140
	caaggtcgcc gactgtccta tcgcgacact gcgaatgcac tgctgtccgc gccggagact	1200
25	gtgcggcgcg aattgagggc aaagtcgtgc atttgcgaaa cggtatccgc tcgaagggcc	1260
	cacgatagac ctccaccggc ctcaaacttg gcgacagggt cgcttcgcac gccggacagc	1320
	aagttaggct ttggcgtcgt cgctcgcata cgctttgcgg gacccttat cgcgactgcc	1380
30	ggattcgatt ggcgatatct ctcgctcgtc ggcctcgtg gacagctgga cagtctctgc	1440
	agcgtcgaag cgacgtcgat aaagtcagcg acgtcctcgc gaaccaagaa gaatcaccgc	1500
35	cgcctatgg	1509
	<210> 68	
	<211> 1606	
	<212> ADN	
	<213> Rhodotorula glutinis ATCC 204091	
	<220>	
40	<221> misc_feature	
	<222> (1)..(6)	
	<223> sitio de clonación de Spel	
	<220>	
	<221> misc_feature	
45	<222> (1601)..(1606)	
	<223> sitio de clonación de NcoI	
	<400> 68	
	actagtcgcc agggaacgca gagaaggcgg gacgagcgag ggtgagtcgc gcaagtcgac	60
50	gaagaagccg aggtcaacgt cgagggttgt caagttctgg cagacggaca gcagccacac	120
	taatcgtctc cgacttctga tgcacgtcgc cgcgggtgtc agtccttagc cgtggaagga	180
	ggaagccagg acgttctcac cgggcgacgc attgcttgtg ccagacctcg aggcgagctt	240

gtcgcgtcga aagaccggcg acctcgccg gctcaaattc cagcccgagt gattgccctg 300
5 acaacaaagg cgaaaagctg aagccgggtac caaaggctcg tccatcgaaa gtcgcgtccc 360
gaagactggc gtogacggat ctgaccatcg ctgcccctcc tgcgtctgct ttgaggcacc 420
ttacagcctg ctgctctcgt tcggagcctc cgcacccgct tggcaggacc acctcgcgac 480
10 cagtgcctc ccttgcgatg gctcgccaag tcttgcatat tccggcgacg ttgcggaagg 540
cgcaggtggg gcaggagatg cggagcgttg tgaactgtcc gttgacgagg agtgcggcg 600
aaagaagggc ggtggtgaga gagtaggtga ggatcttgag gaggagttca ggaggaagtg 660
15 aggaagggc tgccgggtgac tggttaaggct gaagcatgat ggcgagtgtg gccaaagtat 720
ccgagcgacg atcaagagac gaaggacgag acaacgcttc agcgcgcgaa gagagcgagc 780
gaggaccctc ctggtcgaga ggctatccag tcgccaaccg gtacccatcc agtttgagg 840
20 gttgaaacac agctgagagg atcagcgagt ggtagcgcaa actcctaagg cgctgaacgt 900
caaggacagc gagcgtgagc gtgtggaagc gacttgcgaa ggccaaactc gtgtcgcgct 960
ggccaaccgc cgtgcgcgtt tgacgcgctt ctgcccctc cgcctattca gagagtatgc 1020
25 ttcgtcacgg cgtgggcgcc aacatcggcg caggagctgg cgggacggga agaaagccgc 1080
aaccgcggtc ctgacacctc aacgtcccgc gagggccgct cagcactccc agacgtctct 1140
gcttggtgtt ctacgtcgtc gcggcggtgt gcagagtcca gcgcgcgccc gtcgtcgact 1200
30 tctgacaagc gataaattcc gagaccagcg ggagaaggcg gaacgagagg aggaggcgag 1260
ctggcgctct tgcgacctcg ttgagcagtt caagcgagca gattgagcag cagtgcgtcg 1320
35 agtgagccaa ctacggttct catatcggtc cctgagcgat atcgatgagg cgaaggacga 1380
cgacgagcga actgatctcg cgctctccct cttccccttc actctttcca ctcaaaaca 1440
acacgtgcgt cttctctgaa cgctatcaga caatccagga ccatcgctga ccgcgcgctc 1500
40 actcgtgcc tgacttcato gcccaaccaa cccgctcgtc accgactgga tctctccctc 1560
ccctcacacc acctgttgcg ctgcgatact cctctcacag ccatgg 1606
<210> 69
<211> 1408
45 <212> ADN
<213> Rhodotorula glutinis ATCC 204091
<220>
<221> misc_feature
<222> (1)..(6)
50 <223> sitio de clonación de SpeI
<220>
<221> misc_feature
<222> (1403)..(1408)
<223> sitio de clonación de BamHI
55 <400> 69

	actagtcggc gaagaggagg ggtgttagga gggagagctg tgcgagggg agaggtcagt	60
	cagagcacga tacgcgagca ggccaagcgg cttcatgttc actccaagct cgatgcggtc	120
5	gagaagtacg agctccttga cgagtgaag ggagagaaga gagactcacc aaccccggcg	180
	cgaccttcac gaaccgcaac ttccctcac acgcctcact acacgcctc tcaaacacct	240
10	cagcgtcgc gtagtgcgag agcttgaaga ataggtaggc gaggaaggag gaggcgaaga	300
	gggcggccat ctagcgaaca agttgggtta gctgggtggg ggaaaggga gaggggagga	360
	agagggggag ggtaccatga tccagcctat gtcgacttgc agttggaggg gcatcctgtg	420
15	cgcgtactgg cgtcagcggc ggacgacaag gcaggtagac gagagagggg gccggggaac	480
	gcactcgcag ttgcgcacgt ccgagaggaa catgtacgag tagccggccc acatgacgct	540
	cagcagcgcc gcaagcgtgt agtggagat gagcctgcc attggcagtc agcgcgcagc	600
20	cggacaaaact ctgggtgaag agaaaggaga gaacgcacca cttctcctcc ttcagcactc	660
	ccctcgccca gaccgcccc aacacaatac agacgagatc gatagttgcc ccagtcacag	720
	cagcgcgag ttggtgttag acgaggtgtt tgagccgtct gtggcgggga gaggtgcgga	780
25	tgcgtggat ctgcgcggg gagagttttg ggacgaagtt ggggtcgtcg gccatggtga	840
	gcccgctgta gtagctgtct gagcgagcct agtagtgcgc tggacgagca gagcccagag	900
30	tgcgagcag cgtgagcagg agacgaggtt cggaggttcc gcggagggcg acgagacgac	960
	gagcgagcct gggagaagcg cgagcatgtc cagcagcgta gtctcgaggc cgccagcagt	1020
	agtagagcac agcaatgagg caggagggag cgcaaggag ggaaagagcg cgacgaagg	1080
35	tgcaggtgat gaagtccaag gacaggggga ccaccctcgc ccgcttctcc ctgctctcc	1140
	ccacgaagt accacttcta aggttgtaa ttcatccat acagtctaca tacacttgca	1200
	gccatccgct tccctcgca tgccagttc ggtcacctg ggactccgat gcgatgatgc	1260
40	ggccgagttg gcttcctcga ccgctctca cacgctcata ccagcctctc ccagcctgct	1320
	accgctctct ggctctgca aacacccact cgagcacacc caccacaacca gcgaactcgc	1380
	ccagcctttg aaccgcaatg gcggatcc	1408
45	<210> 70	
	<211> 1480	
	<212> ADN	
	<213> cepa de Rhodotorula graminis WP1	
	<220>	
50	<221> misc_feature	
	<222> (1)..(6)	
	<223> sitio de clonación de Spel	
	<220>	
	<221> misc_feature	
	<222> (1475)..(1480)	
55	<223> sitio de clonación de Ncol	
	<400> 70	
60		
65		

5 actagtctgtg cagaaggaac ccgaggaggt cagtgcgcgg tcgagagagg aaaaggagag 60
 agagacgcac cgagcgaggc attgcgcgtg ccagcgctca gcctgtacct cgcgactgct 120
 gatgaaagcg tcgtcgccgc ccgacatgac cgcgcgctcg ctcgactggt cgcgcgacgac 180
 gtcgacgccc gaaggaggtg cgcgccagcc gcaggcggcg cgcctcaggc tcacgtcgac 240
 gtgcgggacg ttcgacgctc gccacgggtc gtgctcgagg gctttgacga cgagctcgtc 300
 10 gcgctcgcg ctgcccacc cattggcgaa cgtgacgtcc tggccgagca cctcctcgcc 360
 gatggcgcg caccaggcgc agacggacgc gacctggcgg taggcgcacg tgcgacaggt 420
 gaccgtgagg tgcgcaaaact ggccgtggac gaggagaggc tcagagatga gcgaggtcgt 480
 15 gaggggcgtg cggatgacct tgaggaggag ctgctgcgc acggcagaga gggtcggcat 540
 gatggtgtgc ggcgcggtcg gcagtcctga gagagatgtg tagaggaaga acgatgtgc 600
 20 cagatcggtc gagcaggagc cggctgcgagg cggctcgagg accgtcgcg tcgaggaccg 660
 gtcacggctg gacgatcgag gagacgcgc cccgtcgagc gcagcggcca gacgcaagcg 720
 agcacctttg aggtctgtact ccaaaacccg gagcgccgc tcgggagccg tgcctcgc 780
 25 ggatcctcgg tcgacagcgc cgagtcggag agggccagcc gacctcgggc cgcgcgacgc 840
 ccggccgcag ctctcgggt ccgacctgca gctcatcca gcagatcgac tttgagagcg 900
 aagcccccag gaagctgcct gagcgacctc gaggttggg aaggtcgccg agccacggct 960
 30 gggagagcga gctccctcac agtcgagacc ggctccaagt cgaatcgac actcgtagct 1020
 gcaccgcaaa agtgtgtgca gagctggagc gagcgaccgc gcgaggcgcg agggtcgcga 1080
 gaaagcgggc gagcgggtgc agtgcgcccg agacgccgag agagggcgcg agggcgagcg 1140
 35 ggcctcgcga gccctctgga gcgtgcagag gcggcgggga ggagcagagt gaggagggga 1200
 agaccctcca gagctggcag gagccaacgg agcgcggaaa tcagtgagat cgatgcggtt 1260
 40 ctcgagacga ttcgaccgcc ctgctcgta acgtcgccc ctgctccctc tcctcttccc 1320
 accacctctc cggtaacctc acacgagtgc gttctgtccc gagatctgat ctgacgccc 1380
 caggcactg actgaccgcc cacctgtct cctcgcccg tccacactc tccttcga 1440
 45 cctccacct cctcgctcaa cccctctgc ctgccatgg 1480
 <210> 71
 <211> 1532
 <212> ADN
 <213> cepa de Rhodotorula graminis WP1
 50 <220>
 <221> misc_feature
 <222> (1)..(6)
 <223> sitio de clonación de SpeI
 <220>
 55 <221> misc_feature
 <222> (1527)..(1532)
 <223> sitio de clonación de NcoI
 <400> 71

actagtctggg aggattgatg atcgggttgc ggtacaaggc gaggccgagg acgagctgga 60
 5 gggcgccgag gacgagggtc gagatgccga gctcgcaggg gacgaggccg ccaatcgagg 120
 tcgtgtcaga cggccgtacc gagtgatctc atagcgggtc cgcgcacgca gaaagacgaa 180
 gagaaaactca cgacgagggg gctcagcttg acaaagttga gctttgtcgt gcagccagac 240
 10 gagctgcgat aggagcgagg gtcagtgcgc ccgctctcgg atgaatgcgc agtcgaggag 300
 ggacaggggc gcaccaggtg acgatgaaca tgtcaggagc gccgtactgc agcatcttgc 360
 ggaagaggta gaaatgccac gccgaagcga aggtgacggc aaactgcgag agggacgagg 420
 15 tcagtgcgag ggtccgcaac agggagaggg cactcgtcac ggaccatcat ccagcccagg 480
 tcgatctgta gctgatcagg gaagctgtgc cgggtgcgag agaggccaat gtcgaagctt 540
 ggcagctcgt cgaggaagaa gaggacggcg agggacgcac gccatgaact tgatgtgacg 600
 20 ctogaagacg accgagtagc cgagccacat cattgtgagg acgcggcgcg agacgtagtg 660
 gaagatgaac ctgagagggc aagaggtcag tctcgaaacg agggaggaag ccggctcgag 720
 caggacgagg cgggcgcaac ggacgcacca tttctcggcg gcgacgatat tgcgagccca 780
 25 cacggccccg agcactagac aggtgatgtt ggcgaccgcg ccgccgacgg cgagagcgag 840
 aacgctgtag acgaagtgtc tgatgcgtcg gtggattggc gctcgtcgga tcctcgcgat 900
 ctgggccttg gtgagcgggtg gtggcgggccc aagcgggtggg ccagcagctg tgcctatcgc 960
 30 agcagcgggtg cggcgcaaga gcgactgtgg agctcgaggg agaggagcgc ggcaggggaa 1020
 agcgagaccg aggaggagcg agcgcggaaca ggcgaggcgg accggacgtt ccggtgcggc 1080
 35 tcgactggcg tgcgagacga gcaggccgtc gccggaagca gccgtgtccg gcggaagag 1140
 ccaggcgcgc gagcggggcg gagcagacag cggcgggtccg agcgcgcggg gcaggttcga 1200
 cgaaagtcgg gctcgggtca ggctcgcgcg agcgcatgag atgccgtcga gcgagcccat 1260
 40 gtacagagtc gagcgagaga gcgaagtgcg tggaaggaga gtggtccaag agtggagcgc 1320
 cgtggagatg agacagatga tggcgaacct cggccacagc ctctcgggtcc tgccacagca 1380
 gctctgtgag tctccctgac ccgccagccc gcgcttcaga actcacagac cacctacaca 1440
 45 gactcgcgca ccagctcgaa ccgcgccaga ccaccgcctc gccgcctccc cacctcgact 1500
 gcttcggaac ctcaaaagct cgaccaccat gg 1532
 <210> 72
 <211> 1222
 50 <212> ADN
 <213> cepa de Rhodotorula graminis WP1
 <220>
 <221> misc_feature
 <222> (1271)..(1222)
 55 <223> sitio de clonación de NcoI
 <400> 72

5 ggtaccgcga ggggaagggc aggagagtcg ccggaacacc gggcaaggag acaaggacaa 60
 agagcgagcg cgcacgcacc gtctggcggc cgggctgggt ggggcgggtc cgagtagcgc 120
 tgaccgaaga gcgaggagga ggaaaagcga gcgcggcggg cgtgggaagg agggcaagga 180
 gggcgggcgc gacgccaaaa aaggccagat tcgcggggca gtgatcgagt gccgtccgcg 240
 10 agctcaacca gcgagcgctc tctcgcgcg cgcgggtgcgc ttctttcgcc agccgatgcg 300
 caccgttcag aagcacgtcg cccgcaccga gagcgcctcc tcgcgagcct gtgaccacct 360
 cgtcgacccg ctcccgcggt cttttctcgc cggcctggac gcgcgctatc agatcgtgcc 420
 15 catgagacaa gcgactcgtc gaaaggacga cgatctcgta gtactgggtc cctgcgcaac 480
 gctcagccgt ccgtccccgt caaagtgcct cggcggaggg gaccgtgcgc gagacgcca 540
 agttggcctc ctcaagtcgg tagatccagc ttaacgctat caagggttc atggtgtagt 600
 20 tggtcacac gtcagtttaa cattcagttc actgaaggtc ctcaagttca acctgggtgc 660
 gatcaccttt ttggctcggc ggcattgcgc ccttacacc gcacgggtct acttcccttt 720
 gcaagcgacc aagcgaagca tcctctcgct cgtaaagctg ccggcgagga ggtcagacgg 780
 25 gttggcgggc cgtcgaaggt cggctcacc tcaacgctgc cggctgacca cgcagggcga 840
 gctatcattg ctttgaaagc ttcgaaaacg ccagggcatg cacagaaagc cgcgcgcgag 900
 aggctcaagt tggcgccgag ctgcggtcga gagacgacga cgacgtggga gctccctcgc 960
 30 ctctcctcct ttctctccca ccccatcagc ccaagtgagt cgcctcgtct tccgcaaggg 1020
 tcagcgacg cgttgctccg cgacagggca gcgcgtgcgc tcaccagggt ccccggttcg 1080
 35 ccggcgaggt tggcaactgac gaggtgcctt gccccctcc gctccctcc cctttggcct 1140
 cctctctcgc acgcacactc tctccctgca ccccttgca cttcccgaca ctctccccc 1200
 ccttccacc gtccgacat gg 1222
 40 <210> 73
 <211> 1507
 <212> ADN
 <213> cepa de Rhodotorula graminis WP1
 <220>
 45 <221> misc_feature
 <222> (1)..(6)
 <223> sitio de clonación de KpnI
 <220>
 <221> misc_feature
 <222> (1502)..(1507)
 50 <223> sitio de clonación de NcoI
 <400> 73

	ggtaccacta cctcgtcacg acccaggggtg cgttcccctt ctgcgcggt ccacagccac	60
	gacgttgcca gtccctaacg cgtccgctcc cccgtgcagg tggtectccg caccggcgct	120
5	acacgcagac atcgtcgacg ccgtcgcaag acgcccactt tgccgtccag accctgtcgc	180
	gcttcggcct cgcgtacctc ctgcgctcgt ccaacacgct caaggacacg tggtcgagcg	240
10	tgtgcgcgcc cgcggcgccc aagggccccc aaccgcacgt cgacgacctc gagctcgaga	300
	agcgcgagca ccgcgacaag tggctcctcg ggcgcacat gggccagggc aagcaggact	360
	cggcgctcgg ggacgcagtc gctgtcgtga gtccctctact ctgcggcgtt ctcgagactt	420
15	ggggtgcgag attgaccttg cgtcccgccc tctcgttgc agcaattcca caagcacttt	480
	ccgcacctgc gctcggcgca cctcttcccc ggctttgtct ttacgtgcgt ccctctcgtt	540
	ccccctctc tccacgtgcg ccaccagcct gactcgcccc tcgcctgccc cgtcccgcag	600
20	caacgccctc gcgtcgacct cgtcgtccc ctgcgcgac ctgtcgtcgt acaacctcgt	660
	cgggccccctc gcggcgcgca tcctgcctt tggcaacctg cccgagacgt acgccgacgt	720
	gcccggttac gtgcgggcca acccggcagc gcgcagccaa gggctcgagt actgcaacga	780
25	gcgcacgaag ccgctcggga gcccgcgctg ggcgagggc gcgacgggcg caaaggtgtg	840
	ggacgggctg agggccatga tcgaggagtg agctgggtgg cggcgagcg aggagccgga	900
	gaggaggggc ggaacgtgtt tgagaaggtc gcgctttgct cgtcggtcgc gggcgagcc	960
30	gtggctgtag ccagtctcgc tttgcagtgt cactcttcta catagctgag caaggcctag	1020
	cgtcgcgaga gagctcgct gtggcgctcgt gtcgagggcc gagagcgtcg cgtcagggg	1080
	cgagctgctc gcgctcacc aaggggctcg agcgggtcgc gctcgacagg ggaccgagag	1140
35	ctgcaggaga cagacgggag gaaaaagctc tggcgagcga ggagcgggac cactgagt	1200
	ctggggaagc gacggacgag gatgagcgca tccactcttg agtttcgccg aggcgcgagc	1260
40	tggcggtcga caacgcagca agctcctct ctctctccac cacactcgcc cctagcacac	1320
	gtgagctctc ctccctcgcc actgtcgacc agcacacgct cgtccaccgc cctgtgcgcc	1380
	ctgtgcggct tgcggtcgag cgaggccgcg ggtcgggtct ctgccaccgc aggaaccatc	1440
45	gatgtcgtc acgcttcgct cctcgtcctc ctctctctcc caccgcgcgc agctacctac	1500
	accatgg	1507
	<210> 74	
	<211> 1646	
50	<212> ADN	
	<213> Rhodotorula glutinis ATCC 204091	
	<220>	
	<221> misc_feature	
	<222> (1)..(6)	
	<223> sitio de clonación de AscI	
55	<220>	
	<221> misc_feature	
	<222> (1641)..(1646)	
	<223> sitio de clonación de NcoI	
60	<400> 74	
65		

	ggcgcgcctg aagctgtaca tcgaggtgga cgacagcgcg ggcaaggatg cgccagcaat	60
	cgtttgtggg tcgcttcttt cctcgagca cgcttttgtc ggctccctga tcagcacaca	120
5	agctaactaa cgctctggtt tcgctggcag tcatgcacgg ccttggtctg tcaacctcgt	180
	tctgggaagc gcccttctcc cgctcaaac tgtcctccc cttccgcctc atccgctacg	240
	acttcgacgg ccacggtctc tcgcccgtct cgtccctcga cgcagcagat gacggcgcca	300
10	tgatcccgtc cgaacacctc gtcggggact tggcggtgt gatcgagtgg gctggggtgg	360
	agaaggttgc gggagtgtt ggacactcga tgagcgggt ggtggcgagc acatttgagg	420
	ccaagtaccc gcagaagctc gacaagctcg gtgagtcga ttgaacctc ctcgcgcgtc	480
15	tcttctccgc tgacgattcg tcgacttggc cctgcttctc gcgcagtcct ctcggcgca	540
	atcgctctc tgaacctac cgtccaaagc aacatgctca agcgagccga tacagtcctc	600
20	gaatccggcc tctcagcaat cgtcgacaaa gtcgtctccg ccgctttgtc cgacaagtca	660
	aagcaggact cgcctctctc ggcagcgatg gtgcgaacgc tcgtgcttgg aacggaccgg	720
	agagggtacg cggcgcgctg tagggcgctt gcgggtgcga aggaccggga ttactcgagc	780
25	atcaaggcgg agacgttggg tgcgttcgct tgttctcctt cctctgcttt tctccagca	840
	actgacgcaa gcgtctgcaa cacagtcgtc gcaggcgagt ttgactacct ctgacaacag	900
	gagacgaccg acgcgctggt caacgacatc ccggcgcgcg agaaggtcca gatggacagt	960
30	gtcggccact ggcacgcctg cgaggacccc gttggactcg ccaagatcct cgatgggttc	1020
	ttcttgacag ggaaatgagg ttgggaaggg gggatagact ggggagaacg gcaggtgctg	1080
	acgcagcgga cgtcggtcgg gaggactttt tcggggagga tattcgctga ctgactccga	1140
35	cgtcgcttcc ctccttgag tatcttcaga agggatggga ggaggcgaac tgcaagggtg	1200
	atgaacgaga caacgcgag ggaggaagcg ccggaactct cgggggcgaa gaaggagtgg	1260
40	tgtcttcgcc agcgaacagc ttccgggtg ggttgacag cgccagtaga attccagcgt	1320
	cgcaacagag ctctagtoga ccgcgatcac ccacaaggac gagagcgggt cgcgccttgt	1380
	ccgcttcccc atcctcgtcc tgccttgcct ctcttccta ccacactctc ccgcttgagg	1440
45	gctctcttcc tcgcttggcg ctcctgctac cgctactcta gactctcta gtctccctgc	1500
	acaaccatcc ctatccctc cgctctctc gcacacccc cacagcttcg ttccccaact	1560
	tcacttcoga tgcgtgctg cgctccctt tcgctggcg ggcccgcgcc tgcttcgag	1620
50	gacaactact gattgtggga ccatgg	1646
	<210> 75	
	<211> 1450	
	<212> ADN	
55	<213> Rhodosporidium toruloides	
	<220>	
	<221> misc_feature	
	<222> (1)..(6)	
	<223> sitio de clonación de KpnI	
	<220>	
60	<221> misc_feature	
	<222> (1445)..(1450)	
	<223> sitio de clonación de NcoI	
	<400> 75	
65		

	ggtacggcg cgctagtctg cagaactacg ccctctcaca cccaacttcc gactcgaccg	60
	gcggtagcag cagcagctac ttctactgcc tgccatcgac atccggggcg gtcgctgcct	120
5	accctgtgcg ttctgogccc tccctcgtct cgggaggcag tgtctgacag aagctttgcg	180
	cgcagtagcc cgtcaagatg caactctacg caacgttcgg cacagaagtc gccaaagctcc	240
10	gcgcacgccc gcccaagct ctccgctgc ccgacgggtg cgtctattac gaggcggaga	300
	agctcgagtt gccggctttg ccagcggcgg tcaaggttga ggtggagacg gagaaggcgg	360
	gagtagcggg ggaggacaat gaggcgaagg gtgagatggt gctggtggag actcttacgg	420
15	tggagcagga ggagattgaa ttgggctcgg gagtctgca gattgaggag tcgttgctcg	480
	tcaagctgga ggtcagcggc tgatccttcc gtctgttga aggatcgtct gcatgtttcg	540
	cttctctcaa tgacacaacc tggagagcgc tcccgctcagc gagaatcgag gacattccgc	600
20	agctcgtgag caagcggagg tgcgaggtc cctcgaaagc tgcgcctctt cagacggctt	660
	gttctctcct gctctggtg gctggcctga catgtaatgt gctccgccgc aagtccgtcg	720
	tccgtctcaa ttcgacgttg aaagggcata gcgcaaggaa gaaccctctg cggacatgca	780
25	gaattactgg ctgcctgct cctctgtcta ctggaataag tccgtctctg ttaaagcccc	840
	aacgtcgttt ttcgacgttt gtaaggcgca agaggtgcta tgggctacgc aggaagctga	900
	gaggacatag aagtcggggg aggaacggcg cagagcggca gttgcggaag catgaggaaa	960
30	gcgagacggt ccagcatctg cagcggcaat ccgcaatctc ctggttgagc ctgcaccgga	1020
	agcgtcggaa cagtatgcgc agagtcgaac gcaagtaaga aagacgcacc ctacactcg	1080
35	cttacttcga gccatacaac ggatcaaagc tgcgcgtatc tcggcttgta agggccggaa	1140
	agcaacctcg gagatggaca cgtcacatca ccaacttatc gatctcgcc gtcgacgtcg	1200
	cagagagggc gagagaagcg gtgaaggagg gaaacaacct ctcgagagca tgatccgacc	1260
40	gaatctgcag cgcaggaagc cgttacaagc ccgcctcgag cgcaggtcgg gtccagccgg	1320
	gggacgaaac gcgcgaggct gattcgtgag cgaaggaagc cgcacgaca agttcgtcc	1380
	cctttgccct ctttccatc acccgttctc gcottaccog ctcagaacaa caccagatca	1440
45	ctcaccatgg	1450
	<210> 76	
	<211> 672	
	<212> ADN	
	<213> Ustilago maydis	
50	<220>	
	<221> misc_feature	
	<222> (1)..(6)	
	<223> sitio de clonación de KpnI	
	<220>	
55	<221> misc_feature	
	<222> (667)..(672)	
	<223> sitio de clonación de NcoI	
	<400> 76	
60		
65		

	ggtaccggcg cgctagtcct taaaactgaa ggcgggaaac gacaatctga tccaagctca	60
	agctaagctc tagtgattag atcttgctga taggcagggtt tgcttggaga atggggggaa	120
5	aagactgacc gaagaaacag cgagatctag aagtataag cggaagaat ctgacttgct	180
	gtgatcagca gccaatTTTT ttttcgtttt ttttttttca ctccacatcg tcgtgcgtgc	240
	acggctcgca tgtgtaaatt gtattcatcg aaagccacag ttgaatacat cagcccgatg	300
10	tggatttcga aaaccaatta atcttggaat tcacgcgctc agatcagtc ctagagtcga	360
	cttcggctgt ttccaagagc ttcttctctg cgagggtggtt gcccggtgtt ctgcgtggga	420
	aaaaaggatc gattattatt cgcttctacc tcgctcgcac ccttggcctg ctgaaggaaa	480
15	cagcgccgag actcggtcac ggttgctggg ctccgtgttg atgctgggac ggcgcaaagt	540
	ggggcccgcg cactcttcga gccaaaggac tcactcttca agaacaagcg ctgtcgccat	600
20	cgtcttcttc tttctgctcc accatogaat ctttctttct cgtttcgaaa ccaaaacact	660
	cttcacccat gg	672
	<210> 77	
	<211> 963	
25	<212> ADN	
	<213> cepa de Rhodotorula graminis WP1	
	<220>	
	<221> misc_feature	
	<222> (1)..(6)	
30	<223> sitio de clonación de AscI	
	<220>	
	<221> misc_feature	
	<222> (958)..(963)	
	<223> sitio de clonación de NcoI	
	<400> 77	
35	ggcgcgcta cgtctacgta aagggaatg ccgacgtgac caaggccatc ggccaggacc	60
	tcgccttctt ctcggtccct gtcgagctcg gcgctcgctc cgcgcgtctc tctctctctt	120
	tctctcggcc gcgcctcacg tgatccacga cgtcgtactg acccttgcca atgtgcgcgc	180
40	cgcagccca acggcgctga gaaggtgcac ccgctcggcg acctgacggc gttcgagaag	240
	gagctcctcg aggcgtgcct cggcgagctg cccgggtcca tctccaaggg cgagtcgttc	300
	atccagggtc ccaagctctg actcgccggc gcacgcagcg gcgcgagcca caaggcgagg	360
45	atgtgagagg aggcgtttcc tccaccttg accccatctg ccgcctccct ttctctctct	420
	ttctttccct tctctctct ctctctctct ctctgtctcc tcttctggg cctctcgac	480
	ctcttcctcg ccgtcgactc gtgaaaatgc agtgcgcggt tctgtacctt gtctgcgag	540
50	agagatctgg ttctgcgagg gtgagtcgtt gccttggcgg tggcacgcct cgcgcgagcg	600
	agagagaaga ggccacggtc caggacgacg acgacgagga ggaagcgcaa aaggcgagac	660
	accgagtgcc atcgattccc cgctcgaacc tgctcacggc tgctgaaggc ggtgcgccac	720
55	ggtgcttgcg ggagcgaaa caagctggcg tcgtcctctt gaactgggtc gagtcogtga	780
	gggcgcgac gagaactcag gcgaggtgct cgcgtcggaa caagccgggc ttgtggtcga	840
60	gggagcgaga gcgaggcagc gccgtcgtcg ccgaggcaag agcggcatcg acaagttggc	900
	ccgtcgccct tcgtccctc ttctcctcct cccaccacca cctttctcca gctcgaacca	960
	tgg	963
65	<210> 78	

5 <211> 2148
 <212> ADN
 <213> Rhodotorula glutinis ATCC 204091
 <220>
 <221> misc_feature
 <222> (1)..(6)
 <223> sitio de clonación de Ascl
 10 <220>
 <221> misc_feature
 <222> (2143)..(2148)
 <223> sitio de clonación de NcoI
 15 <400> 78

15

20

25

30

35

40

45

50

55

60

65

	ggcgcgccac catctcctcg tcgcttcttc cctctccttc ggcgcccaca ccgcttcgca	60
	gggctcaccg actgctcaca tcgtttgtgt gcgtcgctgt gcatgtccac gcaccactcc	120
5	cagccccac gagcgctca aaagacggg acgcagacgc ccgcgaacg acggcacgcc	180
	cctcttctca ctagcgcgac gaaccagctg cgacgattcg tgcgcttatg ttagccggac	240
	ttctggcttg ctttgcgctg ctgcgtccgt cttgtggtgc ggatcggtc gatgggggtt	300
10	tgctcgcttg ctgggagacg gtgcctctc cctcctctc ttactctc gttagcttcc	360
	tacgctcatt ggttctcgga accatctaca tcacgctcgc tcgtcatgct cgtactacga	420
	tcaacacccc tgcctcgctg gcttccctc ctctcctcc ttctggcgc gtccagcgac	480
15	ttgccagcc aacttcccc gcacgcggg gagtctcca cacttcttg cgacccaac	540
	ccagcatctg acatcgcgat cagcgagccc tcccgcttc cactctctc ctcttcacg	600
	actcctctc ctctccctt gattcctgt cctcaaagc ccgcagcct ctcccttca	660
20	aaatcaagc ccccgctcg ctcgacaag tgcagcagaa cctcggaag aggctggcga	720
	agcgcgcgga ggaggggagt aagacggaga ggggtccgtt tggtcagagg agtgcgacgg	780
	cggcgagtgc ggggtgacaa ggtggagcgg gcagcgggag ggcgacgag cgcgttacgg	840
25	gcggagggaag cagaggtgca ggaggggag gagggagtgt cgcggctgct cagcctgtcc	900
	cttcgactac ccagacggtc gagacaggct ctaagatgt ctgactggt cttctgacg	960
	tagcgtcgcc gtcgacggca gatggaggag gcgggacggt caccaggtc gagacggcct	1020
30	cctcaggggt attgatcacc agcacggcg gagcggcgag ttcagcgcg gcgtcgacg	1080
	tcgctagcgc acaggcagcg gaggcgacgt cgagtacgag catgatcagc ggaggagcgg	1140
	cggctggcgg gagtttaagc aggatgctgg cgggaggagt tgcgggtgca gcctgatcc	1200
35	tcctcgtgcg gtgagcaggc gaagcgagga gctcatgtag atacagcata gacagtatat	1260
	atcgccagga tagcttgcaa cagccgccg tcggtttatt ccattgtcct cgaccccatg	1320
	cgaaggcgag ctctgctgt cagctggcca agctggccag cagacgagcg ttgggggtggc	1380
40	ggaacgcca cggcatggag taaagcagc gtgaggatga cggaggagct cgggcgaggt	1440
	gatggggatt ctagcaggaa cagcagagcg gcgaggagga gaggaaccg aagcacagtc	1500
	tcgtggcgc ttgttgaga tccagtgct gctagagtgc tcgtcgtcat cagagcgagt	1560
45	gaacaaagc atgcccgtga gaacgatgag cgaatgagtc gaagcggcgt ctaccggtga	1620
	actcgggggt tggcaaatga gcgagacgag gagtgcgcgc cagagttgcc acgtcgaccc	1680
	cacgtcggaa tcgacgttga tagagtgaac gaagccattg cagaccccag aaggtggcca	1740
50	tggttggaag gcgagggcag gagcagggg agaaggcgag gaggaggagg ggctgggaa	1800
	gcccgtccg gaatggcgca gctgggtgcc ggggatgtgc gcgagtggcg gaggagtcca	1860
	gcgtgagagt tctggaacac ggggcgcgca caagggtoga gggcogtgac gagtctcgcg	1920
55	ggcgggtggtc gggctgaggg cgagcgcgcg ttggggacga cgacgccga cgcctcgtc	1980
	cttcgtctc accgcttccc ggagaacttt gctgtactct gcttctccct tcacactctc	2040
60	acaccactc acacaccctt ccattccacac acaagctatc cgcacacctc tcacaccga	2100
	ccccagctcg ccccatctc ttgcacccg gctcatogaa aacctgg	2148
	<210> 79	
	<211> 1009	
	<212> ADN	
65		

<213> Sporobolomyces roseus
 <220>
 <221> misc_feature
 <222> (1)..(6)
 <223> sitio de clonación de Ascl
 <220>
 <221> misc_feature
 <222> (1004)..(1009)
 <223> sitio de clonación de NcoI
 <400> 79
 ggcgcgccga agttatacct cagaggtctc aaaaacgaaa aagtcattgca agaattctcct 60
 ttgacgtgag gggtatttct ctctctctag tagtctacga gaatcgcaaa gatcggaata 120
 ctgatgcata tttgtgttca cgggttagcg atttgatctt ttcgattccc aaaatcgtat 180
 cgttctgtc gcagggaact acgtcaaaag ccggcactct gatcatcacg ggagttagtt 240
 ttgagctctc cctctatgag agtgcaaggt tcgtcgctga tgggtgaatc cgctcatgcc 300
 ttccccctta ccttctcctt tgtccattct ctctactacg gttgtcacat cttccttctc 360
 cgacagaccc cgcacggaat tggagcgtac tcgaatctc cggaattctt caaggacgga 420
 gacgtcttca gggctcgagat ctggggaggg atcgggaggt tggtaacaa gatcgaatat 480
 gaaaagtaga taatccgtta ctacaggtcaa tggatggct tcgaagatgc tggaatcagc 540
 cggaaagcaa agctggagag aaaaatcgag attgcgaaac gtgcgatgtc atttcgtttc 600
 gagctcgcaa ccattctgta tccctctgag ctacatacaa acgtcactac ggctcggag 660
 tgactcctg cgagcggatt gaaggagatc acggtcgaat cagctagacc ttgcgaacgt 720
 ttctcgctc gcacgttctt atcgatctac tgagattgac tcgaaaaagt cttctctcac 780
 ggtcgattga actttgaatg aactctcagg ttgcgcgaga gccaatatga gccgaccaga 840
 ggcaattcgg agcttcccg aacgttccaa ggagagggat ttcccgagag attacgattg 900
 cgagatagaa aaaaggctag ctttcgattt cgagagagat tactttcaag ttogctgctt 960
 ccaactcttg ctccaacccc ctccactcct tctctacaaa acaccatgg 1009
 <210> 80
 <211> 1858
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 80

atggggccagc aggcgacgct cgaggagctg tacacacgct cagaaatctc caagatcaag 60
 caagtgcagc cagctcttct cctcaccacc ccacaacata ccccgagcc caccagcgt 120
 5 ctcaccacagc acccaccagcc tgctgaccag ctgcagtgca tccacagatt tgcgcccttt 180
 ggcgtccgc ggtgcgcgcg gctgcagacg ttctcgtct ttgcctggac gacggcactg 240
 cccatcctac tcggcgtctt ctctcgtctc tggcgctca ggcttggcgt ggattgggag 300
 10 tagcgggcca ctcagctgac ttgcgcaccc gctccagctc gttcccaccg ctctggccgg 360
 ccgtcatcgc ctacctcacc tgggtctttt tcattgacca ggccgcgact caccgtggac 420
 gggcgcagtc ttggctgcgg aagagtcgga tatgggtctg gtttcagga tattatcccg 480
 15 tcaggtgcgt cgtcccgctt gttgcgcgtc ttgcgacctc gctcacggcc aactcgcccg 540
 accggctacc tccgaacttc ccgccaacag cttgatcaag gttcgtccac ctttccttca 600
 20 gcttgagtga tctgtagagg agctgcagga tcaagcccaa cccggggagg acctcgagg 660
 acgacgccgc tgacttgctc tctcctaca gagcgccgac ttgccgcctg accggaagta 720
 cgtcttcggc tatcatccgc accgcgtcat aggcattggc gccatcgcca acttcgcgac 780
 25 cgacgcaacc ggcttctcga cactcttccc cggcttgaac cctcacctcc tcacctcca 840
 aagcaacttc aagctccgc tctatcgga gttgctgctc gccctcgga tctgctccgt 900
 ctcgatgaag agctgccaga acatcctgcg gcaagggtgc ccagtcattc cgaacgggcg 960
 30 gtcgagcgtg aactctgggg atgggaagag ctgaccttct gcctcactcc atccatgcag 1020
 gtccctggctc ggctctcacc atcggtgtcg gtggcgcagc cgagagcttg agtgcgcac 1080
 ccggaaccgc cgacctcagc ctcaagcgac gaaaaggatt catcaagctc gcgatccggc 1140
 35 aaggcgcga cctcgtgcc gtcttttcgt tcggcgagaa cgacgtgcgt cctctgctcg 1200
 acttcgcta gcgaagccct tcgctgacgc tcccggtttc ttccccaga tcttcggcca 1260
 gctgcgaaat gagcgaggga cgcggctgta caagttgcag aagcgtttcc agggcgtatt 1320
 40 cggcttact ctccgtacgt tgcgcgtgt cgttcaatc tgcgagcgt ccagtcgctc 1380
 acgcagctac aactcccaca gctctcttct acggtcggg actcttcaac tgtacgcccg 1440
 45 agtctacgt actagtctac cgtgggagc actgaagagc accgctgacg tcccacctct 1500
 ccgcgcagat aacgttggct tgatgccgta ccgccatccg atcgtctcgg tcggtgcgtc 1560
 cccctcgtc cctcctgacc tgcgggcttc agtaacaat tctcgacgac atctagtcgg 1620
 50 tcgaccaatc tcggtgcagc agaaggacca cccaacgaca gggatctcg aagaagtcca 1680
 ggcgcgggtat atcgcagAAC tcaagcgggtg cgtccagac gtctacctt gcccggtgtc 1740
 tcgactcgg taagacagat cactgacgct tcggtcactg gccgcgcagc atctgggaag 1800
 55 actacaagga cgcctacgcc aaagtcgca cgcgggagct caatattatc gcctgacc 1858
 <210> 81
 <211> 1047
 <212> ADN
 <213> Rhodosporidium toruloides
 <220>
 <221> CDS
 <222> (1)..(1047)
 <400> 81

5	atg ggc cag cag gcg acg ccc gag gag cta tac aca cgc tca gag atc Met Gly Gln Gln Ala Thr Pro Glu Glu Leu Tyr Thr Arg Ser Glu Ile 1 5 10 15	48
10	tcc aag atc aaa ttc gca ccc ttt ggc gtc cgg cgg tgg cgc cgg ctg Ser Lys Ile Lys Phe Ala Pro Phe Gly Val Pro Arg Ser Arg Arg Leu 20 25 30	96
15	cag acc ttc tcc gtc ttt gcc tgg acg acg gca ctg ccc atc cta ctc Gln Thr Phe Ser Val Phe Ala Trp Thr Thr Ala Leu Pro Ile Leu Leu 35 40 45	144
20	ggc gtc ttc ttc ctc ctc tgc tgg ttc cca cgg ctc tgg cgg gct gtc Gly Val Phe Phe Leu Leu Cys Ser Phe Pro Pro Leu Trp Pro Ala Val 50 55 60	192
25	att gcc tac ctc acc tgg gtc ttt ttc att gac cag gcg cgg att cac Ile Ala Tyr Leu Thr Trp Val Phe Phe Ile Asp Gln Ala Pro Ile His 65 70 75 80	240
30	ggt gga cgg gcg cag tct tgg ctg cgg aag agt cgg ata tgg gtc tgg Gly Gly Arg Ala Gln Ser Trp Leu Arg Lys Ser Arg Ile Trp Val Trp 85 90 95	288
35	ttt gca gga tac tat ccc gtc agc ttg atc aag agc gcc gac ttg cgg Phe Ala Gly Tyr Tyr Pro Val Ser Leu Ile Lys Ser Ala Asp Leu Pro 100 105 110	336
40	cct gac cgg aag tac gtc ttt ggc tac cac cgg cac ggc gtc ata ggc Pro Asp Arg Lys Tyr Val Phe Gly Tyr His Pro His Gly Val Ile Gly 115 120 125	384
45	atg ggc gcc atc gcc aac ttc gcg acc gac gca acc ggc ttc tgg aca Met Gly Ala Ile Ala Asn Phe Ala Thr Asp Ala Thr Gly Phe Ser Thr 130 135 140	432
50	ctc ttc ccc ggc ttg aac cct cac ctc ctc acc ctc caa agc aac ttc Leu Phe Pro Gly Leu Asn Pro His Leu Leu Thr Leu Gln Ser Asn Phe 145 150 155 160	480
55	aag ctc cgg ctc tac cgc gag ttg ctg ctc gct ctc ggc ata tgc tcc Lys Leu Pro Leu Tyr Arg Glu Leu Leu Leu Ala Leu Gly Ile Cys Ser 165 170 175	528
60	gtc tgg atg aag agc tgt cag aac att ctg cga caa ggt cct ggc tgg Val Ser Met Lys Ser Cys Gln Asn Ile Leu Arg Gln Gly Pro Gly Ser 180 185 190	576

5 gct ctc act atc gtc gtc ggt ggc gcc gcc gag agc ttg agt gcg cat 624
Ala Leu Thr Ile Val Val Gly Gly Ala Ala Glu Ser Leu Ser Ala His
195 200 205

ccc gga acc gcc gat ctt acg ctc aag cga cga aaa ggc ttc atc aaa 672
Pro Gly Thr Ala Asp Leu Thr Leu Lys Arg Arg Lys Gly Phe Ile Lys
210 215 220

10 ctc gcg atc cgg caa ggc gcc gac ctt gtg ccc gtc ttt tcg ttc ggc 720
Leu Ala Ile Arg Gln Gly Ala Asp Leu Val Pro Val Phe Ser Phe Gly
225 230 235 240

15 gag aac gac atc ttt ggc cag ctg cga aac gag cga gga acg cgg ctg 768
Glu Asn Asp Ile Phe Gly Gln Leu Arg Asn Glu Arg Gly Thr Arg Leu
245 250 255

tac aag ttg cag aag cgt ttc caa ggc gtg ttt ggc ttc acc ctc cct 816
Tyr Lys Leu Gln Lys Arg Phe Gln Gly Val Phe Gly Phe Thr Leu Pro
260 265 270

20 ctc ttc tac ggc cgg gga ctc ttc aac tac aac gtc gga ttg atg ccg 864
Leu Phe Tyr Gly Arg Gly Leu Phe Asn Tyr Asn Val Gly Leu Met Pro
275 280 285

25 tat cgc cat ccg atc gtc tct gtc gtc ggt cga cca atc tcg gta gag 912
Tyr Arg His Pro Ile Val Ser Val Val Gly Arg Pro Ile Ser Val Glu
290 295 300

cag aag gac cac ccg acc acg gcg gac ctc gaa gaa gtt cag gcg cgg 960
Gln Lys Asp His Pro Thr Thr Ala Asp Leu Glu Glu Val Gln Ala Arg
305 310 315 320

30 tat atc gca gaa ctc aag cgc atc tgg gaa gaa tac aag gac gcc tac 1008
Tyr Ile Ala Glu Leu Lys Arg Ile Trp Glu Glu Tyr Lys Asp Ala Tyr
325 330 335

35 gcc aaa agt cgc acg cgg gag ctc aat att atc gcc tga 1047
Ala Lys Ser Arg Thr Arg Glu Leu Asn Ile Ile Ala
340 345

<210> 82
<211> 348
<212> PRT
40 <213> Rhodosporidium toruloides
<400> 82
Met Gly Gln Gln Ala Thr Pro Glu Glu Leu Tyr Thr Arg Ser Glu Ile
1 5 10 15

45 Ser Lys Ile Lys Phe Ala Pro Phe Gly Val Pro Arg Ser Arg Arg Leu
20 25 30

Gln Thr Phe Ser Val Phe Ala Trp Thr Thr Ala Leu Pro Ile Leu Leu
35 40 45

50 Gly Val Phe Phe Leu Leu Cys Ser Phe Pro Pro Leu Trp Pro Ala Val
50 55 60

ES 2 724 998 T3

Ile Ala Tyr Leu Thr Trp Val Phe Phe Ile Asp Gln Ala Pro Ile His
 65 70 75 80
 5 Gly Gly Arg Ala Gln Ser Trp Leu Arg Lys Ser Arg Ile Trp Val Trp
 85 90 95
 Phe Ala Gly Tyr Tyr Pro Val Ser Leu Ile Lys Ser Ala Asp Leu Pro
 10 100 105 110
 Pro Asp Arg Lys Tyr Val Phe Gly Tyr His Pro His Gly Val Ile Gly
 115 120 125
 15 Met Gly Ala Ile Ala Asn Phe Ala Thr Asp Ala Thr Gly Phe Ser Thr
 130 135 140
 Leu Phe Pro Gly Leu Asn Pro His Leu Leu Thr Leu Gln Ser Asn Phe
 20 145 150 155 160
 Lys Leu Pro Leu Tyr Arg Glu Leu Leu Leu Ala Leu Gly Ile Cys Ser
 165 170 175
 25 Val Ser Met Lys Ser Cys Gln Asn Ile Leu Arg Gln Gly Pro Gly Ser
 180 185 190
 Ala Leu Thr Ile Val Val Gly Gly Ala Ala Glu Ser Leu Ser Ala His
 30 195 200 205
 Pro Gly Thr Ala Asp Leu Thr Leu Lys Arg Arg Lys Gly Phe Ile Lys
 210 215 220
 35 Leu Ala Ile Arg Gln Gly Ala Asp Leu Val Pro Val Phe Ser Phe Gly
 225 230 235 240
 Glu Asn Asp Ile Phe Gly Gln Leu Arg Asn Glu Arg Gly Thr Arg Leu
 40 245 250 255
 Tyr Lys Leu Gln Lys Arg Phe Gln Gly Val Phe Gly Phe Thr Leu Pro
 45 260 265 270
 Leu Phe Tyr Gly Arg Gly Leu Phe Asn Tyr Asn Val Gly Leu Met Pro
 275 280 285
 50 Tyr Arg His Pro Ile Val Ser Val Val Gly Arg Pro Ile Ser Val Glu
 290 295 300
 Gln Lys Asp His Pro Thr Thr Ala Asp Leu Glu Glu Val Gln Ala Arg
 55 305 310 315 320
 Tyr Ile Ala Glu Leu Lys Arg Ile Trp Glu Glu Tyr Lys Asp Ala Tyr
 325 330 335
 60 Ala Lys Ser Arg Thr Arg Glu Leu Asn Ile Ile Ala
 340 345
 <210> 83
 <211> 2758
 65

<212> ADN
<213> Rhodotorula glutinis ATCC 204091
<400> 83

5	atgcacgct cgctcctcgt tgctgctgct ggggtccaagc gaccgcgcac tgctgctgcc	60
	agtctcgctg ccgcgcgcag ggcgttccag tctgccgcgc atgtcgagaa caagctgcga	120
	ctccaggcca aggaggcgga tgaggagggc gcgcgcttgc acgactcgac catctcgaa	180
10	aaggagtacc tccagcactc gaagaacgag gtgcgtgagg ttctagctgg ccgggacggg	240
	ccggaggcgt tggtggcgaa gtcgggggccc aggaggaggt gaggaggacg agggaggcga	300
	ccgcagctca gggaggggca gtggcaggcg gtaacggtag caagggtac taacagcagt	360
15	agacgacctt ctttggcccc caaacgcgcg ctgataacct tccgcccga tcccggctct	420
	cgctcaactt cctcgacacc ctcttgatcc tccagcgac ccatctggac cgcctcgca	480
	ggctgctgcg tctgaacga gccagcactg aacaaaggcg ccggtttcac gccggaagag	540
20	ccggacacgt ttgggttgac gggactcttg ccgtgcgttc ttgctcgtt cgcttgaaga	600
	agttgtcgag ctgaccttc gcgaacaga catgaggtag actcgcttga ccagcaatgc	660
	aagcgcgctt attcgagct gcaagaacgg ccgagtgtt tgccgaaata cacctgtgcg	720
25	ttcgcgctc ccttcttgog tccctotcaa catggcttcg accgcgaaag ggtcgtctt	780
	gccttaacga cctcaacaag gttccgcgcg ggaaccacat tgcgccatcc gctgacgcg	840
	cgccctctg gcagtcctct cctccctccg cgacccaaac atcatcctct tctacgccct	900
30	ctgcctccgc cacctcaacg agctcctccc cgtcatctac actcccacg tcggcgaagc	960
	gatccaaaag tactcgacaa tctggcgcag gccggacggg ttgttctct cctacgcaca	1020
	ccggcacaag atgcgtgaga tgatgatgca ggcgaagagg cctaaggatg tcgacttgat	1080
35	catcgtcacc gactcggagg gcattctcgg tattggagac caagggtgtg gagggatctt	1140
	gatcgcgcag ggaaaggcga acctctacac cctcggagcc ggtatcgacc cctcgcgcac	1200
	cctctcggtc gtctctgacg tcggaaccga caactctgcg ctcttgaatg atccgctcta	1260
40	cctcgggttg cgacggaagc gcgtccgtgg tgccgagtag gacaagttt tcgaccggtt	1320
	ttgcgagctc gtgagggagg agtaccgcga ggcgttgctg tgagtcggtt cgttcaactg	1380

	cgagcgggtac tgctgagcct gactttgagc tcgccgtatg cagtcacttt gaggactttg	1440
	gcgtctcgaa tgcctcgaag atcctcacga cgtaccggaa caagcagtct gtctgtgcgt	1500
5	tcttcagcgc tcgatttgtc tccctccaaa agctcactga tccatccgca cagtcaacga	1560
	cgacatgcaa ggcaccgcag cagtcgtcct cgcgcgcctc ctttcggccg tcaaggtcac	1620
	gaagagcgag ctcaaggacc aacgcacgtc cgtcttttgtt ttccggcacgg ctgggtacgg	1680
10	gatcgagac gggatttcga acgctttgat gcttgaggcg gggttgtcga gtgaggaggt	1740
	tcggaagatc ttctggtgcg tgcgaggacc cgaagcgcga ggcaggcgag actgacgctt	1800
	cctgtcgatg tgcagggtgtg tcgaccgacc gggctctgctt acgacggagc actctccac	1860
15	cctacgacca ggccaagagc acttcacccg caacgcctcg gaagtctcgt cctgggagcg	1920
	cgacgcgcaa ggacgaatca gcttgctcga ggtcgtcaag caggccaaac cgacgattct	1980
20	catcggttgc tcgaccatgt cgggcgcggtt tgatgaggag gtcgtgaggg agatggcgaa	2040
	gcattgtcag cggccgatcg tcttccctct ctcaagtgcg tcgccttccc tttctgact	2100
	ccgagtaacc gctgataaag tcgatgacgt agaccgcagc aagctcgccg aggctgaccc	2160
25	ggccgacatc aacgagtgga cgaacggact ggcgcttatg gcgaccgggt cgcggttccc	2220
	gcctgtcaag acgcgcgag gcaaggagca caagatcgcc gaggcgtgag ttggcacata	2280
	tccagcattc gtgcacggag cgatggctga catccgcctt ttgcagcaac aacggcctgc	2340
30	tttaccctgg actcgggtctc gggtcagtct agctcttctc agccctcctt gctgtcgcgc	2400
	tgacgggtgc tctcgcagtg tcacgtctc tcgcgcctcg ttcttgacgg agaagatgat	2460
	caccgcgggc gttgctgctc tcgccaagat ggtacgtttg cctgctcacc agcgagccat	2520
35	ccaaacctga ccgtctcgtc cacaggctcc cgcctcgcac gaccccgacg agtcgctcct	2580
	ccctccactc agcgacctcc gtcattgtgc agtcaaggtc gccacggccg tcgcgaacgc	2640
40	cgccaaggag gaaggcgtca gccagatcaa gcgcgacgac cccttctccc aggatgaggt	2700
	tcgcgcggcg cagtgggacc ctgtctaccg cccgctcgag ctgctcgacc gtcactag	2758
	<210> 84	
	<211> 2232	
	<212> ADN	
45	<213> Rhodotorula glutinis ATCC 204091	
	<220>	
	<221> CDS	
	<222> (1)..(2232)	
	<400> 84	
50	atg cgc ggg agt gaa ggg ttg ccg tcg tcg tgt gtt cgt tca gcg tgt	48
	Met Arg Gly Ser Glu Gly Leu Pro Ser Ser Cys Val Arg Ser Ala Cys	
	1 5 10 15	
	tcg ttt gcg ggg ttc gct cgt cgt tcc ttc gct tcc tcg tcc tca etc	96
55		
60		
65		

	Ser	Phe	Ala	Gly	Phe	Ala	Arg	Arg	Ser	Phe	Ala	Ser	Ser	Ser	Ser	Leu	
				20					25					30			
5	gct	cgt	cgc	tgt	cct	cct	cct	gct	cac	cct	gct	ccc	ctc	tgc	cac	gct	144
	Ala	Arg	Arg	Cys	Pro	Pro	Pro	Ala	His	Pro	Ala	Pro	Leu	Cys	His	Ala	
				35				40					45				
10	cac	ctt	caa	ctg	aag	caa	gcg	aag	aag	ccg	ggt	ctc	tgc	ctc	gac	tct	192
	His	Leu	Gln	Leu	Lys	Gln	Ala	Lys	Lys	Pro	Gly	Leu	Ser	Leu	Asp	Ser	
				50				55					60				
15	ggc	ctc	gtt	ccc	cac	cca	gct	cac	ttc	cac	tct	cac	ttg	cac	tca	gta	240
	Gly	Leu	Val	Pro	His	Pro	Ala	His	Phe	His	Ser	His	Leu	His	Ser	Val	
							70					75				80	
20	gta	ctc	gca	ccc	caa	cag	cag	cac	acg	gtg	agc	tgc	cat	ccc	ctc	cac	288
	Val	Leu	Ala	Pro	Gln	Gln	Gln	His	Thr	Val	Ser	Ser	His	Pro	Leu	His	
					85					90					95		
25	tgc	cgc	gac	cct	ccc	atc	ccg	gcg	tgc	gac	cgc	ggc	acc	cag	tgc	tgt	336
	Ser	Arg	Asp	Pro	Pro	Ile	Pro	Ala	Cys	Asp	Arg	Gly	Thr	Gln	Ser	Cys	
				100					105					110			
30	gag	gct	gcc	gca	ccc	tct	ccc	acc	cct	ctt	tgc	aca	cag	aca	cac	gat	384
	Glu	Ala	Ala	Ala	Pro	Ser	Pro	Thr	Pro	Leu	Cys	Thr	Gln	Thr	His	Asp	
				115				120					125				
35	ccc	gct	cgc	ccc	atc	act	cgc	cgg	aac	gct	ctc	ttc	ctt	cat	ccc	ccc	432
	Pro	Ala	Arg	Pro	Ile	Thr	Arg	Arg	Asn	Ala	Leu	Phe	Leu	His	Pro	Pro	
				130				135					140				
40	aac	tgc	ctc	tcc	ctc	tct	ctc	ttt	cct	ctc	act	gag	cgc	cca	aaa	gaa	480
	Asn	Ser	Leu	Ser	Leu	Ser	Leu	Phe	Pro	Leu	Thr	Glu	Arg	Pro	Lys	Glu	
							150					155				160	
45	ctc	ctc	atc	gag	cgc	gcc	ttg	acc	cgc	ctc	cgc	tcc	atc	ccg	tcc	gac	528
	Leu	Leu	Ile	Glu	Arg	Ala	Leu	Thr	Arg	Leu	Arg	Ser	Ile	Pro	Ser	Asp	
					165					170					175		
50	ctc	gag	aaa	tat	acg	ttc	ctt	gcg	gga	ttg	agg	tgc	agg	aac	ccg	gat	576
	Leu	Glu	Lys	Tyr	Thr	Phe	Leu	Ala	Gly	Leu	Arg	Cys	Arg	Asn	Pro	Asp	
				180					185					190			
55	gtg	ttt	tat	gga	ctt	gtt	ggg	ggg	aac	atg	aag	gag	tgt	tgc	ggt	gtt	624
	Val	Phe	Tyr	Gly	Leu	Val	Gly	Gly	Asn	Met	Lys	Glu	Cys	Cys	Gly	Val	
				195				200					205				
60	ggc	gag	cgg	tac	ttg	agc	gag	cgc	ggg	cgc	agg	agg	aga	agg	aac	ggg	672
	Gly	Glu	Arg	Tyr	Leu	Ser	Glu	Arg	Gly	Arg	Arg	Arg	Arg	Arg	Asn	Gly	
				210				215					220				
65	cga	gtt	cga	gcg	gac	ttg	cgg	ttc	gcg	gcg	gct	tct	tcc	ctc	aac	gct	720
	Arg	Val	Arg	Ala	Asp	Leu	Arg	Phe	Ala	Ala	Ala	Ser	Ser	Leu	Asn	Ala	
					225			230				235				240	
70	tcc	ctc	gca	cga	aca	act	act	cac	aca	ctt	cga	gac	gac	gat	act	gac	768
	Ser	Leu	Ala	Arg	Thr	Thr	Thr	His	Thr	Leu	Arg	Asp	Asp	Asp	Thr	Asp	
					245					250					255		
75	gct	ttc	acc	tct	ccc	gct	cgt	ctc	acg	aac	cca	cag	ccg	atc	atc	tac	816
	Ala	Phe	Thr	Ser	Pro	Ala	Arg	Leu	Thr	Asn	Pro	Gln	Pro	Ile	Ile	Tyr	
				260					265					270			

ES 2 724 998 T3

5	acc ccc gtc atc ggt ctc gcg tgc caa aac tgg tct ctc atc cac cct Thr Pro Val Ile Gly Leu Ala Cys Gln Asn Trp Ser Leu Ile His Pro 275 280 285	864
10	ccc ccg ccc gaa tcc gac cca aca atc gaa gcc ctc tac ctc tcc tac Pro Pro Pro Glu Ser Asp Pro Thr Ile Glu Ala Leu Tyr Leu Ser Tyr 290 295 300	912
15	tcg gac ctc cct aac ctt ccc tcg ctc atc aaa gcc ctc aag acc cgt Ser Asp Leu Pro Asn Leu Pro Ser Leu Ile Lys Gly Leu Lys Thr Arg 305 310 315	960
20	ctc ccg cac aac cag atg caa atc tcc gtc gtc act gat ggt tcg cgc Leu Pro His Asn Gln Met Gln Ile Ser Val Val Thr Asp Gly Ser Arg 325 330 335	1008
25	gtt ctt gga ctt gga gat ctc gga gtt gga ggg atg ggt atc tcc caa Val Leu Gly Leu Gly Asp Leu Gly Val Gly Gly Met Gly Ile Ser Gln 340 345 350	1056
30	ggg aaa ctc tcg ctt tac gtc gca gca gga ggg gtc aat cct aaa gcc Gly Lys Leu Ser Leu Tyr Val Ala Ala Gly Gly Val Asn Pro Lys Ala 355 360 365	1104
35	acc ctt ccc atc gct atc gac ttc gga acc gat aac gag aag tta ctt Thr Leu Pro Ile Ala Ile Asp Phe Gly Thr Asp Asn Glu Lys Leu Leu 370 375 380	1152
40	gcg gat ccg ctg tat gtc ggt cag agg atg agg agg ttg agt gag gag Ala Asp Pro Leu Tyr Val Gly Gln Arg Met Arg Arg Leu Ser Glu Glu 385 390 395 400	1200
45	aaa tgc ctc gag ttt atg gac gtc ttc atg agg tgt atg cac gag acg Lys Cys Leu Glu Phe Met Asp Val Phe Met Arg Cys Met His Glu Thr 405 410 415	1248
50	ttc ccg aat atg gtc atc caa cac gaa gac tgg cag acc ccg ctc gcg Phe Pro Asn Met Val Ile Gln His Glu Asp Trp Gln Thr Pro Leu Ala 420 425 430	1296
55	ttc ccg ctc ttg cac aag aac cgc gac ttg tat ccg tgc ttt aat gac Phe Pro Leu Leu His Lys Asn Arg Asp Leu Tyr Pro Cys Phe Asn Asp 435 440 445	1344
60	gac atc caa gcc aca gcc gcc gtc gtc cta gcc gcc gcg atc cgc gcc Asp Ile Gln Gly Thr Gly Ala Val Val Leu Ala Gly Ala Ile Arg Ala 450 455 460	1392
65	ttc cac ctg aac gcc gtc gcg ctc aag gac caa aag att ctc ttc ttc Phe His Leu Asn Gly Val Ala Leu Lys Asp Gln Lys Ile Leu Phe Phe 465 470 475 480	1440
70	ggc gcg ggt tcg tcg gcc gtt gga gtc gcc gag acg att tgc aag tac Gly Ala Gly Ser Ser Gly Val Gly Val Ala Glu Thr Ile Cys Lys Tyr 485 490 495	1488
75	ttt gag ctg cag ggg atg agt gag cag gag gcc aag agc aag ttt tgg Phe Glu Leu Gln Gly Met Ser Glu Gln Glu Ala Lys Ser Lys Phe Trp 500 505 510	1536
80	ctt gtt gac tcg aag gcc ctc gtc gcg cac aac cga gcc gac acc ctc Leu Val Asp Ser Lys Gly Leu Val Ala His Asn Arg Gly Asp Thr Leu 515 520 525	1584

65

ES 2 724 998 T3

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

ES 2 724 998 T3

	245	250	255
5	Ala Phe Thr Ser Pro Ala Arg Leu Thr Asn Pro Gln Pro Ile Ile Tyr		
	260	265	270
10	Thr Pro Val Ile Gly Leu Ala Cys Gln Asn Trp Ser Leu Ile His Pro		
	275	280	285
15	Pro Pro Pro Glu Ser Asp Pro Thr Ile Glu Ala Leu Tyr Leu Ser Tyr		
	290	295	300
20	Ser Asp Leu Pro Asn Leu Pro Ser Leu Ile Lys Gly Leu Lys Thr Arg		
	305	310	315
25	Leu Pro His Asn Gln Met Gln Ile Ser Val Val Thr Asp Gly Ser Arg		
	325	330	335
30	Val Leu Gly Leu Gly Asp Leu Gly Val Gly Gly Met Gly Ile Ser Gln		
	340	345	350
35	Gly Lys Leu Ser Leu Tyr Val Ala Ala Gly Gly Val Asn Pro Lys Ala		
	355	360	365
40	Thr Leu Pro Ile Ala Ile Asp Phe Gly Thr Asp Asn Glu Lys Leu Leu		
	370	375	380
45	Ala Asp Pro Leu Tyr Val Gly Gln Arg Met Arg Arg Leu Ser Glu Glu		
	385	390	395
50	Lys Cys Leu Glu Phe Met Asp Val Phe Met Arg Cys Met His Glu Thr		
	405	410	415
55	Phe Pro Asn Met Val Ile Gln His Glu Asp Trp Gln Thr Pro Leu Ala		
	420	425	430
60	Phe Pro Leu Leu His Lys Asn Arg Asp Leu Tyr Pro Cys Phe Asn Asp		
	435	440	445
65	Asp Ile Gln Gly Thr Gly Ala Val Val Leu Ala Gly Ala Ile Arg Ala		
	450	455	460
70	Phe His Leu Asn Gly Val Ala Leu Lys Asp Gln Lys Ile Leu Phe Phe		
	465	470	475
75	Gly Ala Gly Ser Ser Gly Val Gly Val Ala Glu Thr Ile Cys Lys Tyr		
	485	490	495

Phe Glu Leu Gln Gly Met Ser Glu Gln Glu Ala Lys Ser Lys Phe Trp
 500 505 510
 5
 Leu Val Asp Ser Lys Gly Leu Val Ala His Asn Arg Gly Asp Thr Leu
 515 520 525
 10
 Pro Ser His Lys Lys Tyr Leu Ala Arg Ser Glu Pro Asp Ala Pro Lys
 530 535 540
 15
 Leu Arg Ser Leu Lys Glu Val Val Glu His Val Gln Pro Thr Ala Leu
 545 550 555 560
 20
 Leu Gly Leu Ser Thr Val Gly Gly Thr Phe Thr Lys Glu Ile Leu Glu
 565 570 575
 25
 Ser Met Ala Thr Tyr Asn Lys Arg Pro Ile Val Phe Ala Leu Ser Asn
 580 585 590
 30
 Pro Val Ala Gln Ala Glu Cys Thr Phe Glu Glu Ala Ile Glu Gly Thr
 595 600 605
 35
 Asp Gly Arg Val Leu Tyr Ala Ser Gly Ser Pro Phe Asp Pro Val Glu
 610 615 620
 40
 Tyr Lys Glu Lys Arg Tyr Glu Pro Gly Gln Gly Asn Asn Met Tyr Ile
 625 630 635 640
 45
 Phe Pro Gly Leu Gly Ile Gly Ala Ile Leu Ala Arg Val Ser Lys Ile
 645 650 655
 50
 Pro Glu Glu Leu Val His Ala Ser Ala Gln Gly Leu Ala Asp Ser Leu
 660 665 670
 55
 Thr Pro Glu Glu Thr Ala Arg His Leu Leu Tyr Pro Asp Ile Glu Arg
 675 680 685
 60
 Ile Arg Glu Val Ser Ile Lys Ile Ala Val Thr Val Ile Gln Ala Ala
 690 695 700
 65
 Gln Lys Leu Gly Val Asp Arg Asn Glu Glu Leu Arg Gly Lys Ser Ser
 705 710 715 720
 70
 Ala Glu Ile Glu Ala Tyr Val Arg Lys Gly Met Tyr His Pro Leu Leu
 725 730 735
 75
 Glu Ala Glu Gln Gln Ala Gln
 740
 <210> 86
 <211> 3875
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 86

ES 2 724 998 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50

gggccgaggg caagcagtgg atcgctgagc gcgcgggcaa gcaggttcag gtgcgtttcc 60
 ctcttccgcc tcgcctcctt tttttgcaga aagaacacgc gcaactgacac tgcgaacgaa 120
 cgcgcacgcg caggctcgaga agacgacagg cacgctcaac aacttcatcg tcgagccggt 180
 ctgcccgcac ccttcggacg ccgagtacta catctgcatc aactcgggtcc gcgagggcga 240
 cgtgatcctg ttcacgcacg agggcggtgt cgacgtcggc gatgtcgacg ccaaggcgct 300
 cacgctcctc gtccccgtcg gcggcgagct ccctcgcgc gacgagatcc gctcccagct 360
 cctcaagcac gtcactggcg ccgagcgcca ggaggccctc atcgactaca tcatccgcct 420
 ctactcggtc tacgtcgacc tccactttgc ctacctcgag atcaaccctc tcgtcgcggt 480
 cgagaacccc tcgactggca agaccgacat cttctacctc gacatggccg ccaagctoga 540
 ccagacggcc gagtacgtcg tcgggtccaa gtgggccatc gcgcgcgacc cgtccatcat 600
 caaccccgcc gctgcaccca tgtcgaaagg caagatttcg gccgacaagg gcccgcccat 660
 gttctggccc cctcccttcg gtgctgactt gaccaaggag gaggcctaca tcgccaagct 720
 cgacggctcg accggcgcat ccctcaagct caccgtcttg aacgccgagg gacgcatctg 780
 gacgatgggt gccggtggtg gtgcttcggt cgtctactcg gacgccatcg ccgcgcacgg 840
 attcgcgcac gagctcgcca actacggcga gtactcgggc gcgccacccc agaccagac 900
 ttacgagtac gccaagacca tccgtgcgtt tccgccttct ctcttccgct ctacagacgc 960
 ctcgctgacg ctccgtaacg ctcgagtcg acctgatgac ccgcggcacg cccaaccgcg 1020
 agggcaagct cctcttcacg ggcgcggtg tcgccaactt caccaacgtc gccgcgacct 1080
 tcaagggcat catcacggcg ctcaaggagt accagcaccg tctccaggag cacaaggtcc 1140
 gcatctttgt ccgccgcggc ggccccaact accaggaggg cctcaaggcc atgcgcctcc 1200
 tcggcgagac gctcggcgtc gagatccagg tctttggccc cgaaacccac atcacctoga 1260
 tcgtcccgtc cggcctcggc ttgatcaagt cggtcgacga cgcctcaag gttcccggcg 1320
 cccgctccgc cgcgcacgcg accggcaccc tcaccccggt tcccggctcg cccaagtgcg 1380
 gcgcgcgcca gctcccgacc ggcgcgctga cgcctcgcg ccgcgagccc caggacaaca 1440
 tcgtcagctt ctgggacaag atccacgcgc ccgactcggg ccgcccggtg tacgcgccct 1500
 ttgacgagac gacgcgctcg atcgtctacg gtctccagcc ccgcgcgacg cagggcatgc 1560
 tcgactttga ctttgcctgc ggccgcgaga cgcctcgggt cgcgcgcatg gtctaccctt 1620
 tcggcgcca ccacgtccag aagttctact ggggcaccaa ggagaccctc ctccccgtct 1680

	tcacctogat gaaggaggcc gtgcgcaagt gcccgcagcg cgacgtcgtc gtcaactttg	1740
	cgtcacogcg ctcggtctac cagtcgaccc tcgaggcgct cgagttcccc cagatcaagg	1800
5	ccatcgccct catcgccgag ggtgtcccg agcgccacgc ccgcgagatc ctccacctcg	1860
	ccaagaagaa ggaggtcatc atcatcggtc ccgcgacggt cgggcgcatc aagcccggt	1920
	gcttccgtat cggcaacacg ggcggcatga acgagaacat cctctcgtec aagctttacc	1980
10	gcgcgggttc cgtcggctac gtctccaagt ccggaggcat gtcgaaacgag ctcaacaaca	2040
	tcctttcgct cagcagcgac ggcgcgtacg agggcatcgc catcggcggt gaccgctacc	2100
	cgggcaccac cttcatcgac cactttctcc gctacgagge cgaccogaac tgcaagatgc	2160
15	tcgtcctcct cggagaggtc ggcggtgtcg aggagtaccg cgtcatcgag gccgtcaagt	2220
	cgggccagat caagaagccc atcgtcgctg gggccatcgg cactgcgcgc aagatgtttg	2280
	cgaccgaagt ccagttcggc cagcgcggtt ccattggcaa ctcggaacct gagaccgcgc	2340
20	aggccaagaa caacgccatg cgcgcgcgcg gcttcacgtt ccccccagac ttcgaggagc	2400
	tcccgaggt cctcgccgag acctaccaga agctcgtcgg cgacggcacg atccagccca	2460
	agcccgaggt tcctccccc cagatcccca tggaactaaa ctgggctcag acgctcggca	2520
25	tggtccgcaa gcccgcgcgc ttcatctoga ccattctgga cgagcgcggc caggagctcc	2580
	tctacgcgg catgcccac tcgaaggtct tcgaggagga catcggcacg ggcggcgctg	2640
	ttctcgctct ctggttcaag cgcgcgcctc ccgcctacgc gaccaagttc ctcgagatgg	2700
30	tcctcatgct cagggccgac caggttcccg ccgtctcggg cgccatgacc accgtcatca	2760
	ccaccctgct cggcaaggac ctctgtctgt cgtctgtcgc cgggtctctc accatcggcg	2820
	accgcttcgg tggcgcgctc gacggcgccg cgcaggagtt cagcgcgct ttcgaggccg	2880
35	gcctcacgcc ccgcgagttt gtgcactcga tgcgcaaggc caacaagctc atccccggtc	2940
	tcggccacaa ggtcaagtcc aaggccaacc cggacaagcg cgtcgagctc gtcaagaact	3000
	acgtcttcaa gcaacttccc tcgacaaagc tctctgagta cgtctctgct gtcgaggacg	3060
40	tcacgagcgc gaagaaggac acctcatcc tcaacgtcga cggaggtag ccttcgcttg	3120
	ctctctctcg acttatcctt gtactgacgc ttctctctcg tccagcaatc gccgtggcct	3180
	tttgcgattt gctcaaaaag tagcgctctt tcaactgcgc aaaaacagga cgaaggagaa	3240
45	ctgaccogcc ccgcttctt cccgcagctc ggcgcgctc accaccgagg aggtgcgca	3300
	gtacatgaag gtgtgtcctt gctctctat agcgttcgtg cgaccgaacg aaggagcgtg	3360
	gaacgttcag aggcgcggga aggaagagga ggggagacgg ggcgaacgga attgctgacc	3420
50	tcgcgccttc gcagatcgga accttgaacg gtctctctgt tcttggtcgc tcgatcggt	3480
	tcattgccc ccacctcgac cagaagcgcc tcaagcagcc gctctaccgt cacccgccg	3540
	agtgcgctcc atctctctt tgccatctg acctttcgcg tgacctgctc ctctcgcga	3600
55	gcgacatctt catccagccc ttcaacaccg accgcctct cgtccagcag cgccagtaaa	3660
	tggcgcacgc tcgatcgctc accgcggtgg ccgcctagac ttcgctcgca tccgcctctc	3720
	tcgttctcgc ctcgttcgac attagagttc ccttcttggt ttcgcgctgt tccgtcttct	3780
60	ttgctttctt gttctttctc tttttctcgc tctctttcgc gcagtaacct ctctcatggg	3840
	ctctcagtcg tcgtgcatca aagtttctgt tctcc	3875
	<210> 87	
65	<211> 3489	

<212> ADN

<213> Rhodosporidium toruloides

<400> 87

5	ggggccgaggg caagcagtggt atcgctgagc gcgcggggcaa gcagggttcag gtgcgtttcc	60
	ctcttccgcc tcgcctcctt tttttgcaga aagaacacgc gcactgacac tgcgaacgaa	120
	cgcgacgcgc caggtcgaga agacgacagg cacgctcaac aacttcacgc tcgagccggt	180
10	ctgcccgcac ccttcggacg ccgagtacta catctgcac aactcgggtcc gcgagggcga	240
	cgtgatcctg ttcacgcacg agggcggtgt cgacgtcggc gatgtcgacg ccaaggcgct	300
	cacgctcctc gtccccgtcg gcggcgagct cccctcgcgc gacgagatcc gctcccagct	360
15	cctcaagcac gtcactggcg ccgagcgcca ggaggccctc atcgactaca tcatccgctt	420
	ctactcggtc tacgtcgacc tccactttgc ctacctcgag atcaaccctc tcgtcgccgt	480
	cgagaacccc tcgactggca agaccgacat cttctacctc gacatggccg ccaagctcga	540
20	ccagacggcc gactacgtcg tcggtcccaa gtggggccatc gcgcgcgacc cgtccatcat	600
	caaccccgcc gctgcaccca tgtcgaacgg caagatttcg gccgacaagg gcccgcccat	660
	gttctggtccc cctcccttcg gtctgtactt gaccaaggag gaggcctaca tcgccaagct	720
25	cgacggctcg accggcgcat cctcaagct cacgctcttg aacgcgcagg gacgcatctg	780
	gacgatggtt gccggtggtg gtgcttcggt cgtctactcg gacgccatcg ccgcgcacgg	840
30	attcgcgcac gagctcgcca actacggcga gtactcgggc gcgcccaccc agaccagac	900
	ttacaggtac gccaaagcca tcctcgacct gatgaccgc gccacgcca acccgaggg	960
	caagctcctc ttcacggtcg gcggtatcgc caacttcacc aacgtcgccg cgacctcaa	1020
35	gggcatcatc acggcgctca aggagtacca gcaccgtctc caggagcaca aggtccgcat	1080
	ctttgtccgc cgcggcggcc ccaactacca ggagggcctc aaggccatgc gcctcctcg	1140
	cgagacgctc ggcgtcgaga tcagggtctt tggccccgaa acccacatca cctcgatcgt	1200
40	cccgtcggc ctcggttga tcaagtcggt cgacgacgcc ctcaagggtc ccggcgccc	1260
	ctccgcgcgc gacgcgaccg gcaccctcac ccccgttccc ggctcgcca agtcgcgcgc	1320
45	cgcccagctc ccgaccggcg cgtcgacgcc ctgcgcgccg cagccccagg acaacatcgt	1380

	cagcttctcg gacaagatcc acgcgcccga ctcgggccgc ccgtgggtacc gcccctttga	1440
	cgagacgacg cgctcgatcg tctacggtct ccagccccgc gcgatccagg gcatgctoga	1500
5	ctttgacttt gctcgcgccg gcgagacgcc ctcggtcgcc gccatggtct accccttcgg	1560
	cgggcaccac gtccagaagt tctactgggg caccaaggag accctcctcc cegtcttcac	1620
	ctcgatgaag gaggccgtcg ccaagtgcgc cgacgccgac gtcgctcgta actttgcgtc	1680
10	atcgcgctcg gtctaccagt cgaccctoga ggcgctcgag ttccccaga tcaaggccat	1740
	cgccctcatc gccgagggtg tccccgagcg ccacgcccgc gagatcctcc acctcgccaa	1800
	gaagaaggag gtcacatca tgggtccgc gacggtcggc ggcacaaagc cgggctgctt	1860
15	ccgtatcggc aacacgggcg gcatgaacga gaacatcctc tcgtccaagc tttaccgcgc	1920
	cggttcggtc ggctacgtct ccaagtccgg aggcattgtc aacgagctca acaacatcct	1980
	ttcgctcacg accgacggcg cgtacgaggg catcgccatc ggcggtgacc gctaccgggg	2040
20	caccaccttc atcgaccacc ttctccgcta cgaggccgac ccgaactgca agatgctcgt	2100
	cctcctcgga gaggtcggcg gtgtcgagga gtaccgcgtc atcgaggccg tcaagtccgg	2160
	ccagatcaag aagcccatcg tcgctggggc catcggcacc tgcgccaaaga tgtttcgac	2220
25	cgacgtccag ttcggccaag ccggttccat ggccaactcg gacctcgaga ccgccgaggc	2280
	caagaacaac gccatgcgcg ccgcccggct catcgctccc ccgacctcg aggagctccc	2340
	gcaggtcctc gccgagacct accagaagct cgtcgccgac ggcaacgatc agcccaagcc	2400
30	cgaggttctt cccctcaga tcccgatgga ctacaactgg gctcagacgc tcggcatggt	2460
	ccgcaagccc gccgccttca tctcgaccat ctcgacgag cgcgggccagg agctcctcta	2520
35	cgccggcatg cccatctcca aggtcttcga ggaggacatc ggcacggcg gcgtcgtctc	2580
	gctcctctgg ttcaagcgcc gcctccccgc ctacgcgacc aagttcctcg agatggtcct	2640
	catgctcacg gccgaccacg gtcccgcgt ctcgggcgcc atgaccaccg tcatcaccac	2700
40	ccgtgccggc aaggacctcg tctcgctcgt cgtcgccggt ctctcacca tcggcgaccg	2760
	cttcgggtggc gcgctcgacg gcgcgcgcga ggagttcacg cgcgctttcg aggcggcct	2820
	cacgccccgc gagtttgctg actcgatgcg caaggccaac aagctcatcc ccggtatcgg	2880
45	ccacaaggtc aagtccaagg ccaaccggga caagcgctc gagctcgtca agaactacgt	2940
	cttcaagcac ttcccgctga caaagctcct cgagtacgct ctgctgtcg aggacgtcac	3000
	gagcgcaag aaggacaccc tcatctcaa ccaatcgccg tggcctctct cgatttgctc	3060
50	aaaaactcgg gcgcgttcac caccgaggag gctgccgagt acatgaagat cggaaccttg	3120
	aacggtctct tcgttcttgg tcgctcgatc ggcttcattg cccaccacct cgaccagaag	3180
	cgctcgaag agcgcctcta ccgtcacccg gcgcagaca tcttcatoca gcccttcaac	3240
55	accgaccgca tcctcgcca gcagcgccag taaatggcg acgctcgatg cgtcacccgc	3300
	gtggccgcct agacttcgct cgcacccgcc tctctcgttc tcgctcgtt cgacattaga	3360
	gttcccttct tgttttcggc tcgttcgctc tttcttgctt tcttgttctt tctcttttcc	3420
60	tcgctctctt tcggcagta cctcctctca tgggctctca gtcgtcgtgc atcaaagttt	3480
	cgtttctcc	3489
	<210> 88	
65	<211> 1148	

<212> PRT

<213> Rhodosporidium toruloides

<400> 88

Met Ser Ala Lys Pro Ile Arg Glu Tyr Asp Ala Lys Leu Leu Leu Ala
1 5 10 15

Tyr His Leu Ala Arg Ala Pro Thr Ala Gly Ser Lys Ala Val Ala Arg
20 25 30

Asp Gly Phe Gln Ser Pro Glu Val Lys Val Ala Gln Val Ser Trp Asp
35 40 45

Pro Glu Thr Asn Gln Val Thr Pro Asp Ala Ala Leu Pro His Trp Val
50 55 60

Phe Thr Glu Lys Leu Val Val Lys Pro Asp Gln Leu Ile Lys Arg Arg
65 70 75 80

Gly Lys Ala Gly Leu Leu Ala Leu Asn Lys Thr Trp Ala Glu Gly Lys
85 90 95

Gln Trp Ile Ala Glu Arg Ala Gly Lys Gln Val Gln Val Glu Lys Thr
100 105 110

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

Ser Asp Ala Glu Tyr Tyr Ile Cys Ile Asn Ser Val Arg Glu Gly Asp
130 135 140

Val Ile Leu Phe Thr His Glu Gly Gly Val Asp Val Gly Asp Val Asp
145 150 155 160

Ala Lys Ala Leu Thr Leu Leu Val Pro Val Gly Gly Glu Leu Pro Ser
165 170 175

Arg Asp Glu Ile Arg Ser Gln Leu Leu Lys His Val Thr Gly Ala Glu

ES 2 724 998 T3

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

Tyr Gln Glu Gly Leu Lys Ala Met Arg Leu Leu Gly Glu Thr Leu Gly
 435 440 445
 5
 Val Glu Ile Gln Val Phe Gly Pro Glu Thr His Ile Thr Ser Ile Val
 450 455 460
 10
 Pro Leu Gly Leu Gly Leu Ile Lys Ser Val Asp Asp Ala Leu Lys Val
 465 470 475 480
 Pro Gly Ala Arg Ala Ala Ala Asp Ala Thr Gly Thr Leu Thr Pro Val
 485 490 495
 15
 Pro Gly Ser Pro Lys Ser Arg Ala Ala Gln Leu Pro Thr Gly Ala Ser
 500 505 510
 20
 Thr Pro Ser Arg Gln Gln Pro Gln Asp Asn Ile Val Ser Phe Ser Asp
 515 520 525
 Lys Val His Ala Pro Asp Ser Gly Arg Pro Trp Tyr Arg Pro Phe Asp
 530 535 540
 25
 Glu Thr Thr Arg Ser Ile Val Tyr Gly Leu Gln Pro Arg Ala Ile Gln
 545 550 555 560
 30
 Gly Met Leu Asp Phe Asp Phe Ala Cys Gly Arg Glu Thr Pro Ser Val
 565 570 575
 Ala Ala Met Val Tyr Pro Phe Gly Gly His His Val Gln Lys Phe Tyr
 580 585 590
 35
 Trp Gly Thr Lys Glu Thr Leu Leu Pro Val Phe Thr Ser Met Lys Glu
 595 600 605
 40
 Ala Val Ala Lys Cys Pro Asp Ala Asp Val Val Val Asn Phe Ala Ser
 610 615 620
 Ser Arg Ser Val Tyr Gln Ser Thr Leu Glu Ala Leu Glu Phe Pro Gln
 625 630 635 640
 45
 Ile Lys Ala Ile Ala Leu Ile Ala Glu Gly Val Pro Glu Arg His Ala
 645 650 655
 50
 Arg Glu Ile Leu His Leu Ala Lys Lys Lys Glu Val Ile Ile Ile Gly
 660 665 670
 55
 Pro Ala Thr Val Gly Gly Ile Lys Pro Gly Cys Phe Arg Ile Gly Asn
 675 680 685
 60
 65

ES 2 724 998 T3

Thr Gly Gly Met Asn Glu Asn Ile Leu Ser Ser Lys Leu Tyr Arg Ala
 690 695 700
 5
 Gly Ser Val Gly Tyr Val Ser Lys Ser Gly Gly Met Ser Asn Glu Leu
 705 710 715 720
 10
 Asn Asn Ile Leu Ser Leu Thr Thr Asp Gly Ala Tyr Glu Gly Ile Ala
 725 730 735
 Ile Gly Gly Asp Arg Tyr Pro Gly Thr Thr Phe Ile Asp His Leu Leu
 740 745 750
 15
 Arg Tyr Glu Ala Asp Pro Asn Cys Lys Met Leu Val Leu Leu Gly Glu
 755 760 765
 20
 Val Gly Gly Val Glu Glu Tyr Arg Val Ile Glu Ala Val Lys Ser Gly
 770 775 780
 Gln Ile Lys Lys Pro Ile Val Ala Trp Ala Ile Gly Thr Cys Ala Lys
 785 790 795 800
 25
 Met Phe Ala Thr Asp Val Gln Phe Gly His Ala Gly Ser Met Ala Asn
 805 810 815
 30
 Ser Asp Leu Glu Thr Ala Glu Ala Lys Asn Asn Ala Met Arg Ala Ala
 820 825 830
 Gly Phe Ile Val Pro Pro Thr Phe Glu Glu Leu Pro Gln Val Leu Ala
 835 840 845
 35
 Glu Thr Tyr Gln Lys Leu Val Gly Asp Gly Thr Ile Gln Pro Lys Pro
 850 855 860
 40
 Glu Val Pro Pro Pro Gln Ile Pro Met Asp Tyr Asn Trp Ala Gln Thr
 865 870 875 880
 Leu Gly Met Val Arg Lys Pro Ala Ala Phe Ile Ser Thr Ile Ser Asp
 885 890 895
 45
 Glu Arg Gly Gln Glu Leu Leu Tyr Ala Gly Met Pro Ile Ser Lys Val
 900 905 910
 50
 Phe Glu Glu Asp Ile Gly Ile Gly Gly Val Val Ser Leu Leu Trp Phe
 915 920 925
 Lys Arg Arg Leu Pro Ala Tyr Ala Thr Lys Phe Leu Glu Met Val Leu
 930 935 940
 55
 60
 65

5 Met Leu Thr Ala Asp His Gly Pro Ala Val Ser Gly Ala Met Thr Thr
945 950 955 960

Val Ile Thr Thr Arg Ala Gly Lys Asp Leu Val Ser Ser Leu Val Ala
965 970 975

10 Gly Leu Leu Thr Ile Gly Asp Arg Phe Gly Gly Ala Leu Asp Gly Ala
980 985 990

15 Ala Gln Glu Phe Thr Arg Ala Phe Glu Ala Gly Leu Thr Pro Arg Glu
995 1000 1005

Phe Val Asp Ser Met Arg Lys Ala Asn Lys Leu Ile Pro Gly Ile
1010 1015 1020

20 Gly His Lys Val Lys Ser Lys Ala Asn Pro Asp Lys Arg Val Glu
1025 1030 1035

25 Leu Val Lys Asn Tyr Val Phe Lys His Phe Pro Ser Ala Lys Leu
1040 1045 1050

30 Leu Glu Tyr Ala Leu Ala Val Glu Asp Val Thr Ser Ala Lys Lys
1055 1060 1065

Asp Thr Leu Ile Leu Asn Gln Ser Arg Phe Leu Ser Ser Gly Ala
1070 1075 1080

35 Phe Thr Ala Glu Glu Ala Ala Glu Tyr Met Lys Ile Gly Thr Leu
1085 1090 1095

40 Asn Gly Leu Phe Val Leu Gly Arg Ser Ile Gly Phe Ile Ala His
1100 1105 1110

His Leu Asp Gln Lys Arg Leu Lys Gln Pro Leu Tyr Arg His Pro
1115 1120 1125

45 Ala Asp Asp Ile Phe Ile Gln Pro Phe Asn Thr Asp Arg Ile Leu
1130 1135 1140

50 Val Gln Gln Arg Gln
1145
<210> 89
<211> 7222
<212> ADN
<213> Rhodosporidium toruloides
55 <400> 89

60

65

	accacactct cccgcttgcg ggcctctctt ctcgcttgge gtcctgcta ccgctactct	60
	agactctcct agtctccctg cacaaccatc cctatccctt ccgcctctct cgcacacccc	120
5	ccacagcttc gttccccaac ttcacttccg atgcggtgcg tcgcctccct ttgcctggc	180
	gggcccgcgc ctgcttccga ggacaactac tgattgtggg atcatgogac gacaggttct	240
	ctggcgaggc gaaggcggtc aacggatcgc actcggtcga cgaggcgccg aagaacccca	300
10	agtacgacca tgggcggtc gtaaagtacc tcggtgcgtt tcatggggc ctttgagcac	360
	tctggagact ttctggaccg gcggacgggc agttgggctg acctgcatt tctctcctc	420
	ctogacttog ctacactoga caggcggcaa ctcgctcgag tctgogcccc attccaaggt	480
15	cgctgactgg gtcaggagc gtggtggaca caccgtcctc acaaagggtc gctcactcgg	540
	acccctcttg caccgcccct cccgaccgtc cccaactcgc tcgtcctgac ccaactcggc	600
20	ctccgcacag atcctgatcg ccaacaatgg tatcgtcgca gtcaaggaga tccgctcggc	660
	gcgcaagtgg gcgtacgaga cgttcggaag cgagcgcgcg atcgagttta ccgtcatggc	720
	gaccccgagg gacctcaagg tcaatgcaga ctacatccgc atggccgacc agtacgtoga	780
25	ggtaccgggt ggaaccaaca acaacaacta cgcgaacgtc gatgtcatcg ttgatgttgc	840
	cgagcgcgct ggcgtccacg ccgtctgggc aggatggttc gtcccgtttc cttcccgctc	900
	ccttcagcgg aagtgccctg gctaacttgc cttcgacgga cacaggggcc acgcttcga	960
30	gaacccccgc cttcccgagt cgtcgcgcgc ctogaagcac aagatcgtct tcatcggtcc	1020
	tcccggctcc gccatgcgct cgtcggcgca caagatctcg tcgaccatcg ttgcgcagca	1080
	cgcacagggt ccgtgcatgg actggtccgg acaggcgctc gaccagggtc ccagtcgcc	1140
35	cgagggtac gttaccgtcg ccgacgacgt ctaccagcag gcctgtgtgc acgacgccga	1200
	cgagggtctc gcccgcgct cgaggatcgg ataccccgtc atgatcaagg cgtccgaggg	1260
	aggaggagga aagggtattc gcaaggtcga gaaggagcag gacttcaagc aggccttcca	1320
40	ggcgtcctc accgaggttc ccggtgcgtc gattcgcttt cagtttgaag catccttagc	1380
	tgactcgccg cttccatagg ctgcgccgtc ttcacatga agctcgtcgg cgctgctcgc	1440
45	cacctcgagg ttcaggttct cgcgaccag tacggcaacg ccatctcgt cttcggccgt	1500
	gactgctcgg tocagcgtcg ccaccagaag atcatcgaag aggcgcccgt caccatcgcc	1560
	aagcccgaca cgttcgagca gatggaaaag tcggccgtcc gccttgccaa gctcgtcggc	1620
50	tacgtctcgg cgggaaccgt cgagttcctc tactcggctg ccgacgacaa gttcgccttc	1680
	ctcgagctca acccgctct ccagggtgag cacccgacga ccgagatggt ctcggcgctc	1740
	aaccttcccg ccgtcagct ccaggtcgcc atgggtgtcc ctctccaccg catccgcgac	1800
55	atccgcaccc tctacggcaa ggcacccaac ggcagcagcg agatcgactt cgacttogag	1860

60

65

	aatcccgaat cggccaagac gcagcgcaag cccctcgccga agggtcacgt cgtcgccgtc	1920
5	cgtatcacgg ctgagaaccc tgatgccggc ttcaagccgt cgtatgggtac cctccaggag	1980
	ctcaacttcc gttcgagcac gaacgtctgg ggtaacttct cgtcgggcag cgcgggtgg	2040
	ctccacgagt ttgccgactc gcagtttggc cactcttttg cgtacggctc ggaccgttcc	2100
10	gagtcgcgca agaacatggg cgtcgcgctc aaggagctct cgattcgcg tgacttccgc	2160
	acgaccgtcg agtacctcat caagcttctc gagaccgatg cgttcgagca gaacacgatt	2220
	acgaccgcct ggctcgacag cctcatctcg gctcgccga ccgccgagag gcccgacacg	2280
15	actctcgcca tcatctgggg cgcgctcagc aaggctcacc tcgcttcgca ggccaacatc	2340
	gccgagtaca agcgcatcct cgagaagggg cagagcccg ccaaggagct cctggccacc	2400
	gtcgtccctc tcgagttcgt cctcgaggac gtcaagtacc gcgcgaccgc ctcgcgctcg	2460
20	tcgctctoga gctggtccat ctacgtcaac ggctcgaaag tctcggtcgg catccgccct	2520
	ctcgccgatg gcggtctcct catctctctc gacggccgct catacactg ctacgccaaag	2580
	gaggaggctg gcgcgcttcg cctctcaatc gactcgagga ctgtcctcat cgtccaggag	2640
25	aacgacccca cccagcttcg ctgccatcg ccggcgaagc tcgtccgcta cttcatcgag	2700
	tcgggcgagc acatctcaaa gggcgaggcg tacgccgaga tcgaggtcat gaaggtgct	2760
	cgccccggca ttccctgctc gtgttctgtc cgtcgacgt cgatgcctcg cagatgatca	2820
30	tgccctcat cgtctgtgag gacggtatcg cgcagttcat caagcagccg ggagcgacgc	2880
	tcgaggctgg cgacatcctg ggtatcttgt cgtcgacga ccgagccgc gtccaccaag	2940
	ccaagccgtt cgacggccag cttccggccc ttggcttgcc ttccatcctc ggcaacaagc	3000
35	cgcaccagcg cttcgccctac ctcaaggagc tctctcgaa catctcatg ggctacgaca	3060
	accaggctgt gatgcagtcg agcatcaagg agctcatctc ggtccttcgc aaccocgagc	3120
	tcccgtagcg cgaggccaac gccgtcctct caacgctttc ggcccgctac cccgccaaagc	3180
40	tcgagcagac cctccgccag tacattgacc aagctcagca gtctggcgcc gagtcccggt	3240
	ccgccaaagt ccgcaaggcg atcgacaaga cctcgagca gctccgccct gccgaggcgc	3300
	agaccgtccg caacttctc gtgcggttcg acgacatcgt ctaccgctac cgtcggggcc	3360
45	tcaagcacca cgagtggctg acgctcgccg gcatctttgc cgcgtacgcc gagacggaga	3420
	agccgttcag cggcaaggac ggcgacgtcg tctcgagct ccgcgacgcc caccgcgact	3480
	cgtctgactc ggtgtcaag atgtctctc cgcactacaa ggccgcctcg aagaactcgc	3540
50	tcgtctcgc cctcctcgac attgtcaagg actcggactc ggtcccgctc atcgagcagg	3600
	tcgtcagccc cgcgctcaag gaccttgccg acctcgatc gaaggccacg accaaggctg	3660
	cctcaaggc tcgcgaggtg ctcattccaca tccagctccc ctcgctcgac gacgcctcg	3720
55	gccagcttga gcagatctc aaggcctcgg tcacgcccac cgtctacggc gacgtggcc	3780
60		
65		

ES 2 724 998 T3

	acgatcgac tctcgcggg gaggtcctca aggatgtcat cgactcgcg ttcacgtct	3840
	tgcagttct cccgagcttc ttccagcacc aggaccactg ggtctcgtt gccgcgtcg	3900
5	acacctacgt ccgcgcgct taccgctcgt acaatctcct caacatcgag cacattgagg	3960
	ccgatgctgc cgaggacgag ccgcgcaccg tcgcctggtc gttccgcatg cgttaaggctg	4020
	cgtccgaatc cgagccgct acgcccacga ccggcctcac ctgcagcgc accgcctcgt	4080
10	actcggactt gaccttctc ctcaacaacg cccagtcga gccgatcgc tacggcgga	4140
	tgttctcgtt ccgctcgtc gaccgcttc gccaggagct cggcaccgtc ctccgccact	4200
	tccccgactc gaacaagggc aagctccagc agcagcctgc cgttcgtcg agccaggagc	4260
15	agtggaaagt catcaatgtt gcgctcagc tccccgcag ccgcgcagtc gacgaggagc	4320
	ctctccgtgc cgacttcgcc gctcacgtca acgcgatgag cgcgcgagtc gacgctcgcg	4380
	gcctgcgcg cctcaccctc ctcatctgcc gcgagggcca gtaccgctc tactacaccg	4440
20	tccgcaagca ggacggcacc tggaaggagc tcgagacgat ccgcgcacac gagcctgcc	4500
	tgccttcca gctcagctc ggccgctct ccaacttcca cctcagacg tgcccagtcg	4560
	agaaccgcca ggttcacgtc tactacgcga ccgccaggc caactcgtc gactgcgct	4620
25	tcttcgtcgc tgcgctcgtc cgcctggcc gcctccgcg caacatgaag acggccgact	4680
	accttgtgtc cgaggccgac cgcctcgtca ccgacgtcct cgactcgtc gaggtcgcca	4740
30	gctcgcagc ccgcgcgcgc gacggcaacc acatctcgtt gcgtactgc agccttgccc	4800
	tccctcgttc tcgcgctgac tcgctctcgt ttgcagctc aacttctgt actcgtcgc	4860
	tctcgaactt gacgaggttc aggtgcctt ccgcggcttc atcgaccgc accgcaagcg	4920
35	cttctggcgt ctccgctca ccggcgccga gatccgcac gtccctcagg acgcgcagg	4980
	caacatccag cccatccgc ccatcattga gaacgtctc ggttcgtcg tcaagtacga	5040
	ggcgtaccgc gaggtcacga ctgacaagg ccaggtcac ctcaagtoga tcggtccga	5100
40	gggcgcctg cacttcagc cggctcaact cccctaccg accaaggagt ggcttcagcc	5160
	gaagcgctac aaggcccacg tcgtcggcac gacgtacgtc tacgacttc ccgaccttt	5220
	ccgccaggct atccgcaagc agtggaggc ccgcggcaag accgcgccg ccgagctcct	5280
45	cgtcgccaag gagctcgtc tcgacgaatt cggcaagccc caggaggctg cccgccgc	5340
	tggcactaac aacatcgga tggctcgctg gatctacaca atcttcacgc ccgagtacc	5400
	gaccggccgc cgcgtcgtc tcctcgccaa cgacatcac ttcaagattg gttcgttcg	5460
50	ccccgaggag gaccgctact tcttcgcgt cagcagctc gcgcgccagc ttggttgcc	5520
	gcgcgtctac ctctcgcca actcgggcgc tcgcctcgtt attgccagg agctcgtga	5580
55	cttgttcagc gtcgctggg tcgacagctc gcggccggag aagggttca agtacctca	5640
60		
65		

	cctcaccgcc gagaagctcg gcgagctcaa gaacaagggc gagaagagcg tcattacgaa	5700
5	gcgcatcgag gacgagggcg agacgcgcta ccagatcacc gacatcatcg gcttgcagga	5760
	gggtctcggg gtogagtgc tcaagggctc tggcctcatc gccggtgaga cctcgcgcg	5820
	gtacgacgac atcttcacga tcacgctcgt caccgccgcg tcggtcggta tcggtgcgta	5880
10	cctcgtcgcg ctgggccagc gtgcggtcca ggtcgagggc cagccgatca tcctcaccgg	5940
	tgcggggcgcg ctcaacaagg tcctcggtcg cgaggtctac tcgtccaact tgcagctcgg	6000
	cggcacgcaa atcatgtaca agaacgggtgt ctgcacttg acggccgcca acgacctoga	6060
15	gggtgtcctc agcatcgctc agtggtcgcg ctttgtcccc gagcaccgcg gcgcgcctct	6120
	cccgatcatg ccttcgcccg tcgaccggtg ggaccgctcg atcgactaca cggccatcaa	6180
	gggcgcgtac gacccgcgct ggttctcgcg cggcaagacg gacgagggcg acggccgctg	6240
20	gctctctggc ttcttcgaca agggctcgtt ccaggagacg ctctcgggct gggcgcagac	6300
	cgtcgtcgtc ggcgcgctc gcctcggcg catcccatg ggcgctatcg cagtogaac	6360
	ccgcaccatc gagcgcgatg ttcccgccga ccccgccaac cctctctcga acgagcagaa	6420
25	gatcatggag gccggtcagg tgcgcaacaa cctccacttt gtccgaaacc tccactcgtg	6480
	tccaaactcc tcgtcgcaa cttctcctcg acctttctcg tctcctcga cagctgctga	6540
	cctgcttcca ccttgtacca tgcgcaggtc tggtaaccga acagctcgtt caagacggga	6600
30	caggcgatct tcgacttcaa ccgcgagggg ctcgcgctca tcatcttcgc caactggcg	6660
	ggcttctcgg gccggccagca ggacatgttt gacgaggtcc tcaagcgcg ctcgctcatc	6720
35	gtcgacggtc tctcgcgta caagcagccc gtctctcgtc acatcgtccc gaacggcgaa	6780
	cttcgcggcg gtgcttgggt cgtcctcgac ccgtcgatca acgccgagg catgatggag	6840
	atgtacgtcg atgagactgc tcgcgcgggt gtctcagagc ccgagggcat cgtcgagatc	6900
40	aagctccgca aggacaagct cctcgccctc atggaccgcc tcgaccgac ctaccacgcc	6960
	ctccgcgtca agtcgaccga cgtctcgtc tcgccgcgcg acgctgcgca ggccaagacc	7020
	gagctcgccg cgcgcgagaa gcagctcatg ccgatctacc agcaggtcgc gctccagttc	7080
45	gccgactcgc acgacaaggc cggccgcata ctcagcaagg gctgcgcgcg cgaggcgctc	7140
	gagtggtcga acgctcgtcg ctacttctac gcccgccctc gccgcggtgt cgcgaggag	7200
	gccgcgctca agcgccctggg cg	7222
50	<210> 90	
	<211> 6702	
	<212> ADN	
	<213> Rhodosporidium toruloides	
	<220>	
55	<221> CDS	
	<222> (1)..(6702)	
	<400> 90	
60		
65		

5	atg cca ttc tct ggc gag gcg aag gcg gtc aac gga tcg cac tcg gtc Met Pro Phe Ser Gly Glu Ala Lys Ala Val Asn Gly Ser His Ser Val 1 5 10 15	48
10	gac gag gcg ccg aag aac ccc aag tac gac cat ggg cgg gtc gta aag Asp Glu Ala Pro Lys Asn Pro Lys Tyr Asp His Gly Arg Val Val Lys 20 25 30	96
15	tac ctc ggc ggc aac tcg ctc gaa tct gcg ccc cct tcc aag gtc gcc Tyr Leu Gly Gly Asn Ser Leu Glu Ser Ala Pro Pro Ser Lys Val Ala 35 40 45	144
20	gac tgg gtc agg gag cgt ggt gga cac acc gtc atc aca aag atc ctc Asp Trp Val Arg Glu Arg Gly Gly His Thr Val Ile Thr Lys Ile Leu 50 55 60	192
25	atc gcc aac aat ggt atc gcc gca gtc aag gag atc cgc tcg gtg cgc Ile Ala Asn Asn Gly Ile Ala Ala Val Lys Glu Ile Arg Ser Val Arg 65 70 75 80	240
30	aag tgg gcg tac gag acg ttc gga agc gag cgc gcg atc gag ttt acc Lys Trp Ala Tyr Glu Thr Phe Gly Ser Glu Arg Ala Ile Glu Phe Thr 85 90 95	288
35	gtc atg gcg acc ccg gag gac ctc aag gtc aac gca gac tac atc cgc Val Met Ala Thr Pro Glu Asp Leu Lys Val Asn Ala Asp Tyr Ile Arg 100 105 110	336
40	atg gcc gat cag tac gtc gag gtt ccc ggt gga acc aac aac aac aac Met Ala Asp Gln Tyr Val Glu Val Pro Gly Gly Thr Asn Asn Asn Asn 115 120 125	384
45	tac gcc aac gtc gat gtc atc gtc gat gtt gcc gag cgc gca ggc gtc Tyr Ala Asn Val Asp Val Ile Val Asp Val Ala Glu Arg Ala Gly Val 130 135 140	432
50	cac gcc gtc tgg gca gga tgg ggc cac gcc tcc gag aac ccc cgc ctt His Ala Val Trp Ala Gly Trp Gly His Ala Ser Glu Asn Pro Arg Leu 145 150 155 160	480
55	ccc gag tcg ctc gcc gcc tcg aag cac aag atc gtc ttc atc ggt cct Pro Glu Ser Leu Ala Ala Ser Lys His Lys Ile Val Phe Ile Gly Pro 165 170 175	528
60	ccc ggc tcc gcc atg cgc tcg ctc gga gac aag atc tcg tcg acc atc Pro Gly Ser Ala Met Arg Ser Leu Gly Asp Lys Ile Ser Ser Thr Ile 180 185 190	576
65	gtc gcg cag cac gcc cag gtt ccg tgc atg gac tgg tcc ggc cag ggc Val Ala Gln His Ala Gln Val Pro Cys Met Asp Trp Ser Gly Gln Gly 195 200 205	624
70	gtc gac caa gtc acc cag tcg ccc gag ggc tac gtt act gtc gcc gac Val Asp Gln Val Thr Gln Ser Pro Glu Gly Tyr Val Thr Val Ala Asp 210 215 220	672
75	gac gtc tac cag cag gcc tgt gtg cac gac gcc gac gag ggt ctc gcc Asp Val Tyr Gln Gln Ala Cys Val His Asp Ala Asp Glu Gly Leu Ala 225 230 235 240	720

	cgc gcg tcg agg atc gga tac ccc gtc atg atc aag gcg tcc gag gga Arg Ala Ser Arg Ile Gly Tyr Pro Val Met Ile Lys Ala Ser Glu Gly	768
5	gga gga gga aag ggt att cgc aag gtc gag aag gag cag gac ttt aag Gly Gly Gly Lys Gly Ile Arg Lys Val Glu Lys Glu Gln Asp Phe Lys	816
10	cag gcc ttc cag gct gtc ctc acc gag gtt ccc gcc tcg ccc gtc ttt Gln Ala Phe Gln Ala Val Leu Thr Glu Val Pro Gly Ser Pro Val Phe	864
15	atc atg aag ctc gcc gcc gca gct cgc cac ctc gag gtc cag gtt ctc Ile Met Lys Leu Ala Gly Ala Ala Arg His Leu Glu Val Gln Val Leu	912
20	gcc gac cag tac gcc aac gcc atc tcg ctc ttc gcc cgt gac tgc tcg Ala Asp Gln Tyr Gly Asn Ala Ile Ser Leu Phe Gly Arg Asp Cys Ser	960
25	gtt cag cgt cgc cac cag aag atc atc gaa gag gcg ccc gtc acc atc Val Gln Arg Arg His Gln Lys Ile Ile Glu Glu Ala Pro Val Thr Ile	1008
30	gcc aag ccc gac acg ttc gag cag atg gaa aag tcg gcc gtc cgc ctt Ala Lys Pro Asp Thr Phe Glu Gln Met Glu Lys Ser Ala Val Arg Leu	1056
35	gcc aag ctc gtc gcc tac gtc tcg gcg ggt acc gtc gag ttc ctc tac Ala Lys Leu Val Gly Tyr Val Ser Ala Gly Thr Val Glu Phe Leu Tyr	1104
40	tcg gct gcc gac gac aag ttt gcc ttc ctc gag ctc aac ccg cgt ctc Ser Ala Ala Asp Asp Lys Phe Ala Phe Leu Glu Leu Asn Pro Arg Leu	1152
45	cag gtc gag cac ccg acc acc gag atg gtt tcg gcc gtc aac ctt ccc Gln Val Glu His Pro Thr Thr Glu Met Val Ser Gly Val Asn Leu Pro	1200
50	gcc gcc cag ctc cag gtc gct atg ggt gtt ccc ctc cat cgc atc cgc Ala Ala Gln Leu Gln Val Ala Met Gly Val Pro Leu His Arg Ile Arg	1248
55	gac atc cgc acg ctc tac gcc aag gca ccc aac gcc agc agc gag atc Asp Ile Arg Thr Leu Tyr Gly Lys Ala Pro Asn Gly Ser Ser Glu Ile	1296
60	gat ttc gac ttc gag aac ccc gag tcg gcc aag acg cag cgc aag ccc Asp Phe Asp Phe Glu Asn Pro Glu Ser Ala Lys Thr Gln Arg Lys Pro	1344
65	tcg ccg aag ggt cac gtc gtt gcc gta cgt atc acg gct gag aac cct Ser Pro Lys Gly His Val Val Ala Val Arg Ile Thr Ala Glu Asn Pro	1392
	gac gcc gcc ttc aag ccg tcc atg ggt act ctc caa gag ctc aac ttc Asp Ala Gly Phe Lys Pro Ser Met Gly Thr Leu Gln Glu Leu Asn Phe	1440
	cgc tcg agc acg aac gtc tgg ggt tac ttc tcc gtc gcc agc gcc ggt Arg Ser Ser Thr Asn Val Trp Gly Tyr Phe Ser Val Gly Ser Ala Gly	1488

		gga ctg cac gag ttt gcc gac tcg cag ttc ggc cac atc ttt gcg tac	1536
		Gly Leu His Glu Phe Ala Asp Ser Gln Phe Gly His Ile Phe Ala Tyr	
		500 505 510	
5		ggc tcg gac cgt tcc gag tcg cgc aag aac atg gtc gtc gcg ctc aag	1584
		Gly Ser Asp Arg Ser Glu Ser Arg Lys Asn Met Val Val Ala Leu Lys	
		515 520 525	
10		gag ctc tcg att cgc ggt gac ttc cgc acg acc gtc gag tac ctc atc	1632
		Glu Leu Ser Ile Arg Gly Asp Phe Arg Thr Thr Val Glu Tyr Leu Ile	
		530 535 540	
		aag ctt ctc gag acg gac gcg ttc gag cag aac acg atc acg acc gcg	1680
		Lys Leu Leu Glu Thr Asp Ala Phe Glu Gln Asn Thr Ile Thr Thr Ala	
		545 550 555 560	
15		tgg ctc gac agc ctc atc tcg gct cgc ctg acc gcc gag agg ccc gac	1728
		Trp Leu Asp Ser Leu Ile Ser Ala Arg Leu Thr Ala Glu Arg Pro Asp	
		565 570 575	
20		acg act ctc gcc atc atc tgc ggc gcc gtt acc aag gcc cac ctc gct	1776
		Thr Thr Leu Ala Ile Ile Cys Gly Ala Val Thr Lys Ala His Leu Ala	
		580 585 590	
		tcc gag gcc aac atc gcc gag tac aag cgc atc ctc gag aag ggt cag	1824
		Ser Glu Ala Asn Ile Ala Glu Tyr Lys Arg Ile Leu Glu Lys Gly Gln	
		595 600 605	
25		agc ccc gcc aag gag ctc ctc gcc acc gtc gtc ccg ctc gag ttc gtc	1872
		Ser Pro Ala Lys Glu Leu Leu Ala Thr Val Val Pro Leu Glu Phe Val	
		610 615 620	
30		ctc gag gac gtc aag tac cgc gcg acc gcc tcg cgc tcg tcg cct tcg	1920
		Leu Glu Asp Val Lys Tyr Arg Ala Thr Ala Ser Arg Ser Ser Pro Ser	
		625 630 635 640	
		agc tgg tcc atc tac gtc aac ggc tcg aac gtc tcc gtc ggc atc cgc	1968
		Ser Trp Ser Ile Tyr Val Asn Gly Ser Asn Val Ser Val Gly Ile Arg	
		645 650 655	
35		cct ctc gcc gac ggc ggt ctc ctc atc ctc ctt gac ggc cgc tcg tac	2016
		Pro Leu Ala Asp Gly Gly Leu Leu Ile Leu Leu Asp Gly Arg Ser Tyr	
		660 665 670	
40		acc tgc tac gcc aag gag gag gtc ggc gcg ctc cgc ctc tcg atc gac	2064
		Thr Cys Tyr Ala Lys Glu Glu Val Gly Ala Leu Arg Leu Ser Ile Asp	
		675 680 685	
		tcg agg acc gtc ctc att gct cag gag aac gac ccc acc cag ctt cgc	2112
		Ser Arg Thr Val Leu Ile Ala Gln Glu Asn Asp Pro Thr Gln Leu Arg	
		690 695 700	
		tcg cct tca ccc gcc aag ctc gtc cgc tac ttc atc gag tcc gcc gag	2160
		Ser Pro Ser Pro Gly Lys Leu Val Arg Tyr Phe Ile Glu Ser Gly Glu	
		705 710 715 720	
50		cac atc tcg aag gcc gag gcg tac gct gag atc gag gtc atg aag atg	2208
		His Ile Ser Lys Gly Glu Ala Tyr Ala Glu Ile Glu Val Met Lys Met	
		725 730 735	
55		atc atg ccc ctc atc gct gcc gag gac ggt atc gcg caa ttc atc aag	2256
		Ile Met Pro Leu Ile Ala Ala Glu Asp Gly Ile Ala Gln Phe Ile Lys	

ES 2 724 998 T3

	740	745	750	
5	cag ccg gga gcg acg ctc gag gcc ggc gac atc ctc ggt atc ttg tcg Gln Pro Gly Ala Thr Leu Glu Ala Gly Asp Ile Leu Gly Ile Leu Ser 755 760 765			2304
10	ctc gac gac ccg agc cgc gtc cac cac gcc aag ccg ttc gat ggc cag Leu Asp Asp Pro Ser Arg Val His His Ala Lys Pro Phe Asp Gly Gln 770 775 780			2352
15	ctt ccc gcc ctt ggc ttg ccc tcc atc gtc ggc aac aag ccg cac cag Leu Pro Ala Leu Gly Leu Pro Ser Ile Val Gly Asn Lys Pro His Gln 785 790 795 800			2400
20	cgc ttc gcc tac ctc aaa gac gtg ctc tca aac atc ctc atg ggc tac Arg Phe Ala Tyr Leu Lys Asp Val Leu Ser Asn Ile Leu Met Gly Tyr 805 810 815			2448
25	gac aac cag gcc gtc atg cag tcg agc atc aag gag ctc atc tcg gtt Asp Asn Gln Ala Val Met Gln Ser Ile Lys Glu Leu Ile Ser Val 820 825 830			2496
30	ctt cgc aac ccc gag ctc ccc tac ggc gag gcc aac gct gtc ctc tcg Leu Arg Asn Pro Glu Leu Pro Tyr Gly Glu Ala Asn Ala Val Leu Ser 835 840 845			2544
35	acg ctt tcg ggt cgc atc ccc gcc aag ctc gag cag acc ctc cgc cag Thr Leu Ser Gly Arg Ile Pro Ala Lys Leu Glu Gln Thr Leu Arg Gln 850 855 860			2592
40	tac atc gac cag gct cac gag tct ggc gcc gag ttc ccg tcc gcc aag Tyr Ile Asp Gln Ala His Glu Ser Gly Ala Glu Phe Pro Ser Ala Lys 865 870 875 880			2640
45	tgc cgc aag gcg atc gac acg acc ctt gag cag ctc cgc ccc gcc gag Cys Arg Lys Ala Ile Asp Thr Thr Leu Glu Gln Leu Arg Pro Ala Glu 885 890 895			2688
50	gcg cag act gtc cgc aac ttc ctc gtc gcg ttc gac gac atc gtc tac Ala Gln Thr Val Arg Asn Phe Leu Val Ala Phe Asp Asp Ile Val Tyr 900 905 910			2736
55	cgc tac cgc tcg ggc ctc aag cac cac gag tgg tca acg ctc gcc gcc Arg Tyr Arg Ser Gly Leu Lys His His Glu Trp Ser Thr Leu Ala Gly 915 920 925			2784
60	atc ttt gcc gcg tac gcc gag acg gag aag ccg ttc agc gcc aag gac Ile Phe Ala Ala Tyr Ala Glu Thr Glu Lys Pro Phe Ser Gly Lys Asp 930 935 940			2832
65	ggc gac gtc gtc ctc gag ctc cgc gac gcc cac cgc gac tcg ctc gac Gly Asp Val Val Leu Glu Leu Arg Asp Ala His Arg Asp Ser Leu Asp 945 950 955 960			2880
70	tcg gtc gtc aag atc gtt ctc tcg cac tac aag gct gcc tcg aag aac Ser Val Val Lys Ile Val Leu Ser His Tyr Lys Ala Ala Ser Lys Asn 965 970 975			2928
75	tcg ctt gtc ctt gcg ctc ctc gac atc gtc aag gac tcg gac gcg gtt Ser Leu Val Leu Ala Leu Leu Asp Ile Val Lys Asp Ser Asp Ala Val 980 985 990			2976
80	ccg ctc atc gag cag gtc gtc agc cct gcg ctc aag gac ctc gcc gac			3024

	Pro	Leu	Ile	Glu	Gln	Val	Val	Ser	Pro	Ala	Leu	Lys	Asp	Leu	Ala	Asp	
			995					1000					1005				
5	ctc	gac	tcg	aag	gcc	acg	act	aag	gtc	gcc	ctg	aag	gcc	cgc	gag		3069
	Leu	Asp	Ser	Lys	Ala	Thr	Thr	Lys	Val	Ala	Leu	Lys	Ala	Arg	Glu		
		1010					1015					1020					
	gtg	ctc	atc	cac	atc	cag	ctc	ccc	tcg	ctc	gac	gag	cgc	ctc	gga		3114
10	Val	Leu	Ile	His	Ile	Gln	Leu	Pro	Ser	Leu	Asp	Glu	Arg	Leu	Gly		
		1025					1030					1035					
	cag	ctc	gag	cag	att	ctc	aag	gcc	tcg	gtg	acg	ccc	acc	gtt	tac		3159
	Gln	Leu	Glu	Gln	Ile	Leu	Lys	Ala	Ser	Val	Thr	Pro	Thr	Val	Tyr		
		1040					1045					1050					
15	ggc	gag	ccc	ggc	cac	gac	cgc	act	cct	cgc	ggt	gaa	gtc	ctt	aag		3204
	Gly	Glu	Pro	Gly	His	Asp	Arg	Thr	Pro	Arg	Gly	Glu	Val	Leu	Lys		
		1055					1060					1065					
	gac	gtc	atc	gac	tcg	cgc	ttc	acc	gtc	ttt	gac	gtt	ctc	ccg	agc		3249
20	Asp	Val	Ile	Asp	Ser	Arg	Phe	Thr	Val	Phe	Asp	Val	Leu	Pro	Ser		
		1070					1075					1080					
	ttc	ttc	cag	cac	cag	gac	cac	tgg	gtc	tcg	ctc	gcc	gcg	ctc	gac		3294
	Phe	Phe	Gln	His	Gln	Asp	His	Trp	Val	Ser	Leu	Ala	Ala	Leu	Asp		
		1085					1090					1095					
25	acc	tac	gtc	cgc	cgc	gcc	tac	cgc	tcg	tac	aac	ctc	ctc	aac	atc		3339
	Thr	Tyr	Val	Arg	Arg	Ala	Tyr	Arg	Ser	Tyr	Asn	Leu	Leu	Asn	Ile		
		1100					1105					1110					
	gag	cac	atc	gag	gcc	gat	gcc	gcc	gag	gac	gag	ccc	gcg	acg	gtt		3384
30	Glu	His	Ile	Glu	Ala	Asp	Ala	Ala	Glu	Asp	Glu	Pro	Ala	Thr	Val		
		1115					1120					1125					
	gcc	tgg	tcg	ttc	cgc	atg	cgc	aag	gct	gcg	tcc	gag	tct	gag	ccg		3429
	Ala	Trp	Ser	Phe	Arg	Met	Arg	Lys	Ala	Ala	Ser	Glu	Ser	Glu	Pro		
		1130					1135					1140					
35	ccc	acg	ccc	acg	acc	ggc	ctc	acg	tcg	cag	cgc	acc	gcc	tcg	tac		3474
	Pro	Thr	Pro	Thr	Thr	Gly	Leu	Thr	Ser	Gln	Arg	Thr	Ala	Ser	Tyr		
		1145					1150					1155					
	tcg	gac	ttg	acg	ttc	ctc	ctc	aac	aac	gcc	cag	tcc	gag	ccg	atc		3519
40	Ser	Asp	Leu	Thr	Phe	Leu	Leu	Asn	Asn	Ala	Gln	Ser	Glu	Pro	Ile		
		1160					1165					1170					
	cgc	tac	ggc	gcg	atg	ttc	tcg	gtc	cgc	tcg	ctc	gac	cgc	ttc	cgc		3564
45	Arg	Tyr	Gly	Ala	Met	Phe	Ser	Val	Arg	Ser	Leu	Asp	Arg	Phe	Arg		
		1175					1180					1185					
	cag	gag	ctc	ggt	acc	gtc	ctc	cga	cac	ttc	ccc	gac	tcg	aac	aag		3609
	Gln	Glu	Leu	Gly	Thr	Val	Leu	Arg	His	Phe	Pro	Asp	Ser	Asn	Lys		
		1190					1195					1200					
50	ggc	aag	ctc	cag	cag	cag	cct	gcc	gcg	tcg	tcg	agc	cag	gag	cag		3654
	Gly	Lys	Leu	Gln	Gln	Gln	Pro	Ala	Ala	Ser	Ser	Ser	Gln	Glu	Gln		
		1205					1210					1215					
	tgg	aac	gtc	atc	aac	gtc	gcg	ctc	acg	gtc	ccc	gcc	agc	gcg	cag		3699
55	Trp	Asn	Val	Ile	Asn	Val	Ala	Leu	Thr	Val	Pro	Ala	Ser	Ala	Gln		
		1220					1225					1230					
60																	
65																	

	gtc gac gag gac gct ctc cgc gcc gac ttt gcc gct cac gtg aac	3744
	Val Asp Glu Asp Ala Leu Arg Ala Asp Phe Ala Ala His Val Asn	
	1235 1240 1245	
5	gcg atg agc gcc gag atc gac gct cgc gcc atg cgc cgc ctc acc	3789
	Ala Met Ser Ala Glu Ile Asp Ala Arg Gly Met Arg Arg Leu Thr	
	1250 1255 1260	
10	ctc ctc atc tgc cgc gag gcc cag tac ccg tcc tac tac acc gtc	3834
	Leu Leu Ile Cys Arg Glu Gly Gln Tyr Pro Ser Tyr Tyr Thr Val	
	1265 1270 1275	
	cgc aag cag gac gcc acc tgg aag gag ctc gag acg atc cgc gac	3879
	Arg Lys Gln Asp Gly Thr Trp Lys Glu Leu Glu Thr Ile Arg Asp	
	1280 1285 1290	
15	atc gag ccc gcc ctc gcc ttc cag ctc gag ttg gcc cgc ctc tcc	3924
	Ile Glu Pro Ala Leu Ala Phe Gln Leu Glu Leu Gly Arg Leu Ser	
	1295 1300 1305	
20	aac ttc cac ctc gag ccg tgc ccc gtt gag aac cgc cag gtc cac	3969
	Asn Phe His Leu Glu Pro Cys Pro Val Glu Asn Arg Gln Val His	
	1310 1315 1320	
	gtc tac tac gcg acc gcc aag gcc aac tcg tcc gac tgc cgc ttc	4014
	Val Tyr Tyr Ala Thr Ala Lys Gly Asn Ser Ser Asp Cys Arg Phe	
	1325 1330 1335	
25	ttc gtc cgc gca ctc gtc cgc cct gcc cgt ctc cgc ggt aac atg	4059
	Phe Val Arg Ala Leu Val Arg Pro Gly Arg Leu Arg Gly Asn Met	
	1340 1345 1350	
30	aag acg gcc gac tac ctc gtc tcc gag gct gac cgc ctc gtc acc	4104
	Lys Thr Ala Asp Tyr Leu Val Ser Glu Ala Asp Arg Leu Val Thr	
	1355 1360 1365	
	gat gtc ctc gac tcg ctc gag gtc gcc agc tcg cag cgc cgc gct	4149
	Asp Val Leu Asp Ser Leu Glu Val Ala Ser Ser Gln Arg Arg Ala	
	1370 1375 1380	
35	gcc gac gcc aac cac atc tcg ctc aac ttc ctg tac tct ctc cgt	4194
	Ala Asp Gly Asn His Ile Ser Leu Asn Phe Leu Tyr Ser Leu Arg	
	1385 1390 1395	
40	ctc gac ttt gac gag gtc cag gct gcc ctc gcc gcc ttc atc gac	4239
	Leu Asp Phe Asp Glu Val Gln Ala Ala Leu Ala Gly Phe Ile Asp	
	1400 1405 1410	
	cgc cac gcc aag cgc ttc tgg cgt ctc cgc gtc acc gcc gcc gag	4284
	Arg His Gly Lys Arg Phe Trp Arg Leu Arg Val Thr Gly Ala Glu	
	1415 1420 1425	
45	atc cgc atc gtc ctc gag gac gcg cag gcc aac att cag ccc atc	4329
	Ile Arg Ile Val Leu Glu Asp Ala Gln Gly Asn Ile Gln Pro Ile	
	1430 1435 1440	
50	cgc gcc atc atc gag aac gtc tcg ggt ttc gtc gtc aag tac gag	4374
	Arg Ala Ile Ile Glu Asn Val Ser Gly Phe Val Val Lys Tyr Glu	
	1445 1450 1455	
	gcg tac cgc gag gtc acg acc gac aag gcc cag gtc atc ctc aag	4419
	Ala Tyr Arg Glu Val Thr Thr Asp Lys Gly Gln Val Ile Leu Lys	
	1460 1465 1470	
55		
60		
65		

ES 2 724 998 T3

5	tcg atc ggt ccg cag gcc gcg ttg cac ctt cag ccg gtc aac ttc Ser Ile Gly Pro Gln Gly Ala Leu His Leu Gln Pro Val Asn Phe 1475 1480 1485	4464
10	ccc tac ccg acc aag gag tgg ctt cag ccg aag cgc tac aag gcc Pro Tyr Pro Thr Lys Glu Trp Leu Gln Pro Lys Arg Tyr Lys Ala 1490 1495 1500	4509
15	cac gtc gtc gcc acg acg tac gtc tac gac ttc ccc gac ctt ttc His Val Val Gly Thr Thr Tyr Val Tyr Asp Phe Pro Asp Leu Phe 1505 1510 1515	4554
20	cgc cag gca atc cgc aag cag tgg aag gcg gcc gcc aag act gcg Arg Gln Ala Ile Arg Lys Gln Trp Lys Ala Ala Gly Lys Thr Ala 1520 1525 1530	4599
25	ccc gcc gag ctc ctc gtc gcc aag gag ctc gtc ctc gac gag ttc Pro Ala Glu Leu Leu Val Ala Lys Glu Leu Val Leu Asp Glu Phe 1535 1540 1545	4644
30	ggc aag cct cag gag gtc gcc cgc ccg cct gcc acc aac aat atc Gly Lys Pro Gln Glu Val Ala Arg Pro Pro Gly Thr Asn Asn Ile 1550 1555 1560	4689
35	ggc atg gtc gcc tgg atc tac acg atc ttc acg ccc gaa tac ccc Gly Met Val Gly Trp Ile Tyr Thr Ile Phe Thr Pro Glu Tyr Pro 1565 1570 1575	4734
40	tct gcc cgc cgc gtc gtc gtc atc gcg aac gac atc acg ttc aag Ser Gly Arg Arg Val Val Val Ile Ala Asn Asp Ile Thr Phe Lys 1580 1585 1590	4779
45	att ggt tcg ttc gcc ccg gag gag gac cgc tac ttc ttc gcc gtc Ile Gly Ser Phe Gly Pro Glu Glu Asp Arg Tyr Phe Phe Ala Val 1595 1600 1605	4824
50	acg cag ctc gcg cgc caa ctt gcc ttg ccg cgc gtc tac ctc tcg Thr Gln Leu Ala Arg Gln Leu Gly Leu Pro Arg Val Tyr Leu Ser 1610 1615 1620	4869
55	gcc aac tcg ggt gct cgt ctc gcc att gcc gag gag ctc gtc gac Ala Asn Ser Gly Ala Arg Leu Gly Ile Ala Glu Glu Leu Val Asp 1625 1630 1635	4914
60	ttg ttc agc gtc gcg tgg gtc gac agc tcg ccg ccg gag aag gcc Leu Phe Ser Val Ala Trp Val Asp Ser Ser Arg Pro Glu Lys Gly 1640 1645 1650	4959
65	ttc aag tac ctc tac cta acc gcc gag aag ctc gcc gag ctc aag Phe Lys Tyr Leu Tyr Leu Thr Ala Glu Lys Leu Gly Glu Leu Lys 1655 1660 1665	5004
	aac aag gcc gag aag agc gtc atc acg aag cgc atc gag gac gag Asn Lys Gly Glu Lys Ser Val Ile Thr Lys Arg Ile Glu Asp Glu 1670 1675 1680	5049
	ggc gag acg cgc tac cag atc acc gac atc atc gcc ttg cag gag Gly Glu Thr Arg Tyr Gln Ile Thr Asp Ile Ile Gly Leu Gln Glu 1685 1690 1695	5094
	ggt ctc ggt gtc gag tcg ctc aag gcc tct gcc ctc atc gcc ggt Gly Leu Gly Val Glu Ser Leu Lys Gly Ser Gly Leu Ile Ala Gly 1700 1705 1710	5139

ES 2 724 998 T3

	1700	1705	1710	
5	gag acg tcg cgc gcg tac gac Glu Thr Ser Arg Ala Tyr Asp 1715 1720	gac atc ttc acg atc Asp Ile Phe Thr Ile Thr Leu Val 1725	acg ctc gtc 1725	5184
	acc gcc cgc tcg gtc ggt atc Thr Ala Arg Ser Val Gly Ile 1730 1735	ggt gcg tac ctc gtc Gly Ala Tyr Leu Val Arg Leu Gly 1740	cgc ctc gcc 1740	5229
10	cag cgt gcc gtc cag gtc gag Gln Arg Ala Val Gln Val Glu 1745 1750	ggc cag ccg atc atc Gly Gln Pro Ile Ile Leu Thr Gly 1755	ctc acc ggt 1755	5274
15	gcc gcc gcg ctc aac aag gtc Ala Gly Ala Leu Asn Lys Val 1760 1765	ctc ggt cgc gag gtg Leu Gly Arg Glu Val Tyr Ser Ser 1770	tac tcg tcc 1770	5319
	aac ttg cag ctc gcc gcc acg Asn Leu Gln Leu Gly Gly Thr 1775 1780	cag atc atg tac aag Gln Ile Met Tyr Lys Asn Gly Val 1785	aac ggt gtc 1785	5364
20	tcg cac ttg acg gcc gcc aac Ser His Leu Thr Ala Ala Asn 1790 1795	gac ctc gag ggt gtc Asp Leu Glu Gly Val Leu Ser Ile 1800	ctc agc atc 1800	5409
25	gtc cag tgg ctc gcc ttc gtc Val Gln Trp Leu Ala Phe Val 1805 1810	ccc gag cac cgc gcc Pro Glu His Arg Gly Ala Pro Leu 1815	gcg cct ctc 1815	5454
	cog atc atg cct tcg ccc gtc Pro Ile Met Pro Ser Pro Val 1820 1825	gac cog tgg gac cgc Asp Pro Trp Asp Arg Ser Ile Asp 1830	tcg atc gac 1830	5499
30	tac acg ccc atc aag gcc gcg Tyr Thr Pro Ile Lys Gly Ala 1835 1840	tac gac cog cgc tgg Tyr Asp Pro Arg Trp Phe Leu Ala 1845	ttc ctc gcc 1845	5544
35	ggc aag acg gac gag gcc gac Gly Lys Thr Asp Glu Ala Asp 1850 1855	ggt cgc tgg ctc tot Gly Arg Trp Leu Ser Gly Phe Phe 1860	ggc ttc ttc 1860	5589
	gac aag gcc tcg ttc cag gag Asp Lys Gly Ser Phe Gln Glu 1865 1870	acg ctc tcg gcc tgg Thr Leu Ser Gly Trp Ala Gln Thr 1875	gcg cag acc 1875	5634
40	gtc gtc gtc ggt cgc gct cgc Val Val Val Gly Arg Ala Arg 1880 1885	ctc gcc gcc atc ccc Leu Gly Gly Ile Pro Met Gly Ala 1890	atg gcc gcc 1890	5679
45	atc gcg gtc gag acc cgc acc Ile Ala Val Glu Thr Arg Thr 1895 1900	atc gag cgc gtc gtg Ile Glu Arg Val Val Pro Ala Asp 1905	ccc gcc gac 1905	5724
	cct gcc aac cct ctc tcg aac Pro Ala Asn Pro Leu Ser Asn 1910 1915	gag cag aag atc atg Glu Gln Lys Ile Met Gly Ala Gly 1920	gag gcc ggt 1920	5769
50	cag gtc tgg tat ccc aac agc Gln Val Trp Tyr Pro Asn Ser 1925 1930	tcg ttc aag acg gga Ser Phe Lys Thr Gly Gln Ala Ile 1935	cag gcg atc 1935	5814
55	ttc gac ttc aac cgc gag ggt 1935	ctc cog ctc atc atc 1940	ttc gcc aac 1940	5859

	Phe Asp 1940	Phe Asn Arg 1945	Glu Gly 1945	Leu Pro Leu Ile Ile 1950	Phe Ala Asn	
5	tgg cgc Trp Arg 1955	ggc ttc Gly Phe Ser 1960	tgc ggc Gly Gly 1960	cag cag gac Gln Gln Asp Met 1965	ttc gac gag gtc Phe Asp Glu Val 1965	5904
10	ctc aag Leu Lys 1970	cgc ggt Arg Gly Ser 1975	tgc ctc Leu Ile 1975	att gtc gac ggt Ile Val Asp Gly Leu 1980	tcg gcg tac aag Ser Ala Tyr Lys 1980	5949
	cag ccc Gln Pro 1985	gtc ttc Val Phe Val 1990	gtc tac Tyr Ile 1990	atc gtc ccg aac Ile Val Pro Asn Gly 1995	gaa ctt cgc ggc Glu Leu Arg Gly 1995	5994
15	ggg gct Gly Ala 2000	tgg gtc Trp Val Val 2005	gtc ctc Leu Asp 2005	cag cgc Pro Ser Ile Asn 2010	gag ggc atg Glu Gly Met 2010	6039
20	atg gag Met Glu 2015	atg tac Met Tyr Val 2020	gtc gac Val Asp 2020	gag act gct cgc Glu Thr Ala Arg Ala 2025	ggg gtc ctc gag Gly Val Leu Glu 2025	6084
	ccc gag Pro Glu 2030	ggc atc Gly Ile Val 2035	gtc gag Glu Ile 2035	atc aag ctc cgc Ile Lys Leu Arg Lys 2040	gac aag ctc ctc Asp Lys Leu Leu 2040	6129
25	gcc ctc Ala Leu 2045	atg gac Met Asp Arg 2050	cgc ctc Leu Asp 2050	gac ccg acc tac Asp Pro Thr Tyr His 2055	gcc ctc cgc gtc Ala Leu Arg Val 2055	6174
30	aag tcg Lys Ser 2060	acc gac Thr Asp Ala 2065	gct tcg ctc Ala Ser Leu 2065	tcg ccc gcc gac Ser Pro Ala Asp Ala 2070	gcc gcg cag gcc Ala Ala Gln Ala 2070	6219
	aag acc Lys Thr 2075	gag ctc Glu Leu Ala 2080	gcc gcg cgc Ala Ala Arg 2080	gag aag cag ctc Glu Lys Gln Leu Met 2085	atg ccg atc tac Met Pro Ile Tyr 2085	6264
35	cag cag Gln Gln 2090	gtc gcg ctc Val Ala Leu 2095	cag ttc Gln Phe 2095	gcc gac tcg cac Ala Asp Ser His 2100	gag aag gcc ggc Lys Ala Gly 2100	6309
40	cgc atc Arg Ile 2105	ctc agc Leu Ser Lys 2110	aag ggc Gly Cys 2110	tcg gcg cgc Ala Arg Glu Ala 2115	gag gtc gag Glu Trp Ser 2115	6354
45	aac gct Asn Ala 2120	cgt cgc Arg Arg Tyr 2125	tac ttc Phe Tyr 2125	gcc cgc ctc cgc Ala Arg Leu Arg 2130	cgt ctc gcc Arg Leu Ala 2130	6399
	gag gag Glu Glu 2135	gcc gcc Ala Ala Val 2140	aag cgt Lys Arg 2140	ctc gcc gac gcc Leu Gly Asp Ala 2145	gag ccg acc ctc Pro Thr Leu 2145	6444
50	tcg cgc Ser Arg 2150	gac gag Asp Glu Arg 2155	cgc ctc Leu Ala 2155	atc gtc cac gac Ile Val His Asp 2160	gcc gtc gcc cag Ala Val Gly Gln 2160	6489
55	ggg gtc Gly Val 2165	gac ctc Asp Leu Asn 2170	aac aac Asn Asp 2170	gac ctc gct gct Asp Leu Ala Ala 2175	gcc gcc gcg ttc Ala Ala Phe Glu 2175	6534

ES 2 724 998 T3

5
10
15
20
25
30
35
40
45
50
55
60
65

cag ggc gcc gcc gcc atc acc gag cgc gtc aag ctc gcg cgc gcg 6579
Gln Gly Ala Ala Ala Ile Thr Glu Arg Val Lys Leu Ala Arg Ala
2180 2185 2190

acg acc gtc gcc tcg act ctc gcg cag ctc gcg cag gac gac aag 6624
Thr Thr Val Ala Ser Thr Leu Ala Gln Leu Ala Gln Asp Asp Lys
2195 2200 2205

gag gct ttc gcc gcc tcg ctc cag cag gtc ctc ggc gac aag ctc 6669
Glu Ala Phe Ala Ala Ser Leu Gln Gln Val Leu Gly Asp Lys Leu
2210 2215 2220

acc gcc gcc gac ctc gcc cgc atc ctc gcc tag 6702
Thr Ala Ala Asp Leu Ala Arg Ile Leu Ala
2225 2230

<210> 91
<211> 2233
<212> PRT
<213> Rhodosporidium toruloides
<400> 91
Met Pro Phe Ser Gly Glu Ala Lys Ala Val Asn Gly Ser His Ser Val
1 5 10 15

Asp Glu Ala Pro Lys Asn Pro Lys Tyr Asp His Gly Arg Val Val Lys
20 25 30

Tyr Leu Gly Gly Asn Ser Leu Glu Ser Ala Pro Pro Ser Lys Val Ala
35 40 45

Asp Trp Val Arg Glu Arg Gly Gly His Thr Val Ile Thr Lys Ile Leu
50 55 60

Ile Ala Asn Asn Gly Ile Ala Ala Val Lys Glu Ile Arg Ser Val Arg
65 70 75 80

Lys Trp Ala Tyr Glu Thr Phe Gly Ser Glu Arg Ala Ile Glu Phe Thr
85 90 95

Val Met Ala Thr Pro Glu Asp Leu Lys Val Asn Ala Asp Tyr Ile Arg
100 105 110

Met Ala Asp Gln Tyr Val Glu Val Pro Gly Gly Thr Asn Asn Asn Asn
115 120 125

Tyr Ala Asn Val Asp Val Ile Val Asp Val Ala Glu Arg Ala Gly Val
130 135 140

His Ala Val Trp Ala Gly Trp Gly His Ala Ser Glu Asn Pro Arg Leu
145 150 155 160

Pro Glu Ser Leu Ala Ala Ser Lys His Lys Ile Val Phe Ile Gly Pro
165 170 175

5 Pro Gly Ser Ala Met Arg Ser Leu Gly Asp Lys Ile Ser Ser Thr Ile
180 185 190

10 Val Ala Gln His Ala Gln Val Pro Cys Met Asp Trp Ser Gly Gln Gly
195 200 205

Val Asp Gln Val Thr Gln Ser Pro Glu Gly Tyr Val Thr Val Ala Asp
210 215 220

15 Asp Val Tyr Gln Gln Ala Cys Val His Asp Ala Asp Glu Gly Leu Ala
225 230 235 240

20 Arg Ala Ser Arg Ile Gly Tyr Pro Val Met Ile Lys Ala Ser Glu Gly
245 250 255

Gly Gly Gly Lys Gly Ile Arg Lys Val Glu Lys Glu Gln Asp Phe Lys
260 265 270

25 Gln Ala Phe Gln Ala Val Leu Thr Glu Val Pro Gly Ser Pro Val Phe
275 280 285

30 Ile Met Lys Leu Ala Gly Ala Ala Arg His Leu Glu Val Gln Val Leu
290 295 300

Ala Asp Gln Tyr Gly Asn Ala Ile Ser Leu Phe Gly Arg Asp Cys Ser
305 310 315 320

35 Val Gln Arg Arg His Gln Lys Ile Ile Glu Glu Ala Pro Val Thr Ile
325 330 335

40 Ala Lys Pro Asp Thr Phe Glu Gln Met Glu Lys Ser Ala Val Arg Leu
340 345 350

Ala Lys Leu Val Gly Tyr Val Ser Ala Gly Thr Val Glu Phe Leu Tyr
355 360 365

45 Ser Ala Ala Asp Asp Lys Phe Ala Phe Leu Glu Leu Asn Pro Arg Leu
370 375 380

50 Gln Val Glu His Pro Thr Thr Glu Met Val Ser Gly Val Asn Leu Pro
385 390 395 400

Ala Ala Gln Leu Gln Val Ala Met Gly Val Pro Leu His Arg Ile Arg
405 410 415

55

60

65

Asp Ile Arg Thr Leu Tyr Gly Lys Ala Pro Asn Gly Ser Ser Glu Ile
 420 425 430
 5 Asp Phe Asp Phe Glu Asn Pro Glu Ser Ala Lys Thr Gln Arg Lys Pro
 435 440 445
 10 Ser Pro Lys Gly His Val Val Ala Val Arg Ile Thr Ala Glu Asn Pro
 450 455 460
 Asp Ala Gly Phe Lys Pro Ser Met Gly Thr Leu Gln Glu Leu Asn Phe
 465 470 475 480
 15 Arg Ser Ser Thr Asn Val Trp Gly Tyr Phe Ser Val Gly Ser Ala Gly
 485 490 495
 20 Gly Leu His Glu Phe Ala Asp Ser Gln Phe Gly His Ile Phe Ala Tyr
 500 505 510
 Gly Ser Asp Arg Ser Glu Ser Arg Lys Asn Met Val Val Ala Leu Lys
 515 520 525
 25 Glu Leu Ser Ile Arg Gly Asp Phe Arg Thr Thr Val Glu Tyr Leu Ile
 530 535 540
 30 Lys Leu Leu Glu Thr Asp Ala Phe Glu Gln Asn Thr Ile Thr Thr Ala
 545 550 555 560
 Trp Leu Asp Ser Leu Ile Ser Ala Arg Leu Thr Ala Glu Arg Pro Asp
 565 570 575
 35 Thr Thr Leu Ala Ile Ile Cys Gly Ala Val Thr Lys Ala His Leu Ala
 580 585 590
 40 Ser Glu Ala Asn Ile Ala Glu Tyr Lys Arg Ile Leu Glu Lys Gly Gln
 595 600 605
 Ser Pro Ala Lys Glu Leu Leu Ala Thr Val Val Pro Leu Glu Phe Val
 610 615 620
 45 Leu Glu Asp Val Lys Tyr Arg Ala Thr Ala Ser Arg Ser Ser Pro Ser
 625 630 635 640
 50 Ser Trp Ser Ile Tyr Val Asn Gly Ser Asn Val Ser Val Gly Ile Arg
 645 650 655
 55 Pro Leu Ala Asp Gly Gly Leu Leu Ile Leu Leu Asp Gly Arg Ser Tyr

ES 2 724 998 T3

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

Arg Tyr Arg Ser Gly Leu Lys His His Glu Trp Ser Thr Leu Ala Gly
 915 920 925
 5
 Ile Phe Ala Ala Tyr Ala Glu Thr Glu Lys Pro Phe Ser Gly Lys Asp
 930 935 940
 Gly Asp Val Val Leu Glu Leu Arg Asp Ala His Arg Asp Ser Leu Asp
 945 950 955 960
 10
 Ser Val Val Lys Ile Val Leu Ser His Tyr Lys Ala Ala Ser Lys Asn
 965 970 975
 15
 Ser Leu Val Leu Ala Leu Leu Asp Ile Val Lys Asp Ser Asp Ala Val
 980 985 990
 Pro Leu Ile Glu Gln Val Val Ser Pro Ala Leu Lys Asp Leu Ala Asp
 995 1000 1005
 20
 Leu Asp Ser Lys Ala Thr Thr Lys Val Ala Leu Lys Ala Arg Glu
 1010 1015 1020
 25
 Val Leu Ile His Ile Gln Leu Pro Ser Leu Asp Glu Arg Leu Gly
 1025 1030 1035
 Gln Leu Glu Gln Ile Leu Lys Ala Ser Val Thr Pro Thr Val Tyr
 1040 1045 1050
 30
 Gly Glu Pro Gly His Asp Arg Thr Pro Arg Gly Glu Val Leu Lys
 1055 1060 1065
 35
 Asp Val Ile Asp Ser Arg Phe Thr Val Phe Asp Val Leu Pro Ser
 1070 1075 1080
 Phe Phe Gln His Gln Asp His Trp Val Ser Leu Ala Ala Leu Asp
 1085 1090 1095
 40
 Thr Tyr Val Arg Arg Ala Tyr Arg Ser Tyr Asn Leu Leu Asn Ile
 1100 1105 1110
 45
 Glu His Ile Glu Ala Asp Ala Ala Glu Asp Glu Pro Ala Thr Val
 1115 1120 1125
 50
 Ala Trp Ser Phe Arg Met Arg Lys Ala Ala Ser Glu Ser Glu Pro
 1130 1135 1140
 Pro Thr Pro Thr Thr Gly Leu Thr Ser Gln Arg Thr Ala Ser Tyr
 1145 1150 1155
 55
 60
 65

Ser Asp Leu Thr Phe Leu Leu Asn Asn Ala Gln Ser Glu Pro Ile
 1160 1165 1170
 5 Arg Tyr Gly Ala Met Phe Ser Val Arg Ser Leu Asp Arg Phe Arg
 1175 1180 1185
 10 Gln Glu Leu Gly Thr Val Leu Arg His Phe Pro Asp Ser Asn Lys
 1190 1195 1200
 Gly Lys Leu Gln Gln Gln Pro Ala Ala Ser Ser Ser Gln Glu Gln
 1205 1210 1215
 15 Trp Asn Val Ile Asn Val Ala Leu Thr Val Pro Ala Ser Ala Gln
 1220 1225 1230
 20 Val Asp Glu Asp Ala Leu Arg Ala Asp Phe Ala Ala His Val Asn
 1235 1240 1245
 Ala Met Ser Ala Glu Ile Asp Ala Arg Gly Met Arg Arg Leu Thr
 1250 1255 1260
 25 Leu Leu Ile Cys Arg Glu Gly Gln Tyr Pro Ser Tyr Tyr Thr Val
 1265 1270 1275
 30 Arg Lys Gln Asp Gly Thr Trp Lys Glu Leu Glu Thr Ile Arg Asp
 1280 1285 1290
 Ile Glu Pro Ala Leu Ala Phe Gln Leu Glu Leu Gly Arg Leu Ser
 1295 1300 1305
 35 Asn Phe His Leu Glu Pro Cys Pro Val Glu Asn Arg Gln Val His
 1310 1315 1320
 40 Val Tyr Tyr Ala Thr Ala Lys Gly Asn Ser Ser Asp Cys Arg Phe
 1325 1330 1335
 Phe Val Arg Ala Leu Val Arg Pro Gly Arg Leu Arg Gly Asn Met
 1340 1345 1350
 45 Lys Thr Ala Asp Tyr Leu Val Ser Glu Ala Asp Arg Leu Val Thr
 1355 1360 1365
 50 Asp Val Leu Asp Ser Leu Glu Val Ala Ser Ser Gln Arg Arg Ala
 1370 1375 1380
 Ala Asp Gly Asn His Ile Ser Leu Asn Phe Leu Tyr Ser Leu Arg
 1385 1390 1395
 55
 60
 65

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

Leu Asp Phe Asp Glu Val Gln Ala Ala Leu Ala Gly Phe Ile Asp
 1400 1405 1410

Arg His Gly Lys Arg Phe Trp Arg Leu Arg Val Thr Gly Ala Glu
 1415 1420 1425

Ile Arg Ile Val Leu Glu Asp Ala Gln Gly Asn Ile Gln Pro Ile
 1430 1435 1440

Arg Ala Ile Ile Glu Asn Val Ser Gly Phe Val Val Lys Tyr Glu
 1445 1450 1455

Ala Tyr Arg Glu Val Thr Thr Asp Lys Gly Gln Val Ile Leu Lys
 1460 1465 1470

Ser Ile Gly Pro Gln Gly Ala Leu His Leu Gln Pro Val Asn Phe
 1475 1480 1485

Pro Tyr Pro Thr Lys Glu Trp Leu Gln Pro Lys Arg Tyr Lys Ala
 1490 1495 1500

His Val Val Gly Thr Thr Tyr Val Tyr Asp Phe Pro Asp Leu Phe
 1505 1510 1515

Arg Gln Ala Ile Arg Lys Gln Trp Lys Ala Ala Gly Lys Thr Ala
 1520 1525 1530

Pro Ala Glu Leu Leu Val Ala Lys Glu Leu Val Leu Asp Glu Phe
 1535 1540 1545

Gly Lys Pro Gln Glu Val Ala Arg Pro Pro Gly Thr Asn Asn Ile
 1550 1555 1560

Gly Met Val Gly Trp Ile Tyr Thr Ile Phe Thr Pro Glu Tyr Pro
 1565 1570 1575

Ser Gly Arg Arg Val Val Val Ile Ala Asn Asp Ile Thr Phe Lys
 1580 1585 1590

Ile Gly Ser Phe Gly Pro Glu Glu Asp Arg Tyr Phe Phe Ala Val
 1595 1600 1605

Thr Gln Leu Ala Arg Gln Leu Gly Leu Pro Arg Val Tyr Leu Ser
 1610 1615 1620

Ala Asn Ser Gly Ala Arg Leu Gly Ile Ala Glu Glu Leu Val Asp

	1625	1630	1635
5	Leu Phe 1640	Ser Val Ala Trp Val 1645	Asp Ser Ser Arg Pro 1650
10	Phe Lys 1655	Tyr Leu Tyr Leu Thr 1660	Ala Glu Lys Leu Gly 1665
15	Asn Lys 1670	Gly Glu Lys Ser Val 1675	Ile Thr Lys Arg Ile 1680
20	Gly Glu 1685	Thr Arg Tyr Gln Ile 1690	Thr Asp Ile Ile Gly 1695
25	Gly Leu 1700	Gly Val Glu Ser Leu 1705	Lys Gly Ser Gly Leu 1710
30	Glu Thr 1715	Ser Arg Ala Tyr Asp 1720	Asp Ile Phe Thr Ile 1725
35	Thr Ala 1730	Arg Ser Val Gly Ile 1735	Gly Ala Tyr Leu Val 1740
40	Gln Arg 1745	Ala Val Gln Val Glu 1750	Gly Gln Pro Ile Ile 1755
45	Ala Gly 1760	Ala Leu Asn Lys Val 1765	Leu Gly Arg Glu Val 1770
50	Asn Leu 1775	Gln Leu Gly Gly Thr 1780	Gln Ile Met Tyr Lys 1785
55	Ser His 1790	Leu Thr Ala Ala Asn 1795	Asp Leu Glu Gly Val 1800
60	Val Gln 1805	Trp Leu Ala Phe Val 1810	Pro Glu His Arg Gly 1815
65	Pro Ile 1820	Met Pro Ser Pro Val 1825	Asp Pro Trp Asp Arg 1830
70	Tyr Thr 1835	Pro Ile Lys Gly Ala 1840	Tyr Asp Pro Arg Trp 1845
75	Gly Lys 1850	Thr Asp Glu Ala Asp 1855	Gly Arg Trp Leu Ser 1860

5 Asp Lys Gly Ser Phe Gln Glu Thr Leu Ser Gly Trp Ala Gln Thr
 1865 1870 1875

Val Val Val Gly Arg Ala Arg Leu Gly Gly Ile Pro Met Gly Ala
 1880 1885 1890

10 Ile Ala Val Glu Thr Arg Thr Ile Glu Arg Val Val Pro Ala Asp
 1895 1900 1905

15 Pro Ala Asn Pro Leu Ser Asn Glu Gln Lys Ile Met Glu Ala Gly
 1910 1915 1920

Gln Val Trp Tyr Pro Asn Ser Ser Phe Lys Thr Gly Gln Ala Ile
 1925 1930 1935

20 Phe Asp Phe Asn Arg Glu Gly Leu Pro Leu Ile Ile Phe Ala Asn
 1940 1945 1950

25 Trp Arg Gly Phe Ser Gly Gly Gln Gln Asp Met Phe Asp Glu Val
 1955 1960 1965

Leu Lys Arg Gly Ser Leu Ile Val Asp Gly Leu Ser Ala Tyr Lys
 1970 1975 1980

30 Gln Pro Val Phe Val Tyr Ile Val Pro Asn Gly Glu Leu Arg Gly
 1985 1990 1995

35 Gly Ala Trp Val Val Leu Asp Pro Ser Ile Asn Ala Glu Gly Met
 2000 2005 2010

Met Glu Met Tyr Val Asp Glu Thr Ala Arg Ala Gly Val Leu Glu
 2015 2020 2025

40 Pro Glu Gly Ile Val Glu Ile Lys Leu Arg Lys Asp Lys Leu Leu
 2030 2035 2040

45 Ala Leu Met Asp Arg Leu Asp Pro Thr Tyr His Ala Leu Arg Val
 2045 2050 2055

Lys Ser Thr Asp Ala Ser Leu Ser Pro Ala Asp Ala Ala Gln Ala
 2060 2065 2070

50 Lys Thr Glu Leu Ala Ala Arg Glu Lys Gln Leu Met Pro Ile Tyr
 2075 2080 2085

55 Gln Gln Val Ala Leu Gln Phe Ala Asp Ser His Asp Lys Ala Gly
 2090 2095 2100

5 Arg Ile Leu Ser Lys Gly Cys Ala Arg Glu Ala Leu Glu Trp Ser
 2105 2110 2115
 Asn Ala Arg Arg Tyr Phe Tyr Ala Arg Leu Arg Arg Arg Leu Ala
 2120 2125 2130
 10 Glu Glu Ala Ala Val Lys Arg Leu Gly Asp Ala Asp Pro Thr Leu
 2135 2140 2145
 15 Ser Arg Asp Glu Arg Leu Ala Ile Val His Asp Ala Val Gly Gln
 2150 2155 2160
 Gly Val Asp Leu Asn Asn Asp Leu Ala Ala Ala Ala Ala Phe Glu
 2165 2170 2175
 20 Gln Gly Ala Ala Ala Ile Thr Glu Arg Val Lys Leu Ala Arg Ala
 2180 2185 2190
 25 Thr Thr Val Ala Ser Thr Leu Ala Gln Leu Ala Gln Asp Asp Lys
 2195 2200 2205
 Glu Ala Phe Ala Ala Ser Leu Gln Gln Val Leu Gly Asp Lys Leu
 2210 2215 2220
 30 Thr Ala Ala Asp Leu Ala Arg Ile Leu Ala
 2225 2230
 <210> 92
 <211> 1925
 35 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 92
 accgcgtcca aggagaagga gcacgatgtc ctgcctctt ccgactccga ggacgagcac 60
 40 aaccaggacc cgctcaaggc cctcgagaac gactaccgc cgttcgtcgt gcccaactac 120
 tccatcaagg agttgctcgg cgccatcccc gccactgct tcgagcgcag cgccctccgc 180
 tcgtcgtgt acgtcctcgg cgactttgcc atgctcgcgg gtctcgggta cgccgcgtcg 240
 45 cacatcgacc cggccttttc gttcgacggc ggcaagggtc tcagcggctg ggccggattc 300
 gcagccaagt gggccctctg gtcgcctac tgggtcctcg ccggtcgggt cggaacgggc 360
 gtctggatcc tcggacacga gtgcggccac caggcgttct cgacgtccaa gacgatcaac 420
 50 aacacgatgg gcctcttctt ccactcattc gttctcgtcc cgtaccactc gtggcgcatac 480
 tcgcacgcca agcaccacgc cgccacggga cacatgactc gcgacgaggt ctttgtcccg 540
 cgcaccgcgt cgttcgcga ccccaagccc accggcaaga aactccgcgt ctgcacaaac 600
 55
 60
 65

	atcgagctcg acgagctcct cgaggatgcg cccctgtacc gcctcggtcg gcttctcgtc	660
5	cagcagctct tcggctggcc cgcgtacctc ttctcgaacg catcgggaca gctctggtac	720
	cccaagtgga ccaaccactt cgaccogtcg tcgctcgtct ttgaagcccg ccaccgcggt	780
	caggctctcg tctcggacgc ctctctcgca ggcctggctg gcctcctcgt cgcctttggc	840
10	caggctcgtc gactcgcagg cgctcgtcaag tactacttca ttctctacct ctttgtgtga	900
	gtcccgactt cggatcctcc gaacttgctt tctctctcct tgatctttcg ccgagttccc	960
	gctgactctc ccggacctcc ctgcgcgacg acagcaacca ctggctcgtc atgatcacct	1020
15	acctccagca caccgacccc tcgctccctc actacaacgc cggcatgtgg aacttccagc	1080
	goggcgcgct ctgcaccatg gaccgcaaca tgctcggccc cgtcggctct tacctgatgc	1140
	acggcatctg cgagacccac ggtgcgttgc agtctcgcgc gaatctcgtt tagttgggcg	1200
20	gcagtggctg actttgcttc tctctcctct ctctcctcct tctcactccc ttctctcctt	1260
	cccacgcttc ctacgacccc tctccacatt ccacacatcg atgtgcacca ccgcagtcg	1320
	cgcaccacct ctctcgaag atcccgcact accacgcgtg ggaggcgaca gaggcgctca	1380
25	agaacttcct tggcgagcac tacaactaca ccgatgaggg gatgttcagg tcgctctgga	1440
	aggcgtatag gcagtgcggg ttctcgcgag ctctctcggg ccttgtttga acctttctag	1500
30	acgtcgacga tcataattctg acctcccgct cgcctttcct cgcttcacgc agctacgtcg	1560
	acgatgaggg cgactcctc ttctaccgcg acgcctacgg ccgcgcacgc cgcgtcgcg	1620
	tccccgcgga ggtccctccc gactcgggcg tcgagggact ctacacgatg ccctagagcg	1680
35	agactctttc cttcaccttc cccgcttctc gtagatccct ctctttggtt caagacgga	1740
	ctcctctcgt cgtcgtcgtt agatttctcc ttccctctcc ctctctctca ccactctcct	1800
	ctcgagtcgc tcgagtcgtc gttctgttcg gtctgtgcgg tgtgccctcc cttcttcgag	1860
40	tgaagggagg acgcaaaggt ctttgtgtga agcaattatc cactctctac gctcctttta	1920
	cacct	1925
	<210> 93	
	<211> 1107	
45	<212> ADN	
	<213> Rhodosporidium toruloides	
	<220>	
	<221> CDS	
	<222> (1)..(1107)	
50	<400> 93	
	atg ctc gcg ggt ctc ggg tac gcc gcg tcg cac atc gac ccg gcc ttt	48
	Met Leu Ala Gly Leu Gly Tyr Ala Ala Ser His Ile Asp Pro Ala Phe	
	1 5 10 15	
55	tcg ttc gac ggc ggc aag gtg ctc agc ggc tgg gcc gga ttc gca gcc	96
60		
65		

	Ser	Phe	Asp	Gly	Gly	Lys	Val	Leu	Ser	Gly	Trp	Ala	Gly	Phe	Ala	Ala	
				20					25					30			
5	aag	tgg	gcc	ctc	tgg	tcc	gcc	tac	tgg	gtc	ctc	gcc	ggc	tgg	gtc	gga	144
	Lys	Trp	Ala	Leu	Trp	Ser	Ala	Tyr	Trp	Val	Leu	Ala	Gly	Trp	Val	Gly	
			35					40					45				
10	acg	ggc	gtc	tgg	atc	ctc	gga	cac	gag	tgc	ggc	cac	cag	gog	ttc	tcg	192
	Thr	Gly	Val	Trp	Ile	Leu	Gly	His	Glu	Cys	Gly	His	Gln	Ala	Phe	Ser	
		50					55					60					
15	acg	tcc	aag	acg	atc	aac	aac	acg	atg	ggc	ctc	ttc	ctc	cac	tca	ttc	240
	Thr	Ser	Lys	Thr	Ile	Asn	Asn	Thr	Met	Gly	Leu	Phe	Leu	His	Ser	Phe	
		65				70					75				80		
20	ggt	ctc	gtc	cgg	tac	cac	tcg	tgg	cgc	atc	tcg	cac	gcc	aag	cac	cac	288
	Val	Leu	Val	Pro	Tyr	His	Ser	Trp	Arg	Ile	Ser	His	Ala	Lys	His	His	
				85						90				95			
25	gcc	gcc	acg	gga	cac	atg	act	cgc	gac	gag	gtc	ttt	gtc	cgg	cgc	acc	336
	Ala	Ala	Thr	Gly	His	Met	Thr	Arg	Asp	Glu	Val	Phe	Val	Pro	Arg	Thr	
			100					105					110				
30	gog	tcg	ttc	cgc	aac	ccc	aag	ccc	acc	ggc	aag	aaa	ctc	cgc	gtc	tcg	384
	Ala	Ser	Phe	Arg	Asn	Pro	Lys	Pro	Thr	Gly	Lys	Lys	Leu	Arg	Val	Ser	
			115				120						125				
35	cac	aac	atc	gag	ctc	gac	gag	ctc	ctc	gag	gat	gog	ccc	ctg	tac	cgc	432
	His	Asn	Ile	Glu	Leu	Asp	Glu	Leu	Leu	Glu	Asp	Ala	Pro	Leu	Tyr	Arg	
		130				135						140					
40	ctc	ggc	tgg	ctt	ctc	gtc	cag	cag	ctc	ttc	ggc	tgg	ccc	gog	tac	ctc	480
	Leu	Gly	Trp	Leu	Leu	Val	Gln	Gln	Leu	Phe	Gly	Trp	Pro	Ala	Tyr	Leu	
		145				150					155				160		
45	ttc	tcg	aac	gca	tcg	gga	cag	ctc	tgg	tac	ccc	aag	tgg	acc	aac	cac	528
	Phe	Ser	Asn	Ala	Ser	Gly	Gln	Leu	Trp	Tyr	Pro	Lys	Trp	Thr	Asn	His	
			165					170						175			
50	ttc	gac	ccg	tcg	tcg	ctc	gtc	ttt	gac	ggc	cgc	cac	cgc	ggt	cag	gtc	576
	Phe	Asp	Pro	Ser	Ser	Leu	Val	Phe	Asp	Ala	Arg	His	Arg	Gly	Gln	Val	
			180					185						190			
55	ctc	gtc	tcg	gac	gcc	ttc	ctc	gca	ggc	atg	gtc	ggc	ctc	ctc	gtc	gcc	624
	Leu	Val	Ser	Asp	Ala	Phe	Leu	Ala	Gly	Met	Val	Gly	Leu	Leu	Val	Ala	
			195					200					205				
60	ttt	ggc	cag	gtc	gtc	gga	ctc	gca	ggc	gtc	gtc	aag	tac	tac	ttc	att	672
	Phe	Gly	Gln	Val	Val	Gly	Leu	Ala	Gly	Val	Val	Lys	Tyr	Tyr	Phe	Ile	
		210				215						220					
65	cct	tac	ctc	ttt	gtc	aac	cac	tgg	ctc	gtc	atg	atc	acc	tac	ctc	cag	720
	Pro	Tyr	Leu	Phe	Val	Asn	His	Trp	Leu	Val	Met	Ile	Thr	Tyr	Leu	Gln	
		225				230					235				240		
70	cac	acg	gac	ccc	tcg	ctc	cct	cac	tac	aac	ggc	ggc	atg	tgg	aac	ttc	768
	His	Thr	Asp	Pro	Ser	Leu	Pro	His	Tyr	Asn	Ala	Gly	Met	Trp	Asn	Phe	
			245							250				255			
75	cag	cgc	ggc	gog	ctc	tgc	acc	atg	gac	cgc	aac	atg	ctc	ggc	ccc	gtc	816
	Gln	Arg	Gly	Ala	Leu	Cys	Thr	Met	Asp	Arg	Asn	Met	Leu	Gly	Pro	Val	
			260					265						270			

5 ggt cct tac ctg atg cac ggc atc tgc gag acc cac gtc gcg cac cac 864
Gly Pro Tyr Leu Met His Gly Ile Cys Glu Thr His Val Ala His His
275 280 285

10 ctc tcg tcc aag atc ccg cac tac cac gcg tgg gag gcg aca gag gcg 912
Leu Ser Ser Lys Ile Pro His Tyr His Ala Trp Glu Ala Thr Glu Ala
290 295 300

15 ctc aag aac ttc ctt ggc gag cac tac aac tac acc gat gag ggg atg 960
Leu Lys Asn Phe Leu Gly Glu His Tyr Asn Tyr Thr Asp Glu Gly Met
305 310 315 320

20 ttc agg tcg ctc tgg aag gcg tat agg cag tgc cgc tac gtc gac gat 1008
Phe Arg Ser Leu Trp Lys Ala Tyr Arg Gln Cys Arg Tyr Val Asp Asp
325 330 335

25 gag ggc gac gtc ctc ttc tac cgc gac gcc tac ggc cgc gca cgc cgc 1056
Glu Gly Asp Val Leu Phe Tyr Arg Asp Ala Tyr Gly Arg Ala Arg Arg
340 345 350

30 gtc gcc gtc ccc gcc gag gtc ccc tcc gac tcg ggc gtc gag gga ctc 1104
Val Ala Val Pro Ala Glu Val Pro Ser Asp Ser Gly Val Glu Gly Leu
355 360 365

35 tag 1107
<210> 94
<211> 368
<212> PRT
<213> *Rhodospiridium toruloides*
<400> 94

40 Met Leu Ala Gly Leu Gly Tyr Ala Ala Ser His Ile Asp Pro Ala Phe
1 5 10 15

45 Ser Phe Asp Gly Gly Lys Val Leu Ser Gly Trp Ala Gly Phe Ala Ala
20 25 30

50 Lys Trp Ala Leu Trp Ser Ala Tyr Trp Val Leu Ala Gly Trp Val Gly
35 40 45

55 Thr Gly Val Trp Ile Leu Gly His Glu Cys Gly His Gln Ala Phe Ser
50 55 60

60 Thr Ser Lys Thr Ile Asn Asn Thr Met Gly Leu Phe Leu His Ser Phe
65 70 75 80

65 Val Leu Val Pro Tyr His Ser Trp Arg Ile Ser His Ala Lys His His
85 90 95

70 Ala Ala Thr Gly His Met Thr Arg Asp Glu Val Phe Val Pro Arg Thr
100 105 110

75 Ala Ser Phe Arg Asn Pro Lys Pro Thr Gly Lys Lys Leu Arg Val Ser

ES 2 724 998 T3

	115	120	125
5	His Asn Ile Glu Leu Asp 130	Glu Leu Leu Glu Asp 135	Ala Pro Leu Tyr Arg 140
10	Leu Gly Trp Leu Leu Val 145	Gln Gln Leu Phe 150	Gly Trp Pro Ala Tyr Leu 155 160
15	Phe Ser Asn Ala Ser 165	Gly Gln Leu Trp Tyr 170	Pro Lys Trp Thr Asn His 175
20	Phe Asp Pro Ser Ser 180	Leu Val Phe Asp 185	Ala Arg His Arg Gly Gln Val 190
25	Leu Val Ser Asp Ala Phe 195	Leu Ala Gly Met 200	Val Gly Leu Leu Val Ala 205
30	Phe Gly Gln Val Val Gly 210	Leu Ala Gly Val Val 215	Lys Tyr Tyr Phe Ile 220
35	Pro Tyr Leu Phe Val Asn 225	His Trp Leu Val Met 230	Ile Thr Tyr Leu Gln 235 240
40	His Thr Asp Pro Ser 245	Leu Pro His Tyr Asn 250	Ala Gly Met Trp Asn Phe 255
45	Gln Arg Gly Ala Leu Cys 260	Thr Met Asp Arg Asn 265	Met Leu Gly Pro Val 270
50	Gly Pro Tyr Leu Met His 275	Gly Ile Cys Glu Thr 280	His Val Ala His His 285
55	Leu Ser Ser Lys Ile Pro 290	His Tyr His Ala Trp 295	Glu Ala Thr Glu Ala 300
60	Leu Lys Asn Phe Leu Gly 305	Glu His Tyr Asn Tyr 310	Thr Asp Glu Gly Met 315 320
65	Phe Arg Ser Leu Trp Lys 325	Ala Tyr Arg Gln Cys 330	Arg Tyr Val Asp Asp 335
70	Glu Gly Asp Val Leu Phe 340	Tyr Arg Asp Ala Tyr 345	Gly Arg Ala Arg Arg 350
75	Val Ala Val Pro Ala Glu 355	Val Pro Ser Asp Ser 360	Gly Val Glu Gly Leu 365
80	<210> 95 <211> 1371 <212> ADN <213> Phytophthora infestans <220> <221> CDS <222> (1)..(1371) <400> 95		

ES 2 724 998 T3

5	atg gtg gac ggc ccc aag acc aag cgc aag att tcg tgg cag gag gtc Met Val Asp Gly Pro Lys Thr Lys Arg Lys Ile Ser Trp Gln Glu Val 1 5 10 15	48
	aag cag cac gcc tcg tac gac aac gcg tgg atc gtc atc cac cac aag Lys Gln His Ala Ser Tyr Asp Asn Ala Trp Ile Val Ile His His Lys 20 25 30	96
10	gtc tac gac atc tcc aag tgg gac gcg cac ccc ggc ggc atg gtc atg Val Tyr Asp Ile Ser Lys Trp Asp Ala His Pro Gly Gly Met Val Met 35 40 45	144
15	ctc tcg cag gcg ggc gag gac gcg acc gac atc ttc acc gtc tgc cac Leu Ser Gln Ala Gly Glu Asp Ala Thr Asp Ile Phe Thr Val Cys His 50 55 60	192
	ccc acc tcg tcg tgg aag cag ctt gag cag ttc tac atc ggc gac gtg Pro Thr Ser Ser Trp Lys Gln Leu Glu Gln Phe Tyr Ile Gly Asp Val 65 70 75 80	240
20	gac gag tcc acc gcg acc gtc aac gag gac ctc tcg gag gag gag aag Asp Glu Ser Thr Ala Thr Val Asn Glu Asp Leu Ser Glu Glu Glu Lys 85 90 95	288
25	gcc aag aag gcg aag acc gac gag ttc atc tcg gcg tac cgc cgc ctc Ala Lys Lys Ala Lys Thr Asp Glu Phe Ile Ser Ala Tyr Arg Arg Leu 100 105 110	336
	cgc atc aag atc aag ggc atg ggc ctt tac gac gcg tcg atg gtc ttc Arg Ile Lys Ile Lys Gly Met Gly Leu Tyr Asp Ala Ser Met Val Phe 115 120 125	384
30	tac gcg tgg aag atc ctc tcg acc ttc ggc ctc tgg atg gcg tcg gcg Tyr Ala Trp Lys Ile Leu Ser Thr Phe Gly Leu Trp Met Ala Ser Ala 130 135 140	432
35	gcg atc tgc tgg cac ttc gac tcg tgg ccc atg tac atg ctt gcg gcg Ala Ile Cys Trp His Phe Asp Ser Trp Pro Met Tyr Met Leu Ala Ala 145 150 155 160	480
40	tgc gtc atg gga ctt ttc tgg cag cag tcg ggc tgg ctc gcg cac gac Cys Val Met Gly Leu Phe Trp Gln Gln Ser Gly Trp Leu Ala His Asp 165 170 175	528
	gtg ctc cac cac cag gtc tgg gac aac cac atg atc ggc aac gtc atg Val Leu His His Gln Val Trp Asp Asn His Met Ile Gly Asn Val Met 180 185 190	576
45	ggc gtc atc atc gga gac gtc tgg atg ggc ttc tcc gtc cag tgg tgg Gly Val Ile Ile Gly Asp Val Trp Met Gly Phe Ser Val Gln Trp Trp 195 200 205	624

5 aag aac aag cac aac ttc cac cac gcg gtc ccc aac ctc atc ggt gac 672
Lys Asn Lys His Asn Phe His His Ala Val Pro Asn Leu Ile Gly Asp
210 215 220

gag aag acg aag tac ctc gcc gac ccc gac atc gac acc atg ccc ctc 720
Glu Lys Thr Lys Tyr Leu Gly Asp Pro Asp Ile Asp Thr Met Pro Leu
225 230 235 240

10 ctc gcg tgg tgc aag cac atg gcg tgc aag gcg tac gag tgc tgc tgg 768
Leu Ala Trp Ser Lys His Met Ala Ser Lys Ala Tyr Glu Ser Ser Trp
245 250 255

15 gcc ccc ttc ttc gtc gcc cac cag gcg gtc atc tac ttc ccc ctc ctc 816
Gly Pro Phe Phe Val Gly His Gln Ala Val Ile Tyr Phe Pro Leu Leu
260 265 270

ctc ttc gcg cgc ttc tgc tgg ctc ctc cag tgc tac tac tac gtc ttc 864
Leu Phe Ala Arg Phe Ser Trp Leu Leu Gln Ser Tyr Tyr Tyr Val Phe
275 280 285

20 aag gcc ttc gcg ttc gcc aag tac gac ccc gtg gac ctc ccc aac gcc 912
Lys Gly Phe Ala Phe Gly Lys Tyr Asp Pro Val Asp Leu Pro Asn Gly
290 295 300

25 gag aag gtc gcc ctc atg ctc cac tac atc tgg aac gtc atg ctc ccc 960
Glu Lys Val Gly Leu Met Leu His Tyr Ile Trp Asn Val Met Leu Pro
305 310 315 320

gtc gtc acc gga atg tcc gtc gcg cag gcc ctc gcg ttc ttc atg ctt 1008
Val Val Thr Gly Met Ser Val Ala Gln Gly Leu Ala Phe Phe Met Leu
325 330 335

30 gcg cag atg tgc tgc gcc gcc ttc ctc gcg gcg gtc ttc tgc gtc gcc 1056
Ala Gln Met Ser Cys Gly Gly Phe Leu Ala Ala Val Phe Ser Val Gly
340 345 350

35 cac aac gcc atg tgc gtc tac gag cgc gag gac aag ccc gac ttc tgg 1104
His Asn Gly Met Ser Val Tyr Glu Arg Glu Asp Lys Pro Asp Phe Trp
355 360 365

cag ctc cag gtc acc acc acc cgc aac atc acc ccc gcc ttc ttc atg 1152
Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Thr Pro Gly Phe Phe Met
370 375 380

40 gac tgg ttc tgc gcc gcc ctc aac tac cag atc gag cac cac ctc ttc 1200
Asp Trp Phe Cys Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Phe
385 390 395 400

45 ccc atg atg ccc cgc cac aac ctc cag aag gtc aac ccc ctc gtc aag 1248
Pro Met Met Pro Arg His Asn Leu Gln Lys Val Asn Pro Leu Val Lys
405 410 415

tgc ctc tgc aag cag tac gac gtg cgc ttc cac gag acg gcc ttc tac 1296
Ser Leu Cys Lys Gln Tyr Asp Val Arg Phe His Glu Thr Gly Phe Tyr
420 425 430

50 cgg gcc ctc gtc gag gtc gtg gac gag ctt gcg gac atc tgc aag gag 1344
Arg Gly Leu Val Glu Val Val Asp Glu Leu Ala Asp Ile Ser Lys Glu
435 440 445

55 ttc ctc ctt gag ttc ccc gcg atg tag 1371
Phe Leu Leu Glu Phe Pro Ala Met
450 455
<210> 96
<211> 456
<212> PRT
<213> Phytophthora infestans
60 <400> 96

65

ES 2 724 998 T3

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

Glu Lys Thr Lys Tyr Leu Gly Asp Pro Asp Ile Asp Thr Met Pro Leu
 225 230 235 240
 5 Leu Ala Trp Ser Lys His Met Ala Ser Lys Ala Tyr Glu Ser Ser Trp
 245 250 255
 10 Gly Pro Phe Phe Val Gly His Gln Ala Val Ile Tyr Phe Pro Leu Leu
 260 265 270
 Leu Phe Ala Arg Phe Ser Trp Leu Leu Gln Ser Tyr Tyr Tyr Val Phe
 275 280 285
 15 Lys Gly Phe Ala Phe Gly Lys Tyr Asp Pro Val Asp Leu Pro Asn Gly
 290 295 300
 20 Glu Lys Val Gly Leu Met Leu His Tyr Ile Trp Asn Val Met Leu Pro
 305 310 315 320
 Val Val Thr Gly Met Ser Val Ala Gln Gly Leu Ala Phe Phe Met Leu
 325 330 335
 25 Ala Gln Met Ser Cys Gly Gly Phe Leu Ala Ala Val Phe Ser Val Gly
 340 345 350
 30 His Asn Gly Met Ser Val Tyr Glu Arg Glu Asp Lys Pro Asp Phe Trp
 355 360 365
 35 Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Thr Pro Gly Phe Phe Met
 370 375 380
 Asp Trp Phe Cys Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Phe
 385 390 395 400
 40 Pro Met Met Pro Arg His Asn Leu Gln Lys Val Asn Pro Leu Val Lys
 405 410 415
 45 Ser Leu Cys Lys Gln Tyr Asp Val Arg Phe His Glu Thr Gly Phe Tyr
 420 425 430
 Arg Gly Leu Val Glu Val Val Asp Glu Leu Ala Asp Ile Ser Lys Glu
 435 440 445
 50 Phe Leu Leu Glu Phe Pro Ala Met
 450 455
 <210> 97
 <211> 1374
 <212> ADN
 <213> Mortierella alpina
 <220>
 <221> CDS
 60 <222> (1)..(1374)
 <400> 97
 65

5	atg gcg gcg gcg ccc tcg gtc cgc acc ttc acc cgc gcg gag atc ctc Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu Ile Leu 1 5 10 15	48
10	aac gcc gag gcc ctc aac gag ggc aag aag gac gcg gag gcg cgc ttc Asn Ala Glu Ala Leu Asn Glu Gly Lys Lys Asp Ala Glu Ala Pro Phe 20 25 30	96
15	ctt atg atc atc gac aac aag gtc tac gac gtg cgc gag ttc gtc ccc Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Val Pro 35 40 45	144
20	gac cac ccc ggc ggc tcg gtc atc ctc acc cac gtc ggc aag gac ggc Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly 50 55 60	192
25	acc gac gtg ttc gac acc ttc cac ccc gag gcg gcg tgg gag acg ctc Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu Thr Leu 65 70 75 80	240
30	gcg aac ttc tac gtc ggc gac att gac gag tcc gac cgc gcg atc aag Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Ala Ile Lys 85 90 95	288
35	aac gac gac ttc gcg gcg gag gtc cgc aag ctc cgc acc ctc ttc cag Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu Phe Gln 100 105 110	336
40	tcg ctc ggc tac tac gac tcg tcg aag gcg tac tac gcg ttc aag gtg Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val 115 120 125	384
45	tcg ttc aac ctc tgc atc tgg ggc ctc tcg acc ttc atc gtc gcg aag Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Phe Ile Val Ala Lys 130 135 140	432
50	tgg ggc cag acc tcg acc ctc gcg aac gtc ctc tcg gcc gcg ctc ctc Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu 145 150 155 160	480
55	ggc ctt ttc tgg cag cag tgc ggc tgg ctc gcc cac gac ttc ctc cac Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His 165 170 175	528
60	cac cag gtc ttc cag gac cgc ttc tgg ggc gac ctc ttc ggc gcg ttc His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe 180 185 190	576
65	ctc ggc ggc gtc tgc cag ggc ttc tcg tcg tcg tgg tgg aag gac aag Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Trp Trp Lys Asp Lys 195 200 205	624
70	cac aac acc cac cac gcg gcg ccc aac gtc cac ggc gag gac ccc gac	672

ES 2 724 998 T3

	His	Asn	Thr	His	His	Ala	Ala	Pro	Asn	Val	His	Gly	Glu	Asp	Pro	Asp	
	210						215					220					
5	atc	gac	acc	cac	ccc	ctc	ctc	acc	tgg	tgc	gag	cac	gcg	ctt	gag	atg	720
	Ile	Asp	Thr	His	Pro	Leu	Leu	Thr	Trp	Ser	Glu	His	Ala	Leu	Glu	Met	
	225					230					235				240		
	ttc	tgc	gac	gtg	ccc	gac	gag	gag	ctt	acc	cgc	atg	tgg	tgc	cgc	ttc	768
10	Phe	Ser	Asp	Val	Pro	Asp	Glu	Glu	Leu	Thr	Arg	Met	Trp	Ser	Arg	Phe	
					245					250					255		
	atg	gtc	ctc	aac	cag	acc	tgg	ttc	tac	ttc	ccc	atc	ctc	tgc	ttc	gcg	816
	Met	Val	Leu	Asn	Gln	Thr	Trp	Phe	Tyr	Phe	Pro	Ile	Leu	Ser	Phe	Ala	
				260				265						270			
15	cgc	ctc	tgc	tgg	tgc	ctc	cag	tcc	atc	atg	ttc	gtc	ctc	ccc	aac	ggc	864
	Arg	Leu	Ser	Trp	Cys	Leu	Gln	Ser	Ile	Met	Phe	Val	Leu	Pro	Asn	Gly	
			275					280					285				
	cag	gcg	cac	aag	ccc	tgc	ggc	gcg	cgc	gtc	ccc	atc	tgc	ctc	gtc	gag	912
20	Gln	Ala	His	Lys	Pro	Ser	Gly	Ala	Arg	Val	Pro	Ile	Ser	Leu	Val	Glu	
	290						295					300					
	cag	ctc	tgc	ctc	gcg	atg	cac	tgg	acc	tgg	tac	ttg	gcg	acc	atg	ttc	960
	Gln	Leu	Ser	Leu	Ala	Met	His	Trp	Thr	Trp	Tyr	Leu	Ala	Thr	Met	Phe	
	305				310					315					320		
25	ctc	ttc	atc	aag	gac	ccc	gtc	aac	atg	atc	gtc	tac	ttc	ctc	gtc	tgc	1008
	Leu	Phe	Ile	Lys	Asp	Pro	Val	Asn	Met	Ile	Val	Tyr	Phe	Leu	Val	Ser	
				325					330						335		
	cag	gcg	gtc	tgc	ggc	aac	ctc	ctc	gcg	atc	gtc	ttc	tcc	ctc	aac	cac	1056
30	Gln	Ala	Val	Cys	Gly	Asn	Leu	Leu	Ala	Ile	Val	Phe	Ser	Leu	Asn	His	
			340					345					350				
	aac	gga	atg	ccc	gtc	atc	tgc	aag	gag	gag	gcg	gtg	gac	atg	gac	ttc	1104
	Asn	Gly	Met	Pro	Val	Ile	Ser	Lys	Glu	Glu	Ala	Val	Asp	Met	Asp	Phe	
		355						360					365				
35	ttc	acc	aag	cag	atc	atc	acc	ggc	cgc	gac	gtg	cac	ccc	ggc	ctc	ttc	1152
	Phe	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Arg	Asp	Val	His	Pro	Gly	Leu	Phe	
		370					375					380					
	gcg	gac	tgg	ttc	acc	ggc	ggc	ctc	aac	tac	cag	atc	gag	cac	cac	ctc	1200
40	Ala	Asp	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Tyr	Gln	Ile	Glu	His	His	Leu	
	385					390					395					400	
	ttc	ccc	tgc	atg	ccc	cgc	cac	aac	ttc	tgc	aag	atc	cag	ccc	gcc	gtc	1248
45	Phe	Pro	Ser	Met	Pro	Arg	His	Asn	Phe	Ser	Lys	Ile	Gln	Pro	Ala	Val	
					405					410					415		
	gag	acg	ctc	tgc	aag	aag	tac	ggc	gtc	cgc	tac	cac	acc	acc	ggc	atg	1296
	Glu	Thr	Leu	Cys	Lys	Lys	Tyr	Gly	Val	Arg	Tyr	His	Thr	Thr	Gly	Met	
			420					425					430				
50	atc	gag	ggc	acc	gcc	gag	gtc	ttc	tgc	cgc	ctt	aac	gag	gtg	tcc	aag	1344
	Ile	Glu	Gly	Thr	Ala	Glu	Val	Phe	Ser	Arg	Leu	Asn	Glu	Val	Ser	Lys	
		435					440					445					
	gcg	gcg	tgc	aag	atg	ggc	aag	gcg	cag	tag							1374
55	Ala	Ala	Ser	Lys	Met	Gly	Lys	Ala	Gln								
	450						455										
	<210> 98																
	<211> 457																
	<212> PRT																
	<213> Mortierella alpina																
60	<400> 98																
65																	

ES 2 724 998 T3

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

Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu Met
 225 230 235 240
 5
 Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser Arg Phe
 245 250 255
 10
 Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala
 260 265 270
 15
 Arg Leu Ser Trp Cys Leu Gln Ser Ile Met Phe Val Leu Pro Asn Gly
 275 280 285
 20
 Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu Val Glu
 290 295 300
 25
 Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe
 305 310 315 320
 30
 Leu Phe Ile Lys Asp Pro Val Asn Met Ile Val Tyr Phe Leu Val Ser
 325 330 335
 35
 Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His
 340 345 350
 40
 Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe
 355 360 365
 45
 Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe
 370 375 380
 50
 Ala Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu
 385 390 395 400
 55
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val
 405 410 415
 60
 Glu Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met
 420 425 430
 65
 Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys
 435 440 445
 Ala Ala Ser Lys Met Gly Lys Ala Gln
 450 455
 <210> 99
 <211> 1408
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 99

	atgacctcgt acgcgcgcca tcctcgcgcg tcgtccttcc tcgcctcgtt cgcagacggg	60
5	cccaagccac caacaccgac aggcaccccc gcgccccctg cttogacgta cgacttggtc	120
	ttgaaccccc tcacgccccct cgttttcgga ctgctctact tcggacggc caaaaccctc	180
	tcccacttcc agaacggcaa gaaccgcac aagggcaagg gctgggacgt cgtgtctctc	240
10	gtgcacaata tcctccttgc ggtctactcg gcttgacgt tcctcggaac cgcgcgcgag	300
	atcttcggcg ctttcgttcg cggctacatg ggggacgggt tcgcgggatt gacccaacgc	360
	tgagttccgc gcgcttttac attcgagcgt catctctttg caggaggggc gtogaattct	420
15	gaaagacgag cggcgctaac tcgctcgcca actccgacag cttctgcgac tcgtcgttcg	480
	cgatctggca gcagacgacg ttccccaaat tcgcctacct cttctacgtc agcaagttct	540
	acgtgcgtct ctctcgtctc catctcatag cagtcactca ctgggctcgc tcgcgcagga	600
20	aatcgtcgac acggccatct tgctgttgaa gggcaagaag gtcggcatgc tcagtcgta	660
	ccaccacatg ggcgcgatct ggactatgta cgcgcgtac gctacgcaag ccattgcctgt	720
	cgtgcgtccc gcttcctccc tcgctcgaaa acggagaact ctgctgacg ctgcactccc	780
25	gcgcgtgcag tggctcttcg tcgtcttcaa ctgcttcac cactcaatca tgtacaccta	840
	ctacgccttc tcgaccgtct cgtccccctt cccgcgcttc ctcaagaaat ccctcacgcg	900
30	actccaaatc acccagttcc tcgtcggcgg ttgcctcgcc gcctcgtacc tcttcattaa	960
	gcttcccgaa ctcccttcgg cggaggagat gtccgctcgc gcgacttoga gcttcgaggc	1020
	ggcgctcggc gcgctcaagc gcgagggtcc gacgtgcctc gtcaatgcgg cgcagaggca	1080
35	tgcgacgttg ctcaattgcg cgtacctcgt gccgttgacg tacctctttg tcgcattctt	1140
	cttcaagacg taaccagaaga actcggcggc taacgcgcgc gcaaaggcca aggccaacgc	1200
	gaagaaggcc aactagcgcg gtcccttcct ccctctcttt cgtcctttc gcatccgttt	1260
40	caactgtcgt gatttcgcac cctcttcgct tgcatacctc tccttagccc tcgtcgtttc	1320
	gtcgcagcta gatctcctct cctctctttc tcgtgtcgt attcaccagc tctcttctct	1380
	tgggtctctc gtcgtaatcc atagcctc	1408
45	<210> 100	
	<211> 990	
	<212> ADN	
	<213> Rhodosporidium toruloides	
	<220>	
50	<221> CDS	
	<222> (1)..(990)	
	<400> 100	
55		
60		
65		

	atg acc tcg tac gcc gcc cat cct cgc gcg tcg tcc ttc ctc gcc tcg	48
	Met Thr Ser Tyr Ala Ala His Pro Arg Ala Ser Ser Phe Leu Ala Ser	
	1 5 10 15	
5	ttc gca gac ggt ccc aag cca cca aca cgg aca ggc atc ccc gcg ccc	96
	Phe Ala Asp Gly Pro Lys Pro Pro Thr Pro Thr Gly Ile Pro Ala Pro	
	20 25 30	
10	ctc gct tcg acg tac gac ttg ttc ttg aac ccc gtc acg ccc ctc gct	144
	Leu Ala Ser Thr Tyr Asp Leu Phe Leu Asn Pro Val Thr Pro Leu Ala	
	35 40 45	
15	ttc gga ctc gtc tac ttc gcg acg gcc aaa acc ctc tcc cac ttc cag	192
	Phe Gly Leu Val Tyr Phe Ala Thr Ala Lys Thr Leu Ser His Phe Gln	
	50 55 60	
20	aac ggc aag aac cgc atc aag ggc aag ggc tgg gac gtc gct gtc ctc	240
	Asn Gly Lys Asn Arg Ile Lys Gly Lys Gly Trp Asp Val Ala Val Leu	
	65 70 75 80	
25	gtg cac aat atc ctc ctt gcg gtc tac tcg gcg tgg acg ttc ctc gga	288
	Val His Asn Ile Leu Ala Val Tyr Ser Ala Trp Thr Phe Leu Gly	
	85 90 95	
30	acc gcg ccg cag atc ttc gcc gct ttc gtt cgc gcc tac atg gcg gac	336
	Thr Ala Pro Gln Ile Phe Gly Ala Phe Val Arg Gly Tyr Met Ala Asp	
	100 105 110	
35	ggg ttc gcc gga ttg acc cac gcc ttc tgc gac tcg tcg ttc gcg atc	384
	Gly Phe Ala Gly Leu Thr His Ala Phe Cys Asp Ser Ser Phe Ala Ile	
	115 120 125	
40	tgg cag cag acg acg ttc ccc aag ttc gcc tac ctc ttc tac gtc agc	432
	Trp Gln Gln Thr Thr Phe Pro Lys Phe Ala Tyr Leu Phe Tyr Val Ser	
	130 135 140	
45	aag ttc tac gaa atc gtc gac acg gcc atc ttg ctg ttg aag ggc aag	480
	Lys Phe Tyr Glu Ile Val Asp Thr Ala Ile Leu Leu Lys Gly Lys	
	145 150 155 160	
50	aag gtc gcc atg ctc cag tcg tac cac cac atg gcc gcg atc tgg act	528
	Lys Val Gly Met Leu Gln Ser Tyr His His Met Gly Ala Ile Trp Thr	
	165 170 175	
55	atg tac gcc gcg tac gct acg caa gcc atg cct gtc tgg ctc ttc gtc	576
	Met Tyr Ala Ala Tyr Ala Thr Gln Ala Met Pro Val Trp Leu Phe Val	
	180 185 190	
60	gtc ttc aac tcg ttc atc cac tca atc atg tac acc tac tac gcc ttc	624
	Val Phe Asn Ser Phe Ile His Ser Ile Met Tyr Thr Tyr Tyr Ala Phe	
	195 200 205	
65	tcg acc gtc tcg ctc ccc ttc ccg cgc ttc ctc aag aaa tcc ctc acg	672
	Ser Thr Val Ser Leu Pro Phe Pro Arg Phe Leu Lys Lys Ser Leu Thr	
	210 215 220	
70	cga ctc caa atc acc cag ttc ctc gtc gcc ggt tcg ctc gcc gcc tcg	720
	Arg Leu Gln Ile Thr Gln Phe Leu Val Gly Gly Ser Leu Ala Ala Ser	
	225 230 235 240	
75	tac ctc ttc att aag ctt ccc gaa ctc cct tcg gcg gag gag atg tcc	768
	Tyr Leu Phe Ile Lys Leu Pro Glu Leu Pro Ser Ala Glu Glu Met Ser	

ES 2 724 998 T3

				245				250					255																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
--	--	--	--	-----	--	--	--	-----	--	--	--	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

ES 2 724 998 T3

	130	135	140	
5	Lys Phe Tyr Glu Ile Val Asp Thr Ala Ile Leu Leu Leu Lys Gly Lys 145 150 155 160			
10	Lys Val Gly Met Leu Gln Ser Tyr His His Met Gly Ala Ile Trp Thr 165 170 175			
15	Met Tyr Ala Ala Tyr Ala Thr Gln Ala Met Pro Val Trp Leu Phe Val 180 185 190			
20	Val Phe Asn Ser Phe Ile His Ser Ile Met Tyr Thr Tyr Tyr Ala Phe 195 200 205			
25	Ser Thr Val Ser Leu Pro Phe Pro Arg Phe Leu Lys Lys Ser Leu Thr 210 215 220			
30	Arg Leu Gln Ile Thr Gln Phe Leu Val Gly Gly Ser Leu Ala Ala Ser 225 230 235 240			
35	Tyr Leu Phe Ile Lys Leu Pro Glu Leu Pro Ser Ala Glu Glu Met Ser 245 250 255			
40	Ala Ala Ala Thr Ser Ser Phe Glu Ala Gly Val Gly Ala Leu Lys Arg 260 265 270			
45	Glu Gly Pro Thr Cys Leu Val Asn Ala Ala Gln Arg His Ala Thr Leu 275 280 285			
50	Leu Asn Cys Ala Tyr Leu Val Pro Leu Thr Tyr Leu Phe Val Ala Phe 290 295 300			
55	Phe Phe Lys Thr Tyr Gln Lys Asn Ser Ala Ala Asn Ala Ala Ala Lys 305 310 315 320			
60	Ala Lys Ala Asn Ala Lys Lys Ala Asn 325			
65	<210> 102 <211> 1620 <212> ADN <213> Rhodosporidium toruloides <400> 102			
	ttagttgggc ttcttggcga cggtcgcggc ggccgccttt gcggcgctga ccttggtctg	60		
	gccggccttc ttgtacgtgc gttggtaaaa tgcgataaag aggaagaggt aactggctgt	120		
	tttcgccagc cctcggggcg tcagccttgc tcgtagaaaa aaccgagagg aaaggggaaa	180		

	atggacgacg aggaggaggg gggaaggact cactgaggag tgcgcagccg aagatcgag	240
5	ctccctccgt cccagcgac gacccgaggg tcgggagcga cccgcgggcg tacgtccagg	300
	cgaagtagct gtagctggcg aagtagaccg tgaagaggtc gatcacgaac tgggtgattt	360
	ggagcgctgt gaggtacctt taccacgggg gctcgtcagg tctagcacia agggagagga	420
10	ggggagagac ggacttcttc caccagatct tctagccggc ggccgtcata aagtagtagt	480
	agtactttac gacggccac aagcggcgct agcttcggcg tcgactccag gatgacagga	540
	ctcgccgcaa cgcacatga ggacatgcac aaagaggttg aggggtgatga cgaaccacga	600
15	gacggaggtg cgtccattga gctgctgta gcagaggacg gcggtcgcg tgtggtggaa	660
	gacgtggagg aactgaaggg gcttcttctt gacgacgagg aagaccgtgt cgaacaactc	720
	ccagtcctgc cgcgaggttg cgttaggttg cgcaagcagg ggggcgtgag aaggtcggag	780
20	agggcgtaga cttgaactgc gcggcgtagt aacgtcagtt tcggcatctt cgagggacag	840
	agggacaaca cagtagtag ttgaagatgt agtaggtctc gagggggggc gtccaggcct	900
	cgttgccgca gatggcggtg aagaggccgt gcttcagat gatggggacg atctgggccc	960
25	gggagagcgt aagcgacgtc ctccaaagaa tcagatcgtc gaacgcacct cctcaagcat	1020
	cagggcgaga agaaggccgg agccggtcga gaggaggaag ttgtgcagca tgaaaagggg	1080
	cttgaagcct gtgcgtcgca ggacgaggtc caagtcagtt tctgtagact ctgcaagctc	1140
30	aagactttgc tgggtgaagct tctgtgcgac aggcgagccc atccgctcga cgggtgcatg	1200
	acctcgagcg gttgaagcgt gatcggaacg cgttctcttc gtctatccgg ccttttgctg	1260
	cgggtccgag gctcttcagc ctgccagcgt cccattgacg cgtctgcata tctgatccga	1320
35	cgcgcgcgac gcgttgctcc tgcaagagcc ccaacttctt gtogaagcaa accccaaact	1380
	cacggtacgg ctttgatccg cgcagtagcg cctgacctcc gaagatgacc gcgaggtaga	1440
40	cgcgaacggc aatgacgacc tcgggcgtgg tcgagagggg tgtttggccc gggatccagt	1500
	gctcgatggt gcgcgggagg gaggggatgg ggagcgctg caggagcgag tagacggggc	1560
	cgggggacgg tgcgaccata gtgaatggac ctgagcgggt cgagagcgta gcctgggtgc	1620
45	<210> 103	
	<211> 879	
	<212> ADN	
	<213> Rhodosporidium toruloides	
	<220>	
	<221> CDS	
	<222> (1)..(879)	
50	<400> 103	
	atg gtc gca ccg tcc ccc gcc ccc gtc tac tgg ctc ctg cac gcg ctc	48
	Met Val Ala Pro Ser Pro Gly Pro Val Tyr Ser Leu Leu His Ala Leu	
	1 5 10 15	
55		
60		
65		

	ccc atc ccc tcc ctc ccg cgc acc atc gag cac tgg atc ccg ggc caa	96
	Pro Ile Pro Ser Leu Pro Arg Thr Ile Glu His Trp Ile Pro Gly Gln	
	20 25 30	
5	aca ccc ctc tcg acc acg ccc gag gtc gtc att gcc gtt ggc gtc tac	144
	Thr Pro Leu Ser Thr Thr Pro Glu Val Val Ile Ala Val Ala Val Tyr	
	35 40 45	
10	ctc gcg gtc atc ttc gga ggt cag gcg ctc atg cgc gga tca aag ccg	192
	Leu Ala Val Ile Phe Gly Gly Gln Ala Leu Met Arg Gly Ser Lys Pro	
	50 55 60	
15	tac cgc ttc aag ccc ctt ttc atg ctg cac aac ttc ctc ctc tcg acc	240
	Tyr Arg Phe Lys Pro Leu Phe Met Leu His Asn Phe Leu Leu Ser Thr	
	65 70 75 80	
20	ggc tcc ggc ctt ctt ctc gcc ctg atg ctt gag gag atc gtc ccc atc	288
	Gly Ser Gly Leu Leu Ala Leu Met Leu Glu Glu Ile Val Pro Ile	
	85 90 95	
25	atc tgg aag cac ggc ctc ttc cac gcc atc tgc gcc aac gag gcc tgg	336
	Ile Trp Lys His Gly Leu Phe His Ala Ile Cys Ala Asn Glu Ala Trp	
	100 105 110	
30	acg ccc cgc ctc gag acc tac tac atc ttc aac tac tac ttc aag tac	384
	Thr Pro Arg Leu Glu Thr Tyr Tyr Ile Phe Asn Tyr Tyr Phe Lys Tyr	
	115 120 125	
35	tgg gag ttg ttc gac acg gtc ttc ctc gtc gtc aag aag aag ccc ctt	432
	Trp Glu Leu Phe Asp Thr Val Phe Leu Val Val Lys Lys Lys Pro Leu	
	130 135 140	
40	cag ttc ctc cac gtc ttc cac cac acc gcg acc gcc gtc ctc tgc tac	480
	Gln Phe Leu His Val Phe His His Thr Ala Thr Ala Val Leu Cys Tyr	
	145 150 155 160	
45	acg cag ctc aat gga cgc acc tcc gtc tcg tgg gtc gtc atc acc ctc	528
	Thr Gln Leu Asn Gly Arg Thr Ser Val Ser Trp Val Val Ile Thr Leu	
	165 170 175	
50	aac ctc ttt gtg cat gtc ctc atg tac tac tac tac ttt atg acg gcc	576
	Asn Leu Phe Val His Val Leu Met Tyr Tyr Tyr Tyr Phe Met Thr Ala	
	180 185 190	
55	gcc gcc tac aag atc tgg tgg aag aag tac ctc acg acg ctc caa atc	624
	Ala Gly Tyr Lys Ile Trp Trp Lys Lys Tyr Leu Thr Thr Leu Gln Ile	
	195 200 205	
60	acc cag ttc gtg atc gac ctc ttc acg gtc tac ttc gcc agc tac agc	672
	Thr Gln Phe Val Ile Asp Leu Phe Thr Val Tyr Phe Ala Ser Tyr Ser	
	210 215 220	
65	tac ttc gcc tgg acg tac gcc cgc ggg tcg ctc ccg acc ctc ggg tcg	720
	Tyr Phe Ala Trp Thr Tyr Ala Arg Gly Ser Leu Pro Thr Leu Gly Ser	
	225 230 235 240	
70	tgc gct ggg acg gag gga gct gcg atc ttc gcc tgc gca ctc ctc acc	768
	Cys Ala Gly Thr Glu Gly Ala Ala Ile Phe Gly Cys Ala Leu Leu Thr	
	245 250 255	
75	agt tac ctc ttc ctc ttt atc gca ttt tac caa cgc acg tac aag aag	816
	Ser Tyr Leu Phe Leu Phe Ile Ala Phe Tyr Gln Arg Thr Tyr Lys Lys	
	260 265 270	
80	gcc gcc cag aac aag gtc gac gcc gca aag gcc gcc gcc gcg acc gtc	864
	Ala Gly Gln Asn Lys Val Asp Ala Ala Lys Gly Ala Ala Thr Val	
	275 280 285	
85	gcc aag aag acc aac	879
	Ala Lys Lys Thr Asn	
	290	
	<210> 104	
	<211> 293	

<212> PRT

<213> *Rhodospiridium toruloides*

<400> 104

```

5 Met Val Ala Pro Ser Pro Gly Pro Val Tyr Ser Leu Leu His Ala Leu
  1          5          10          15

10 Pro Ile Pro Ser Leu Pro Arg Thr Ile Glu His Trp Ile Pro Gly Gln
   20          25          30

15 Thr Pro Leu Ser Thr Thr Pro Glu Val Val Ile Ala Val Ala Val Tyr
   35          40          45

20 Leu Ala Val Ile Phe Gly Gly Gln Ala Leu Met Arg Gly Ser Lys Pro
   50          55          60

25 Tyr Arg Phe Lys Pro Leu Phe Met Leu His Asn Phe Leu Leu Ser Thr
   65          70          75          80

30 Gly Ser Gly Leu Leu Leu Ala Leu Met Leu Glu Glu Ile Val Pro Ile
   85          90          95

35 Ile Trp Lys His Gly Leu Phe His Ala Ile Cys Ala Asn Glu Ala Trp
   100         105         110

40 Thr Pro Arg Leu Glu Thr Tyr Tyr Ile Phe Asn Tyr Tyr Phe Lys Tyr
   115         120         125

45 Trp Glu Leu Phe Asp Thr Val Phe Leu Val Val Lys Lys Lys Pro Leu
   130         135         140

50 Gln Phe Leu His Val Phe His His Thr Ala Thr Ala Val Leu Cys Tyr
   145         150         155         160

55 Thr Gln Leu Asn Gly Arg Thr Ser Val Ser Trp Val Val Ile Thr Leu
   165         170         175

60 Asn Leu Phe Val His Val Leu Met Tyr Tyr Tyr Tyr Phe Met Thr Ala

```

ES 2 724 998 T3

	180	185	190
5	Ala Gly Tyr Lys Ile Trp Trp	Lys Lys Tyr Leu Thr	Thr Leu Gln Ile
	195	200	205
10	Thr Gln Phe Val Ile Asp	Leu Phe Thr Val Tyr	Phe Ala Ser Tyr Ser
	210	215	220
15	Tyr Phe Ala Trp Thr Tyr	Ala Arg Gly Ser Leu Pro	Thr Leu Gly Ser
	225	230	235 240
20	Cys Ala Gly Thr Glu Gly	Ala Ala Ile Phe Gly Cys	Ala Leu Leu Thr
	245	250	255
25	Ser Tyr Leu Phe Leu Phe	Ile Ala Phe Tyr Gln Arg	Thr Tyr Lys Lys
	260	265	270
30	Ala Gly Gln Asn Lys Val	Asp Ala Ala Lys Gly Ala	Ala Ala Thr Val
	275	280	285
35	Ala Lys Lys Thr Asn		
	290		
40	<210> 105		
	<211> 527		
45	<212> PRT		
	<213> Agaricus bisporus		
50	<400> 105		
	Met Val Asp Leu Lys Tyr Thr Pro Val	Glu Glu Ile Glu Lys Ile Gln	
	1	5	10 15
55	Ala Thr Leu Arg Asn Gly Phe Arg	Ser Gly Arg Thr Lys Asn Ile Glu	
	20	25	30
60	Tyr Arg Lys Tyr Gln Leu Leu Gln	Leu Ala Tyr Met Leu Gln Asp Asn	
	35	40	45
65	Val Lys Arg Leu Glu Glu Ala Leu	Ala Ala Asp Leu Gly Arg Pro Pro	
	50	55	60
70	Leu Glu Ser Gln Phe Leu Glu Ile	Gly Pro Ser Met Met Asp Ala Arg	
	65	70	75 80
75	Asn Ala Trp Ala Gly Val Asp Lys	Trp Ala Lys Thr Glu Arg Ala Pro	
	85	90	95
80	Phe Ser Ile Asn Gly Phe Ala Met	Arg Pro Val Ile Tyr Lys Glu Pro	

ES 2 724 998 T3

	100	105	110
5	Lys Gly Val Val Leu Ile Ile Ser Pro Phe Asn Tyr Pro Val Trp Leu		
	115	120	125
10	Cys Met Ser Pro Leu Ala Gly Ala Ile Ala Ala Gly Asn Ala Val Leu		
	130	135	140
15	Leu Lys Pro Ser Glu Ser Thr Pro His Val Ser Ser Leu Phe Ala Glu		
	145	150	155
20	Gly Val Pro Glu Thr Thr Lys Leu Leu Asp Leu Pro Trp Asp His Ile		
	180	185	190
25	Leu Tyr Thr Gly Ser Gly Gln Val Gly Arg Ile Val Ser Ala Ala Ala		
	195	200	205
30	Ala Lys His Leu Thr Pro Val Ser Leu Glu Leu Gly Gly Lys Ser Pro		
	210	215	220
35	Val Phe Ile Asp Pro Asn Cys Asp Ile Glu Leu Ala Ala Lys Arg Ile		
	225	230	235
40	Leu Trp Gly Lys Cys Val Asn Ala Gly Gln Thr Cys Thr Ala Pro Asp		
	245	250	255
45	Tyr Val Leu Val Pro Arg Glu Val Gln Asp Lys Phe Val Asn Ala Leu		
	260	265	270
50	Lys Asn Ser Met Asp Asn Phe Tyr Pro Glu Ser Val Ala Thr Pro Gly		
	275	280	285
55	Val Phe Ser Arg Leu Val Thr Pro Gln Ala Phe Asn Arg Ile Lys Gly		
	290	295	300
60	Leu Leu Asp Asn Thr Lys Gly Thr Ile Val Ile Gly Gly Glu Met Asp		
	305	310	315
65	Glu Ala Thr Lys Phe Ile Ala Pro Thr Ile Val Lys Asp Val Pro Thr		
	325	330	335
70	Asn Asp Ser Leu Met Asn Glu Glu Ile Phe Gly Pro Val Leu Pro Ile		
	340	345	350

Val Pro Val Lys Asp Val Glu Glu Ala Ile Ala Tyr Val Asn Ser Asn
355 360 365

Asp His Pro Leu Ala Val Tyr Val Phe Ser Gln Asp Ala Ala Tyr Lys
370 375 380

Gln Lys Val Phe Ser Arg Thr Gln Ser Gly Ser Ala Val Ala Asn Glu
385 390 395 400

Val Val Ile Gln Pro Gly Ile Glu Gly Leu Pro Phe Gly Gly Ile Gly
405 410 415

Pro Ser Gly Ser Gly Tyr His Thr Gly Lys Tyr Thr Phe Asp Met Phe
420 425 430

Thr His Leu Arg Ala Ser Leu Asp Ser Pro Gly Trp Leu Asp Lys Ile
435 440 445

Leu Gly Phe Arg Phe Pro Pro Tyr Thr Asp Lys Ser Ile Lys Ala Ser
450 455 460

Gln Arg Ile Leu Lys Ser Leu Pro Pro Arg Pro Thr Gly Pro Pro Arg
465 470 475 480

Thr Asn Asn Ala Met Ala Asn Gly Ser Ala Thr Lys Trp Trp Gly Lys
485 490 495

Tyr Phe Phe Leu Ala Phe Val Leu Ala Thr Ile Gly Gly Leu Thr Lys
500 505 510

Pro Val Lys Ile Leu Gly Arg Lys Phe Val Pro Lys Ile Leu Gly
515 520 525

<210> 106

<211> 487

<212> PRT

<213> Phanerochaete camosa

<400> 106

Met Thr Arg Leu Glu Tyr Thr Asn Ile Asp Asn Val Pro Lys Ile His
1 5 10 15

Ala Glu Leu Lys Glu Thr Phe Arg Ser Gly Lys Thr Arg Pro Val Pro
20 25 30

Phe Arg Lys Glu Gln Leu Ala Gln Leu Ala Trp Leu Leu Lys Asp Asn
35 40 45

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

Arg Leu Ile Asp Gly Thr Lys Gly Thr Val Val Val Gly Gly Thr Val
 305 310 315 320
 5
 Asp Val Glu Lys Arg Tyr Ile Ala Pro Thr Ile Ile Arg Asp Val Pro
 325 330 335
 10
 Val Asp Asp Ser Thr Met Asp Glu Glu Ile Phe Gly Pro Val Leu Pro
 340 345 350
 15
 Ile Val Pro Val Arg Asp Val Glu Glu Ala Ile Gln Ile Val Asn Ser
 355 360 365
 Arg Asp His Pro Leu Ser Leu Tyr Val Phe Thr Gln Asp Ala Ala Phe
 370 375 380
 20
 Lys Glu Asn Val Phe Ser Arg Thr Gln Ser Gly Ala Ala Leu Ala Asn
 385 390 395 400
 25
 Glu Val Leu Val His Val Gly Ala Thr Gly Leu Pro Phe Gly Gly Ile
 405 410 415
 Gly Pro Ser Gly Ser Gly Ser Leu Thr Gly Lys His Gly Phe Asp Ala
 420 425 430
 30
 Phe Thr His Leu Arg Ser Thr Leu Asp Asn Pro Lys Trp Val Asp Ala
 435 440 445
 35
 Ile Met Lys Gly Arg Tyr Pro Pro Tyr Thr Pro Glu Lys Leu Ala Arg
 450 455 460
 40
 Leu Arg Thr Ala Leu Lys Val Arg Met Pro Pro Arg Pro Gly Gly Ala
 465 470 475 480
 Arg Pro Ala Gln Lys Ser Ala
 485
 <210> 107
 <211> 514
 <212> PRT
 <213> malassezia globosa
 <400> 107
 Met Lys Thr Thr Pro Ala Asp Thr Ile Pro Ala Leu Val Gly Arg Leu
 1 5 10 15
 50
 Arg Ala Thr Phe Leu Thr Gly Lys Thr Arg Cys Val Glu Tyr Arg Lys
 20 25 30
 55
 60
 65

ES 2 724 998 T3

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

Arg Ile Val Ser Asp Asn His Trp Lys Arg Ile His Ser Leu Val Ser
 290 295 300
 5 Ser Ser Lys Ala Glu Leu Val Val Gly Gly Thr Ala Gly Ala Asp Gln
 305 310 315 320
 10 Ala Lys Arg Phe Ile Pro Leu Thr Ile Leu Lys Asn Val Asp Ala Ser
 325 330 335
 15 Asp Ser Val Met Thr Asp Glu Ile Phe Gly Pro Val Leu Pro Ile Val
 340 345 350
 20 Pro Phe Asp Thr Ile Arg Asp Ala Val Asp Phe Val Asn Glu Arg Asp
 355 360 365
 25 Gln Pro Leu Ala Leu Tyr Val Phe Thr Ser Cys Asn Glu Thr Arg Asp
 370 375 380
 30 Tyr Ile Leu Ala Tyr Thr Arg Ser Gly Gly Val Val Arg Gly Asp Cys
 385 390 395 400
 35 Leu Leu His Tyr Ala Ile Asp Ser Leu Pro Phe Gly Gly Thr Gly Pro
 405 410 415
 40 Ser Gly Tyr Gly Ser Tyr His Gly Lys Ala Gly Phe Asp Cys Phe Thr
 420 425 430
 45 His Glu Arg Ala Val Val Asp Ala Pro Ser Tyr Gly Met Leu Gly Lys
 435 440 445
 50 Leu Val Glu Val Val Met Ala Arg Arg Tyr Pro Pro Tyr Ser Lys Ser
 450 455 460
 55 Lys Leu Asp Phe Phe Arg Phe Val Leu Pro Lys Leu Val Trp Phe Gly
 465 470 475 480
 60 Arg Pro Pro Gln Pro Thr Arg Ser Ser Lys Ser Ile Asp His Pro Pro
 485 490 495
 65 Ser Lys Val Arg Thr Thr Ala Pro Arg His Gly Gly His Pro Leu Ser
 500 505 510
 Leu Val
 <210> 108
 <211> 538
 <212> PRT
 <213> Sporisorium reilianum
 <400> 108

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

ES 2 724 998 T3

	225		230		235		240									
5	Asp	Leu	Lys	Ile	Ala	Ala	His	Arg	Ile	Leu	Trp	Gly	Lys	Ser	Phe	Asn
				245					250						255	
10	Cys	Gly	Gln	Thr	Cys	Ile	Ala	Pro	Asp	Tyr	Val	Leu	Ile	Pro	His	Glu
			260						265					270		
15	Leu	Gln	His	Lys	Phe	Val	His	Glu	Leu	His	Lys	Ala	Tyr	Glu	Arg	Phe
		275						280					285			
20	Tyr	Pro	Glu	Ile	Lys	Gly	Gly	Val	Ser	Gln	Ser	Glu	Ser	Tyr	Ala	Arg
	290						295					300				
25	Ile	Ile	Asn	Pro	Gly	His	Trp	Lys	Arg	Leu	Thr	Ala	Met	Leu	Ser	Gly
	305					310					315					320
30	Thr	Lys	Gly	Lys	Ile	Val	Leu	Gly	Gly	Glu	Gly	Asp	Glu	Ala	Thr	Lys
				325					330						335	
35	Phe	Leu	Pro	Pro	Thr	Val	Ile	Ala	Asn	Val	Lys	Thr	Asp	Asp	Pro	Val
			340						345					350		
40	Met	Ala	Gly	Glu	Ile	Phe	Gly	Pro	Leu	Leu	Pro	Ile	Val	Pro	Val	Lys
		355						360					365			
45	Asp	Val	Lys	Ala	Ala	Val	Asp	Phe	Ile	Asn	Ser	Arg	Asp	Gln	Pro	Leu
	370						375					380				
50	Ala	Leu	Tyr	Leu	Phe	Ala	Gly	Asp	Ser	Lys	Val	Lys	Glu	Tyr	Phe	Phe
	385				390						395					400
55	Asp	Asn	Thr	Arg	Ser	Gly	Ala	Cys	Val	Gln	Gly	Asp	Thr	Leu	Leu	His
				405						410					415	
60	Phe	Ala	Val	Asp	Ala	Leu	Pro	Phe	Gly	Gly	Thr	Gly	Pro	Ser	Gly	Tyr
			420						425					430		
65	Gly	Asn	Tyr	His	Gly	Lys	Ala	Ser	Phe	Asp	Gln	Phe	Ser	His	Gln	Arg
		435						440					445			
70	Ala	Ser	Leu	Asp	Ala	Pro	Ser	Thr	Gly	Leu	Leu	Gly	Lys	Ile	Ile	Glu
	450					455						460				
75	Ile	Val	Met	Ser	Ser	Arg	Tyr	Pro	Pro	Tyr	Thr	Asn	Ala	Asn	Leu	Asn
	465					470					475					480

Lys Met Arg Met Leu Ala Glu Tyr Ser Val Ser Phe Lys Arg Pro Ser
 485 490 495
 5
 Asn Pro His Lys Ser Thr Thr Ser Ser Ser Ser Gly Gln Gly Ala Val
 500 505 510
 10
 Ala Lys Arg Leu Ala Val Val Leu Leu Leu Ile Ser Leu Val Leu Gly
 515 520 525
 15
 Ala Arg Asn Arg Gly Leu Ile Gly Trp Ile
 530 535
 <210> 109
 <211> 640
 <212> PRT
 <213> Ustilago maydis
 <400> 109
 20
 Met Arg Glu Trp Leu Gly Gly Asn Leu Arg Phe Ala Ala Ile Phe Ile
 1 5 10 15
 25
 Arg Ile Gln Ser Arg Pro Asp His Ser Leu Arg Phe Thr Val Leu Asp
 20 25 30
 30
 Pro Leu Pro Val His Tyr Arg Leu Ser Pro Pro Ser Ser Thr Leu Ala
 35 40 45
 35
 Arg Leu His Gln Pro Arg Leu Ala Leu Thr Ser Phe Ala His Leu Ala
 50 55 60
 40
 Leu Leu Pro Ser Ser Ile Ser Pro Ser Pro Thr Thr Leu Arg Leu Val
 65 70 75 80
 45
 Val Cys Leu Val Thr Asn Ser His His Ser Val Ser Ser Ser Leu Lys
 85 90 95
 50
 Arg Gln Ser Pro Ile Met Ala Ala Ala Ala Ala Thr Ala Ala Thr Glu
 100 105 110
 55
 Ala Gly Leu Gln Tyr Thr Pro Ile Asp Asp Ile Pro Ser Ile Val Ser
 115 120 125
 60
 Asp Leu Arg Ala Ala Phe Leu Thr Gly Lys Thr Arg Ser Val Glu Tyr
 130 135 140
 65
 Arg Lys Asn Gln Leu Lys Gln Leu Ala Tyr Met Ile Lys Asp Asn Gln
 145 150 155 160

5
10
15
20
25
30
35
40
45
50
55
60
65

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

Glu Ser Ile Phe Ala Glu Leu Met Gly Thr Thr Asn Glu Ile Val Glu
180 185 190

Ala Val Thr Lys Ile Asp Lys Trp Ala Lys Pro Ala Lys Pro Trp Ala
195 200 205

Gly Ala Ala Trp Ala Met His Gly Ala Thr Ile Arg Ser Glu Pro Lys
210 215 220

Gly Thr Val Leu Val Leu Gly Ala Trp Asn Tyr Pro Ile Thr Val Gln
225 230 235 240

Ile Gly Pro Val Ile Gly Ala Ile Ala Ala Gly Asn Thr Val Ile Leu
245 250 255

Lys Pro Ser Glu Val Ala Ser His Thr Ala Lys Leu Ile Ala Glu Leu
260 265 270

Trp Asn Lys Tyr Leu Asp Pro Glu Cys Phe Arg Val Val Asn Gly Gly
275 280 285

Ile Pro Glu Thr Thr Ala Leu Leu Asp Gln Arg Phe Glu His Ile Phe
290 295 300

Tyr Thr Gly Asn Gly Thr Val Gly Arg Ile Ile Ala Glu Lys Ala Ala
305 310 315 320

Lys Trp Leu Cys Pro Thr Thr Leu Glu Leu Gly Gly Lys Ser Pro Val
325 330 335

Tyr Val Asp Lys Ser Ala Asp Leu Ser Ile Ala Ala His Arg Ile Leu
340 345 350

Trp Gly Lys Ser Phe Asn Cys Gly Gln Thr Cys Ile Ala Pro Asp Tyr
355 360 365

Val Leu Ile Gln Pro Asp Leu Gln Asp Lys Phe Val Gln Glu Leu Lys
370 375 380

Lys Ala Tyr Gln Arg Phe Tyr Pro Glu Leu Gln Gly Gly Val Asn Asn
385 390 395 400

Ser Glu Ser Tyr Ala Arg Ile Ile Asn Pro Gly His Trp Lys Arg Leu
405 410 415

5 Asn Ala Met Leu Ser Gly Thr Lys Gly Lys Val Val Leu Gly Gly Glu
 420 425 430
 Gly Glu Glu Ala Thr Lys Phe Leu Pro Pro Thr Val Ile Ala Asp Val
 435 440 445
 10 Lys Pro Asp Asp Ala Ile Met Ser Gly Glu Ile Phe Gly Pro Leu Leu
 450 455 460
 15 Pro Ile Val Pro Val Arg Asp Val Glu Ala Ala Val Asp Leu Ile Asn
 465 470 475 480
 Ser Arg Asp Gln Pro Leu Ala Leu Tyr Leu Phe Ala Gly Asp Asn Arg
 485 490 495
 20 Val Lys Asn Phe Phe Phe Asp Asn Thr Arg Ser Gly Ala Cys Val Gln
 500 505 510
 25 Gly Asp Thr Leu Leu His Phe Ala Val Asp Val Leu Pro Phe Gly Gly
 515 520 525
 Thr Gly Pro Ser Gly Tyr Gly Asn Tyr His Gly Lys Ala Ser Phe Asp
 530 535 540
 30 Gln Phe Ser His Gln Arg Ala Ser Leu Asp Ala Pro Ser Thr Gly Leu
 545 550 555 560
 35 Leu Gly Lys Leu Val Glu Leu Ile Met Ser Ser Arg Tyr Pro Pro Tyr
 565 570 575
 40 Thr Glu Ala Asn Leu Lys Lys Leu Arg Ala Leu Ala Ala Tyr Ser Val
 580 585 590
 Ser Phe Lys Arg Pro Ser Asn Pro His Lys Ser Ile Ala Ser Ser Ser
 595 600 605
 45 Val Ser Leu Cys Leu Ser His Ser Arg Pro Ser Pro Phe Leu Ser Met
 610 615 620
 50 Ser Gln Ser Leu Phe Pro Met Val His Tyr Asn Met Leu Pro Thr Gln
 625 630 635 640
 <210> 110
 <211> 578
 <212> PRT
 <213> Melampsora larici-populina
 55 <400> 110

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

Ala Gly Gln Thr Cys Val Ala Pro Asp Tyr Ile Leu Val Ser Lys Lys
260 265 270

5 Ser Glu Ala Lys Leu Ile Thr Ser Leu Lys Lys Ala Met Gln Glu Leu
275 280 285

10 Tyr Pro Leu Asp Ser Ala Ser Gly Leu Ser Ala Lys Lys Ile Ile Asn
290 295 300

Ala Ser Gly Lys Ser Gly Asp Gly Pro Asp Asp Leu Gln Tyr Ser Lys
305 310 315 320

15 Ile Val Asn Gln Asn Gln Phe Asn Arg Leu Asn Asn Val Leu Gln Glu
325 330 335

20 Thr Lys Gly Glu Phe Ile Pro Thr Asp Asp Thr Phe Gln Ala Gly Thr
340 345 350

Lys Asp Ser Ser Asp Ala Gln Glu Leu Lys Met Pro Leu Thr Leu Val
355 360 365

25 Arg Asn Leu Thr Met Glu Asp Pro Val Met Gln Asn Glu Ile Phe Gly
370 375 380

30 Pro Ile Phe Pro Ile Leu Thr Tyr Asp Leu Gln Ser Glu Ser Met Ala
385 390 395 400

Glu Ile Leu Arg Pro Ile Ala Asp Ala Glu Pro Leu Ala Leu Tyr Val
405 410 415

35 Phe Thr Gln Ser Ser Gln Asn Phe Glu Leu Val Arg Gln His Thr Lys
420 425 430

40 Ser Gly Gln Ile Met Cys Asn Asp Leu Leu Ile Gln Phe Ala Ile Pro
435 440 445

45 Gly Leu Pro Phe Gly Gly Ile Gly Gln Ser Gly Ser Gly Asn Tyr His
450 455 460

Gly Tyr Tyr Ser Phe Leu Thr Phe Thr Tyr Glu Arg Ser Ser Ala Asn
465 470 475 480

50 Leu Pro Thr Trp Ala Asp Phe Leu Phe Asn Ala Arg Tyr Pro Pro Tyr
485 490 495

55 Thr Pro Phe Lys Leu Lys Leu Phe Ser Ala Ile Met Gly Pro Ala Arg

ES 2 724 998 T3

500 505 510

5 Ile Lys Gly Lys Ser Asn Pro Gly Leu Val Pro Lys Ser Ala Glu Val
515 520 525

10 Gly Lys Arg Ser Trp Leu Pro Thr Leu Thr Pro Leu Ser Phe Ser Thr
530 535 540

15 Leu Leu Leu Ala Gly Tyr Tyr Ala Leu Ser Arg Arg Tyr Gly Ser Asp
545 550 555 560

20 Tyr Leu Lys Ile Trp Met Thr Arg Phe Ile Gly Ala Ile Lys Gln Ser
565 570 575

25 Asn Arg
<210> 111
<211> 576
<212> PRT
<213> Puccinia graminis f. sp. tritici
<400> 111
Met Ser Ser Thr Ser Pro Leu Glu Phe Thr Pro Thr Glu Gln Ile Lys
1 5 10 15

30 Glu Ile Tyr Ala Thr Val Ser Gln Gly His Ser Ser Gly Ile Thr Lys
20 25 30

35 Ser Tyr Glu Trp Arg Glu His Gln Leu Lys Gln Leu Gly Tyr Leu Leu
35 40 45

40 Gln Glu Asn Glu Ser Leu Leu Glu Glu Ala Leu Thr Ile Asp Leu Gly
50 55 60

45 Arg Pro Asn Leu Glu Asn His Val Gly Glu Leu Val Gly Thr Arg Asn
65 70 75 80

50 Glu Val Leu Ser Ala Leu Lys Asn Leu Lys Lys Trp Ile Lys Pro Gln
85 90 95

55 Ser Val Lys Thr Glu Leu Thr Trp Leu Ile Ala Lys Pro Arg Val Ser
100 105 110

60 His Glu Pro Lys Gly Ile Val Ala Ile Phe Gly Ala Trp Asn Tyr Pro
115 120 125

65 Val Ala Val Leu Phe Gly Pro Leu Val Gly Ala Ile Ala Gly Gly Asn

ES 2 724 998 T3

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

Asp Ile Val Gln Cys Leu His Arg Ile Ser Gln Ser Ala Pro Leu Ala
 385 390 395 400
 5 Leu Tyr Ala Phe Gly Gln Ser Glu Glu Lys Leu Glu Phe Ile Arg Arg
 405 410 415
 10 Gln Thr Lys Ser Gly Gln Phe Val Cys Asn Asp Leu Leu Ile Gln Phe
 420 425 430
 Asn Ile Pro Gly Leu Pro Phe Gly Gly Val Gly Ala Ser Gly Leu Gly
 435 440 445
 15 Asn Tyr His Gly Tyr Tyr Ser Phe Leu Ala Phe Thr Tyr Glu Arg Pro
 450 455 460
 20 Met Val Asn Phe Pro Phe Trp Ala Asp Val Leu Leu Lys Ser Arg Tyr
 465 470 475 480
 Pro Pro Tyr Thr Ser Phe Lys Phe Lys Phe Met Gln Ala Val Leu Gly
 485 490 495
 25 Pro Ser Lys Leu Lys Gly Lys Ser Asn Pro Asn Pro Pro Ala Leu Thr
 500 505 510
 30 Asp Pro Leu Asp Phe Lys Arg Leu Leu Asp Pro Ser Ser Thr Gly Trp
 515 520 525
 Leu Ala Lys Ile Pro Val Lys Leu Ser Leu Met Ala Leu Leu Phe Ala
 530 535 540
 35 Phe Tyr Cys Ser Arg Arg Gln Asp Ser Leu Gly Gln Lys Gly Leu Phe
 545 550 555 560
 40 Asn Ser Phe Lys Lys Val Gln Asp Gln Val Lys Gln Phe Ile Ser Ser
 565 570 575
 <210> 112
 <211> 28
 <212> PRT
 45 <213> *Puccinia triticina*
 <400> 112
 Met Ser Ser Thr Ser Pro Leu Glu Phe Thr Pro Thr Asp His Asp Lys
 1 5 10 15
 50 Glu Ile Tyr Ala Thr Val Ser Cys Gly His Ser Ser
 20 25
 <210> 113
 <211> 18
 <212> PRT
 55 <213> *Puccinia striiformis*
 <400> 113
 Met Ala Ser Leu Glu Thr Pro Thr Asp Glu Ile Lys Gln Ile Tyr Ala
 1 5 10 15
 60 Thr Val
 <210> 114
 <211> 505
 <212> PRT
 65 <213> *Microbotryum violaceum*

Ser Ser Asp Tyr Thr Pro Val Asp Gln Ile Glu Pro Ala Tyr Asp Glu
1 5 10 15

Ile Val Lys Thr Phe Leu Thr Gly Lys Thr Arg Ser Ile Lys Trp Arg
20 25 30

Lys Gln Gln Leu Lys Gln Leu Gly Phe Leu Ile Gln Asp Asn Glu Ala
35 40 45

Ala Phe Ile Lys Ala Leu Arg Ser Asp Leu Ser Arg Pro Glu Phe Glu
50 55 60

Thr Thr Ile Ala Glu Leu Asn Pro Leu Lys Thr Asp Val Asn Glu Ala
65 70 75 80

Tyr Asp His Leu Asp Lys Trp Ala Lys Pro Val Lys Ala Lys Thr Thr
85 90 95

Ala Val Trp Ala Leu Ala Asn Ala Tyr Ile Tyr Asn Glu Pro Lys Gly
100 105 110

Ala Val Phe Ile Ile Gly Thr Trp Asn Tyr Pro Leu Val Leu Leu Leu
115 120 125

Ser Pro Leu Val Gly Ala Ile Ala Ala Gly Cys Thr Ala Leu Leu Lys
130 135 140

Pro Ala Glu Gln Ala Pro Ala Val Ala Lys Leu Val Gln Glu Leu Leu
145 150 155 160

Pro Lys Tyr Leu Asp Thr Ser Ala Tyr Lys Ile Ile Leu Gly Ala Val

ES 2 724 998 T3

	165	170	175
5	Asp Gln Val Thr Arg Ala Leu Glu Leu Lys Phe Asp His Ile Phe Tyr 180 185 190		
10	Thr Gly Ser Gly Gly Ile Gly Lys Ile Ile Ala Arg Ala Ala Ala Glu 195 200 205		
15	His Leu Thr Pro Phe Thr Leu Glu Leu Gly Gly Lys Ser Pro Ala Val 210 215 220		
20	Val Phe Asp Asp Ala Asn Ile Asp Ile Thr Ala Arg Arg Ile Met Trp 225 230 235 240		
25	Gly Lys Phe Val Asn Ser Gly Gln Thr Cys Ile Ser Pro Asp Tyr Val 245 250 255		
30	Leu Cys Thr Ala Asp Val Gln Asp Lys Leu Val Ala Ala Met Gln Lys 260 265 270		
35	Val Tyr Lys Glu Phe Thr Thr Asp Ala Lys Gly Gln Glu Lys Ser Met 275 280 285		
40	Val Asn Gly Glu Gly Tyr Ala Arg Ile Val Asn Met Asn His Phe Gly 290 295 300		
45	Arg Ile Ser Thr Met Leu Asp Glu Thr Lys Gly Arg Val Val Val Gly 305 310 315 320		
50	Gly Gly Arg Asn Lys Glu Thr Gly Lys Ile Glu Thr Thr Ile Ile Ala 325 330 335		
55	Asp Val Gly Ala Asp Asp Pro Leu Met Lys Gly Glu Ile Phe Gly Pro 340 345 350		
60	Val Met Pro Ile Val Val Lys Gln Thr Lys Glu Glu Met Val Glu Phe 355 360 365		
65	Ile Gln Glu Arg Asp Asn Pro Leu Ala Leu Tyr Val Phe Thr Gln Ser 370 375 380		
70	Thr Lys Asn Arg Asp Tyr Ile Phe Glu Arg Thr Arg Ser Gly Gly Phe 385 390 395 400		
75	Val Gln Asn Asp Thr Ile Leu His Phe Thr Ile Pro Gly Leu Pro Phe 405 410 415		

5 Gly Gly Ala Gly Ala Ser Gly Ile Gly Ala Tyr His Gly Lys Trp Ser
420 425 430

Phe Asp Thr Phe Ser His Gln Arg Ala Ser Ala His Ile Pro Thr Trp
435 440 445

10 Met Asp Val Ala Leu Asn Ser Arg Tyr Pro Pro Tyr Thr Pro Lys Lys
450 455 460

15 Leu Lys Met Met Leu Leu Ala Thr Lys Ala Val Ile Lys Arg Glu Ser
465 470 475 480

Lys Trp Ser Leu Lys Ser Leu Phe Gly Val Leu Ala Val Val Ala Ala
485 490 495

20 Ile Val Arg Tyr Arg Gln Ser Lys Leu
500 505

<210> 115
<211> 15
<212> PRT
25 <213> Rhodotorula glutinis ATCC 204091
<400> 115
Met Gln Asp Thr Pro Ile Asp Ser Ile Pro Cys Ala Thr Tyr Val
1 5 10 15

30 <210> 116
<211> 504
<212> PRT
<213> Rhodosporidium toruloides NP11
<400> 116
35 Met Ala Ala Met Gln Asp Thr Pro Ile Asp Ser Ile Pro Gln Ala Tyr
1 5 10 15

Asp Thr Val Thr Lys Ala Phe Leu Ser Gly Lys Thr Arg Pro Ile Ala
20 25 30

40 Trp Arg Lys Ala Gln Ile Lys Lys Leu Gly Phe Leu Val Gln Asp Asn
35 40 45

45 Glu Asp Ala Phe Val Arg Ala Leu Glu Gln Asp Phe Gly Arg Pro Ala
50 55 60

Phe Glu Thr Ile Thr Ala Glu Ile Asn Pro Val Lys Ala Glu Ile Asn
65 70 75 80

Glu Val Tyr Asp His Leu Glu Lys Trp Ala Lys Pro Arg Arg Val Lys

ES 2 724 998 T3

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

Lys Ile Glu Val Thr Ile Val Arg Gly Val Lys Pro Asp Asp Ser Leu
 340 345 350
 5 Met Ser Glu Glu Ile Phe Gly Pro Val Leu Pro Ile Met Thr Leu Pro
 355 360 365
 10 Thr Leu Asp Asp Met Val Lys Phe Ile Gln Ser Arg Asp Thr Pro Leu
 370 375 380
 15 Ala Leu Tyr Val Phe Thr Gln Ser Lys Lys Asn Arg Asp Phe Ile Phe
 385 390 395 400
 20 Glu Arg Thr Arg Ser Gly Gly Phe Val Gln Asn Asp Val Leu Val Gln
 405 410 415
 25 Phe Met Ile Pro Gly Leu Pro Phe Gly Gly Thr Gly Ala Ala Gly Tyr
 420 425 430
 Gly Asn Tyr His Gly Arg Arg Thr Phe Asp Thr Phe Ser His Glu Arg
 435 440 445
 30 Ala Ser Ala Asn Val Pro Thr Trp Met Asp Met Ile Met Ala Ser Arg
 450 455 460
 Tyr Pro Pro Tyr Thr Gln Lys Lys Leu Lys Met Leu Leu Phe Ala Thr
 465 470 475 480
 35 Lys Ala Val Ile Lys Lys Pro Ser Lys Phe Gly Ser Ile Ser Arg Leu
 485 490 495
 40 Leu Lys Lys Leu Thr Gly Gln Ala
 500
 <210> 117
 <211> 16
 <212> PRT
 <213> Rhodosporidium toruloides ATCC 10657
 <400> 117
 Met Gln Asp Ile Pro Ile Asp Ser Asp Pro Gln Ala Tyr Asp Thr Val
 1 5 10 15
 45 <210> 118
 <211> 21
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 118
 gcgagggatg gcagtaagac g 21
 <210> 119
 <211> 28
 <212> ADN
 55 <213> Rhodosporidium toruloides
 <400> 119
 aaaggatcca acttgctgc ccagtacc 28
 <210> 120
 <211> 30
 <212> ADN
 60 <213> Rhodosporidium toruloides
 <400> 120
 tttaagcttc acgtacagcc tgtgtagcc 30
 <210> 121
 <211> 30
 65

<212> ADN
 <213> Rhodosporidium toruloides
 <400> 121
 5 tttaggcctg gaggagtcga gcgtgagagt 30
 <210> 122
 <211> 29
 <212> ADN
 <213> Rhodosporidium toruloides
 10 <400> 122
 ttccatggc tgctcgtcg gcactcgag 29
 <210> 123
 <211> 29
 <212> ADN
 <213> Rhodosporidium toruloides
 15 <400> 123
 ttgatatcc attacgcctt gaccgtcag 29
 <210> 124
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 20 <400> 124
 aaagagctcg gtgactgcat gctcgttac 30
 <210> 125
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 25 <400> 125
 aaaggatcct gatggagtag ttgggcacga 30
 <210> 126
 <211> 31
 <212> ADN
 <213> Rhodosporidium toruloides
 30 <400> 126
 tttaagcttc ctctccttg atcttcgcc g 31
 <210> 127
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 35 <400> 127
 tttaggcctg accttgctg cctccctca 30
 <210> 128
 <211> 44
 <212> ADN
 <213> Rhodosporidium toruloides
 40 <400> 128
 tcagaacaac accagatcac tcacaatggc cgctaccctc cgcc 44
 <210> 129
 <211> 26
 <212> ADN
 <213> Rhodosporidium toruloides
 45 <400> 129
 ttccatagg ccgctaccct ccgcca 26
 <210> 130
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 50 <400> 130
 ttgatatct ctaggcatc gtctagagtc 30
 <210> 131
 <211> 33
 <212> ADN
 <213> Rhodosporidium toruloides
 55 <400> 131
 aaagagctct attgttcgac tagactgccc cac 33
 <210> 132
 <211> 32

<212> ADN
 <213> Rhodosporidium toruloides
 <400> 132
 5 aaaggatcca aggaggatat tgtgcacgag ga 32
 <210> 133
 <211> 24
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 133
 10 cgactccaaa tcaccagtt cctc 24
 <210> 134
 <211> 28
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 134
 15 tttaggcctg accgacttg acgacgac 28
 <210> 135
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 135
 20 aaagagctcg cggcactgta cttactacg 30
 <210> 136
 <211> 28
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 136
 25 aaaggatcca cgagacctat ccaaacgc 28
 <210> 137
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 137
 30 ttaagctta aggtcaagtc caaggccaac 30
 <210> 138
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 138
 35 tttaggcctg gctgctggag aaacgaaact 30
 <210> 139
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 139
 40 aaaaggcctc attccctcga ctgcacgcat 30
 <210> 140
 <211> 21
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 140
 45 agaggaggaa gttgtgcagc a 21
 <210> 141
 <211> 30
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 141
 50 ttggatccg ttgcggcgag tcctgtcatc 30
 <210> 142
 <211> 28
 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 142
 55 tttagctcc ggagcgagta agacgagg 28
 <210> 143
 <211> 28
 65

<212> ADN
 <213> Rhodosporidium toruloides
 <400> 143
 5 ttccatggc ctgtacgcc gcccatcc 28
 <210> 144
 <211> 29
 <212> ADN
 <213> Rhodosporidium toruloides
 10 <400> 144
 ttgatatcg gaagggaccg cgctagttg 29
 <210> 145
 <211> 24
 <212> ADN
 15 <213> Rhodosporidium toruloides
 <400> 145
 ttccatggt cgcacgtcc cccg 24
 <210> 146
 <211> 24
 20 <212> ADN
 <213> Rhodosporidium toruloides
 <400> 146
 ttccatggt cgcacgtcc cccg 24
 25
 30
 35
 40
 45
 50
 55
 60
 65

REIVINDICACIONES

1. Una célula huésped fúngica en donde los ácidos grasos totales están compuestos por un aumento del nivel de ácidos grasos poliinsaturados (PUFA), en particular ácido alfa-linolénico (ALA) o ácido gamma-linolénico (GLA), a un nivel de al menos 9 %, preferentemente más del 24 %, con la máxima preferencia más del 49 % en la célula huésped fúngica, en donde el genoma de la célula huésped fúngica se ha modificado de manera que la célula huésped fúngica ha reducido la actividad de la enzima aldehído deshidrogenasa (ALD1) nativa en comparación con una célula huésped fúngica que tiene un genoma no modificado, y en donde el huésped fúngico es una especie del género *Rhodospiridium* o el género *Rhodotorula*.
2. La célula huésped fúngica de conformidad con la reivindicación 1, en donde la ALD1 nativa se selecciona del grupo que consiste en:
 - (a) ALD1 que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:3; y
 - (b) una ALD1 que tiene al menos 75 % de identidad o al menos 85 % de identidad o al menos 95 % de identidad con respecto a la secuencia de aminoácidos expuesta en la SEQ ID NO:3.
3. La célula huésped fúngica de conformidad con la reivindicación 2, en donde la ALD1 nativa es codificada por un ácido nucleico seleccionado del grupo que consiste en:
 - (a) un ácido nucleico que tiene la secuencia de nucleótidos expuesta en la SEQ ID NO:1;
 - (b) un ácido nucleico que tiene la secuencia de nucleótidos expuesta en la SEQ ID NO:2;
 - (c) un ácido nucleico que tiene al menos 75 % de identidad o al menos 85 % de identidad o al menos 95 % de identidad con respecto al ácido nucleico de (a); y
 - (d) un ácido nucleico que tiene al menos 75 % de identidad o al menos 85 % de identidad o al menos 95 % de identidad con respecto al ácido nucleico de (b).
4. La célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 1 a 3, en donde la célula huésped fúngica tiene un gen defectuoso que codifica *ALD1* provocando una pérdida en la actividad de la enzima Ald1.
5. La célula huésped fúngica de 4, en donde el gen defectuoso es provocado mediante la inserción del T-ADN, la recombinación homóloga o la mutagénesis dirigida a sitio.
6. La célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 1 a 3, en donde el genoma de la célula huésped fúngica se ha modificado para regular negativamente la expresión del gen de *ALD1* nativa provocando una pérdida o reducción de la actividad de la enzima Ald1.
7. La célula huésped fúngica de conformidad con la reivindicación 6, en donde la expresión es regulada negativamente por un represor de la transcripción artificial, ARNi, siARN, shARN, miARN, supresión antisentido o con sentido.
8. La célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 1-7, en donde el genoma de la célula huésped fúngica se ha modificado más aún para incluir al menos dos casetes de expresión, en donde cada casete de expresión comprende un promotor unido operativamente a un ácido nucleico que codifica una proteína implicada en la biosíntesis de ácidos grasos y en donde las proteínas se seleccionan del grupo que consiste en:
 - (a) una acil-CoA delta-12 desaturasa
 - (b) una estearoil-CoA-delta-9-desaturasa;
 - (c) una omega-3 desaturasa;
 - (d) un ácido graso elongasa;
 - (e) una acil-CoA carboxilasa (ACC);
 - (f) una ATP: citrato liasa (ACL);
 - (g) una diacilglicerol aciltransferasa (DGA);
 - (h) una enzima málica (MAE); y
 - (i) acil-CoA delta-6 desaturasa;
 opcionalmente en donde el casete de expresión comprende además un terminador de la transcripción unido operativamente al ácido nucleico que codifica una proteína implicada en la biosíntesis de ácidos grasos; y opcionalmente en donde las secuencias codificantes de los ácidos nucleicos contienen al menos 55 % de C y G, preferentemente 60 %-70 % de C y G, en donde al menos el 70 % de los codones tienen una C o G en la 3^{ra} posición.
9. La célula fúngica de conformidad con la reivindicación 8, en donde las proteínas se seleccionan del grupo que consiste en:
 - (a) una acil-CoA delta-12 desaturasa que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:5 y/o SEQ ID NO:94 que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;
 - (b) una estearoil-CoA-delta-9-desaturasa que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:8 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(c) una omega-3 desaturasa que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:10 o SEQ ID NO:12 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(d) un ácido graso elongasa que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:101 o SEQ ID NO:104 (Elo2) o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(e) una acil-CoA carboxilasa (*ACC1*) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:91 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(f) una ATP: citrato liasa (*ACL1*) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:88 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(g) una diacilglicerol aciltransferasa (*DGA1*) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:82 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos;

(h) una enzima málica (*MAE1*) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:85 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de aminoácidos; y

(i) una acil-CoA delta-16 desaturasa que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO:96 o SEQ ID NO:98, o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dichas secuencias de aminoácidos.

10. La célula fúngica de conformidad con la reivindicación 9, en donde las proteínas son codificadas por ácido nucleicos seleccionados del grupo que consiste en un ácido nucleico que tiene la secuencia de nucleótidos expuesta en la SEQ ID NO:4, 6, 7, 9, 11, 80, 81, 83, 84 86, 87, 89, 90, 92, 93,96, 98, 99, 100, 102 y 103 o que tiene al menos 80 % o al menos 90 % o al menos 95 % de identidad con respecto a dicha secuencia de nucleótidos.

11. La célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 8 a 10, en donde el promotor es un promotor aislado de un gen que codifica una proteína seleccionada del grupo que consiste en: gliceraldehído 3-fosfato deshidrogenasa (GPD), proteína portadora de acil-CoA (ACP), ácido graso desaturasa, factor de alargamiento de la traducción (TEF), piruvato decarboxilasa (PDC), enolasa (2-fosfoglicerato deshidratasa) (ENO), peptidilprolil isomerasa (PPI), acetil-CoA carboxilasa (ACC) y transaldolasa y opcionalmente en donde el promotor se aísla de una especie del género *Rhodosporidium* o el género *Rhodotorula*.

12. La célula huésped fúngica de conformidad con la reivindicación 11, en donde el promotor es una secuencia promotora seleccionada del grupo de promotores expuesto en las SEQ ID NO:55-79.

13. La célula huésped fúngica de conformidad con la reivindicación 1 en donde los genes *ELO1* o *ELO2*, expuestos en las Seq ID No. 99 y 102, o un ácido nucleico que tiene al menos 75 % de identidad, 85 % de identidad o 95 % de identidad con respecto a las secuencias de estos, se ha manipulado artificialmente para que tenga un nivel aumentado o reducido de la actividad del ácido graso elongasa.

14. Un método para producir ácidos grasos poliinsaturados (PUFA) omega-3 y omega-6, que comprende cultivar la célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 1 a 13 en condiciones adecuadas para producir los PUFA.

15. Un método para producir triacilglicérido (TAG) que comprende cultivar la célula huésped fúngica de conformidad con cualquiera de las reivindicaciones 1 a 13 en condiciones adecuadas para producir TAG.

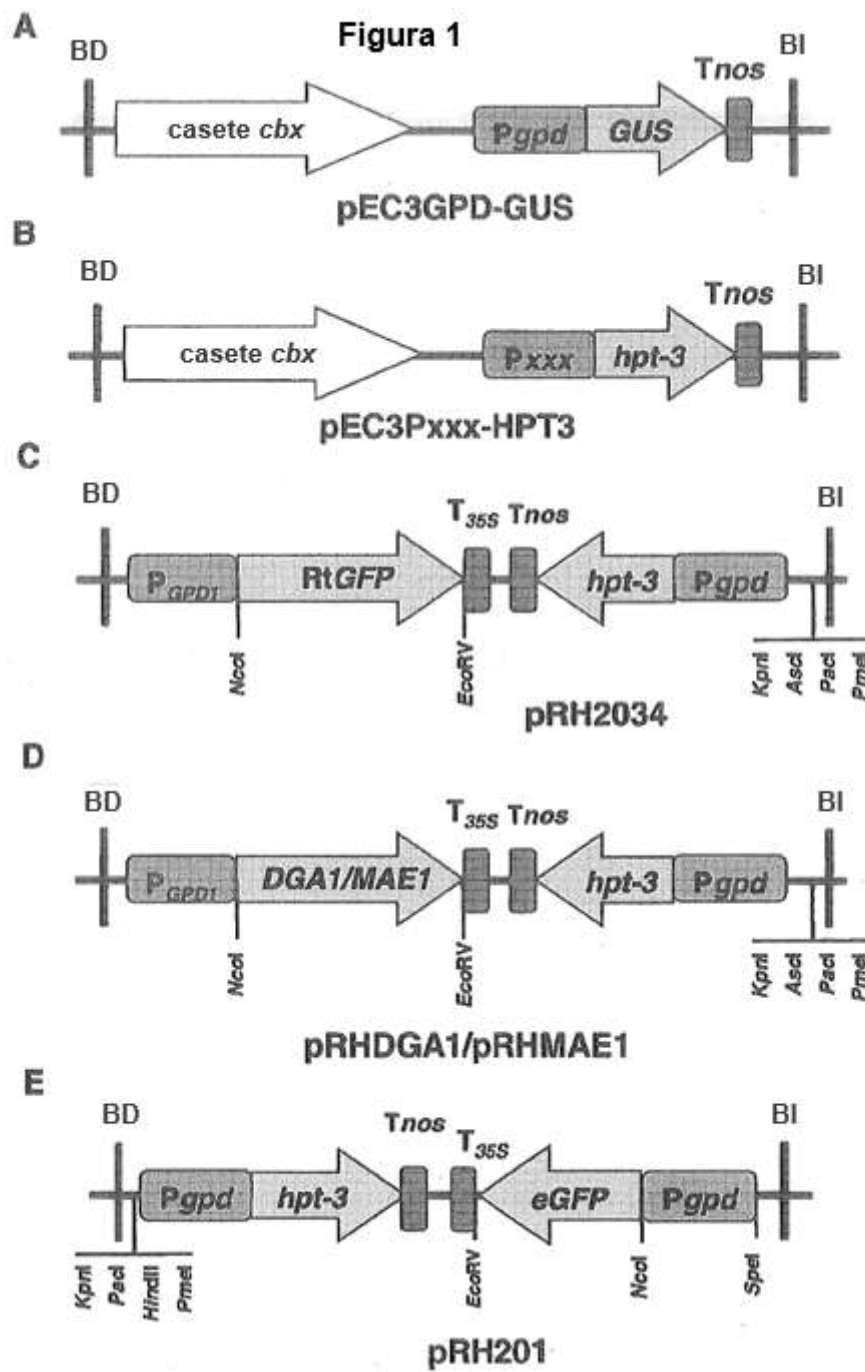


Figura 2

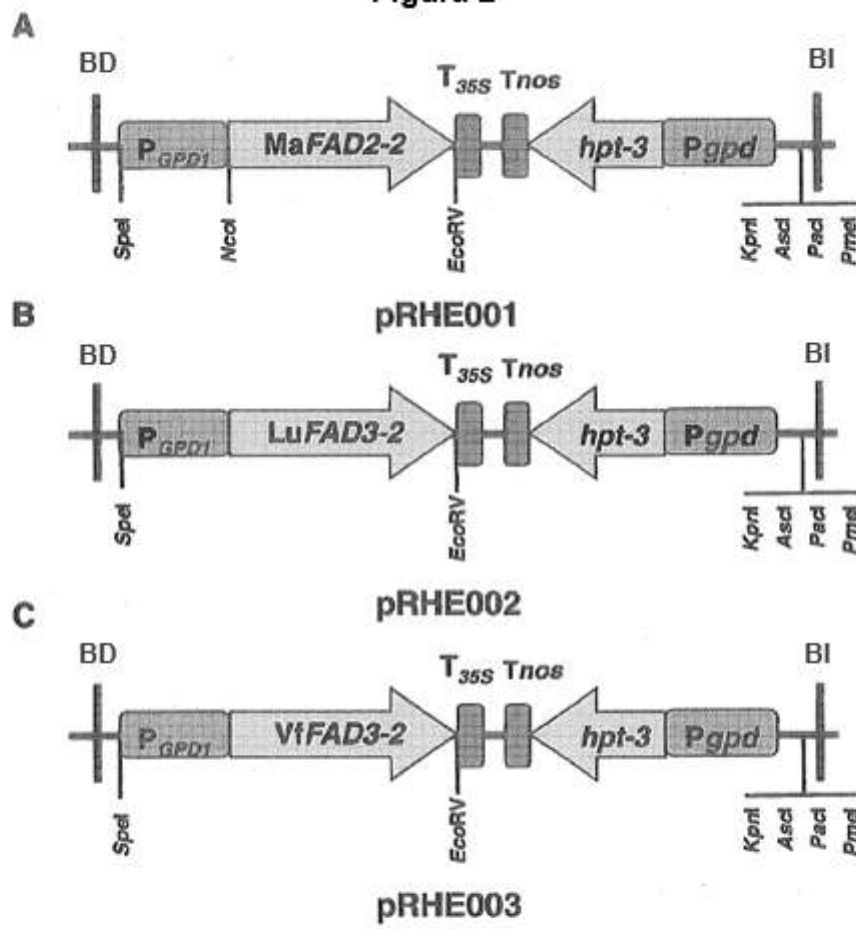


Figura 3

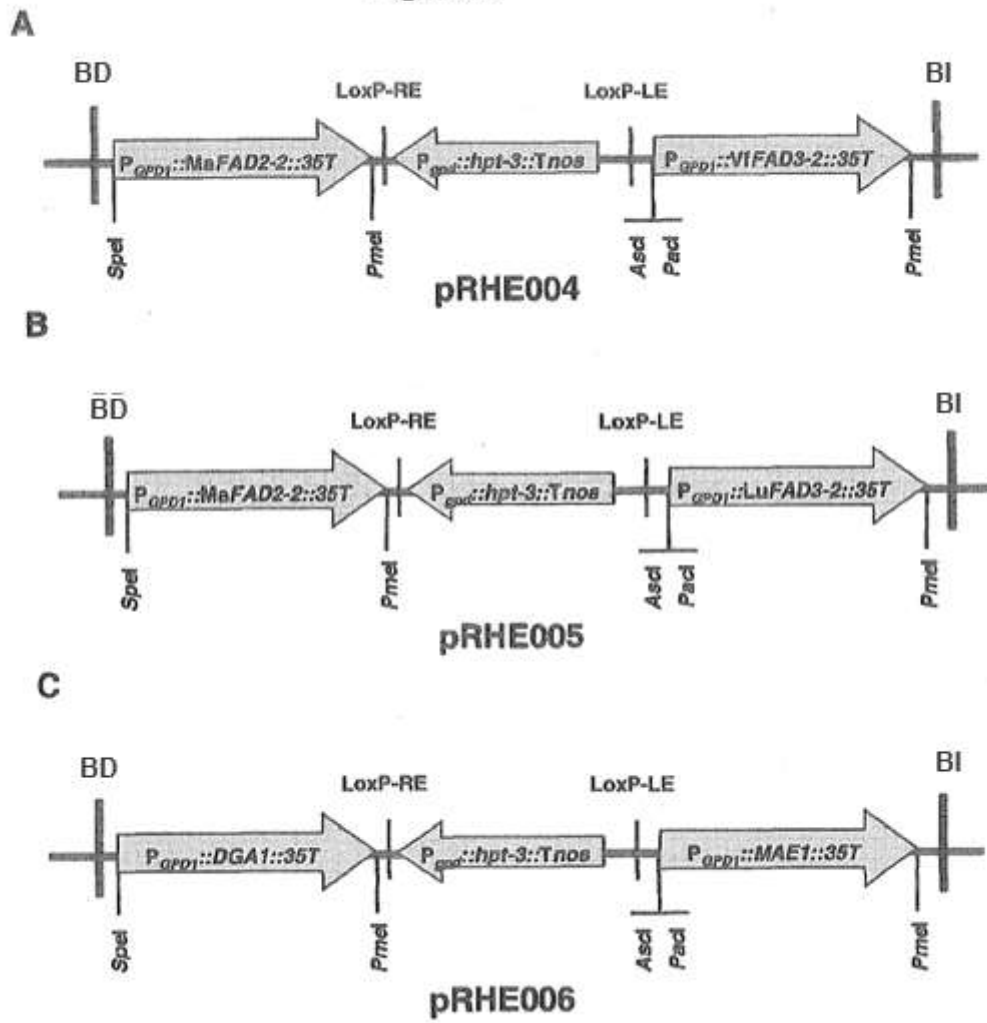


Figura 4

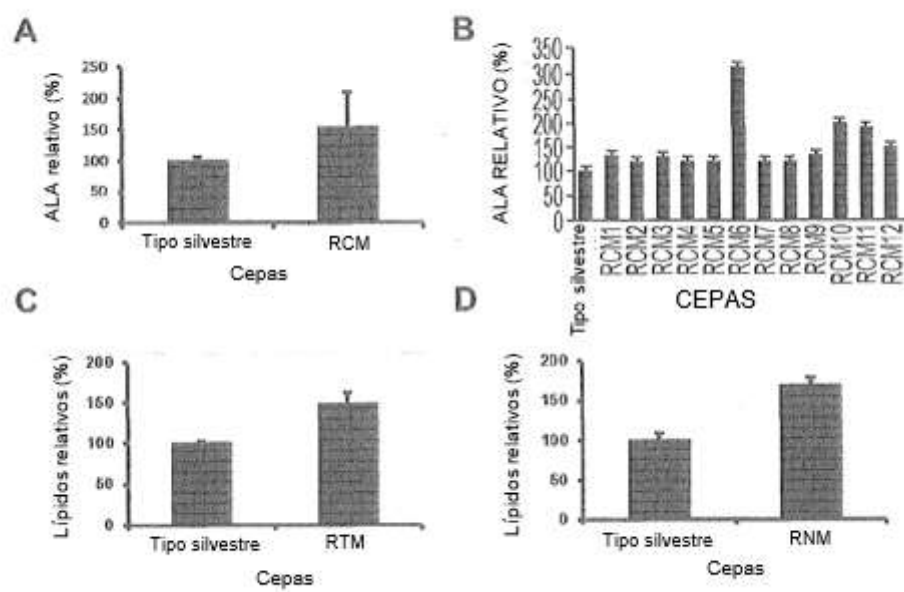
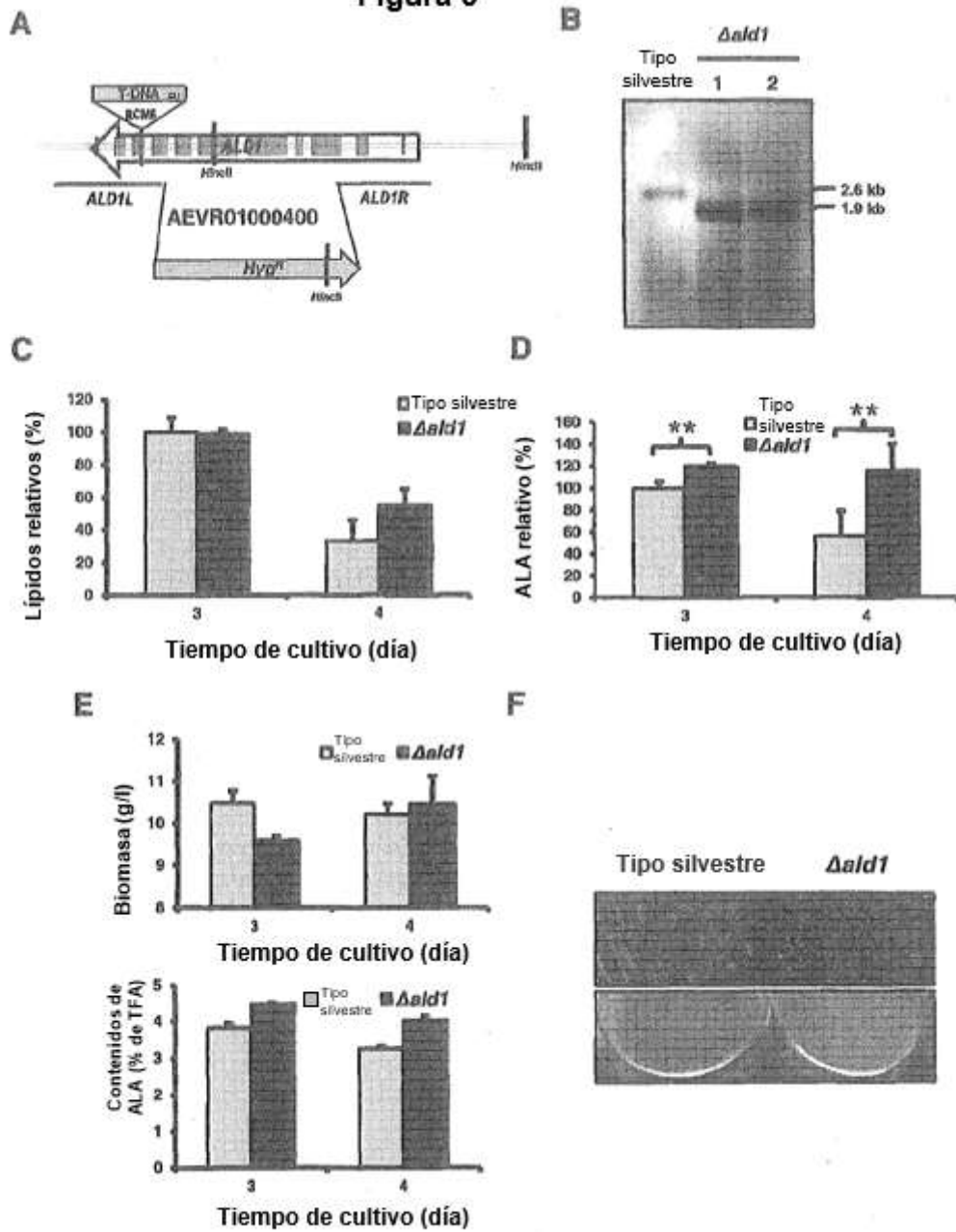


Figura 5



	*	20	*	40	*	60	*
Ab	-	-	-	-	-	-	-
Pc	-	-	-	-	-	-	-
Mg	-	-	-	-	-	-	-
Sr	-	-	-	-	-	-	-
Um	:	MREWLGGNLRFAAIFIRIQSRPDHSLRFTVLDPLPVHYRLSPPSSTLARLHQPRALATSF AHLALLPSSI	-	-	-	-	-
Ml	-	-	-	-	-	-	-
Pg	-	-	-	-	-	-	-
Pt	-	-	-	-	-	-	-
Ps	-	-	-	-	-	-	-
Mv	-	-	-	-	-	-	-
Rg2	-	-	-	-	-	-	-
Rt3	-	-	-	-	-	-	-
Rtl	-	-	-	-	-	-	-

Figura 6

[illegible]

Figura 6
Continuación

Figura 6
Continuación

Ab	140	160	180	
Pc				
Mg				
Si				
Um				
M1				
Pg				
Pt				
Ps				
Mv				
Rg2				
Rt3				
Rt1				

Ab	200	220	240	
Pc				
Mg				
Si				
Um				
M1				
Pg				
Pt				
Ps				
Mv				
Rg2				
Rt3				
Rt1				

Figura 6
Continuación

Figura 6
Continuación

Figura 6
Continuación

Ab : NEEIFGCEVLPIMPVKDVEEATAYNS---NDHPLAVVVFSDAAAYKQKVFSTRQSASAMANEVVIQPGIE *
 Pc : DEEIFGCEVLPIMPVRDVEEATQVNS---RDHPTSLVVFSDAAAFKENVFSTRQSAAAFANEVLVHVGCAT
 Mg : TDEIFGCEVLPIMPFDTRDAMDFVNE---RDQPLALVVFSCNETRDVILAYTRSGGVMRGDOLILHYATD
 Sr : AGEIFGCEVLPIMPVKDVKAANDFIN---RDQPLALVVFAGDSKVKEYFFDNTIRSGACMQGDILHFAVD
 Um : SCEIFGCEVLPIMPVRDVEAANDFIN---RDQPLALVVFAGDNRVKNFFDNTIRSGACMQGDILHFAVD
 M1 : QNEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Pg : ENEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Pt : FEEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Ps : QEEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Mv : KGEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Rg2 : SEEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Rt3 : SEEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 Rt1 : SEEIFGCEVLPIMPVTDQSEMAELLRPIADAEPALVVFVQSSQNFELMRQHTKSGQIMCNDLTIQFAPIP
 E6FGP6 P66
 500 520 540
 * * *
 T SG 6 nd 66

Figura 6
Continuación

560 580 600
 * * *
 : 452
 : 453
 : 457
 : 470
 : 572
 : 492
 : 479
 : 479
 : 481
 : 456
 : 461
 : 464
 : 461

Figura 6
Continuación

Figura 6
Continuación

	620	*	640	*	660	*
Ab	FPVYTDKSIKASQRIKSLPPR	PTGPPRTNNAMANGSATKWGKYFFLAFLATIGGLTKPVKILGRKF				
Pc	YPPYTPKILARLTALKVRMPF	PPGGARPAQKSA				
Mg	YPPYSKSKLDFFRFVLPKLWFG	PPQPTRSSKS				
Sr	YPPYINANINKNRMLAEYSVSPK	PSNPHKSTTSSSSGQGAVALKRLAVLLLSLVLGARNRGLIGWI				
Um	YPPYIEANIKKIDRALAAYS	SPKPSNPHKSIASSSVSLCLSHSRPSPFLSMSQSLSFPMVHYNMLPTQ				
Ml	YPPYIPFKILFSAIMGPARIKKSNP	-----G--LVPKSAEVGKRS--WLPTLTPLSFSTLLLAGYY				
Pg	YPPYISFKFQFQAVLGPSKIKCK	SNPNPPALTDPLDFKRLDDPSSTG--WLAKIPVKLSLMALLFAF				
Pt	YPPYISFKFQFQAVLGPSKIKCK	SNPQPPALTDPLDFKRLDDPSAG--CLAKVPVKFSLVLLFAF				
Ps	YPPYNNIKKILQOSALGPSKIKCK	SNPIVPTLTDPLNFKKLLFNPNNHSSSSLLPKGFPVKLSLLALLSF				
Mv	YPPYIPKILKMLLATKAVIK	RESKWSLSLFGVLAVVAIVRYRQSKL				
Rg2	YPPYIQKILKMLLFATKAVIK	PSKFGSISRLLKKTGQA				
Rt3	YPPYIQKILKMLLFATKAVIK	PSKFGSISRLLKKTGQA				
Rt1	YPPYIQKILKMLLFATKAVIK	PSKFGSISRLLKKTGQA				
	5PPY k k					
		680	*	700	*	
	VPKILG	-----	-----	-----	-----	: 527
	-----	-----	-----	-----	-----	: 487
	-----	-----	-----	-----	-----	: 514
	-----	-----	-----	-----	-----	: 538
	-----	-----	-----	-----	-----	: 640
	ALSR--R	-----YGS DY LKIWMTRFIGAIKQSNR	-----	-----	-----	: 578
	YCSR--RQDS	-----LGQKGLFNSFKKVDQVKQFISS	-----	-----	-----	: 576
	YFSR--RQDS	-----LGQKGLFNSIKKLQDQVKQFISS	-----	-----	-----	: 576
	YYSRNRRESSVFSPLVSSQKAFSDLIQIQEYLLKKNLST	-----	-----	-----	-----	: 591
	-----	-----	-----	-----	-----	: 505
	-----	-----	-----	-----	-----	: 501
	-----	-----	-----	-----	-----	: 504
	-----	-----	-----	-----	-----	: 501

Figura 6
Continuación

Figura 7

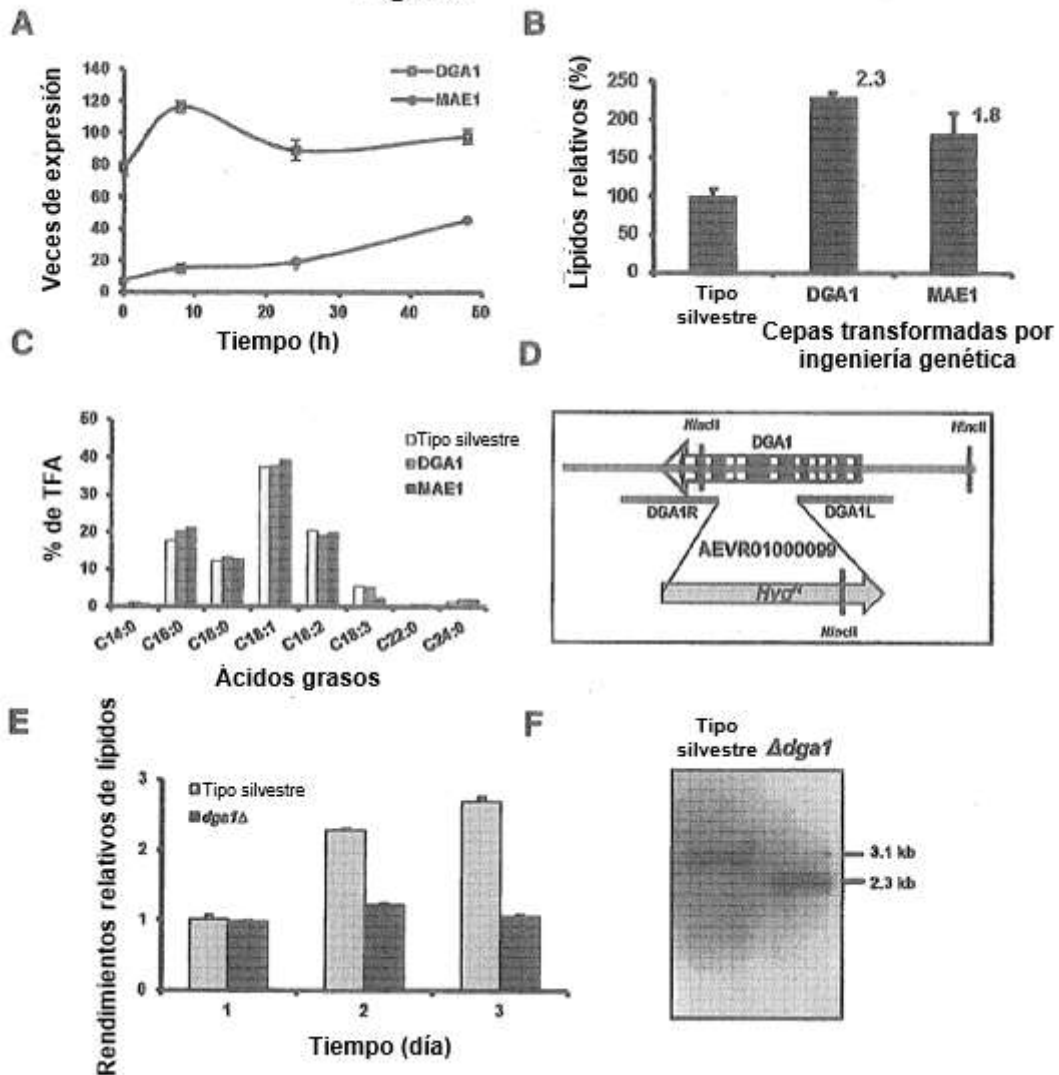


Figura 8

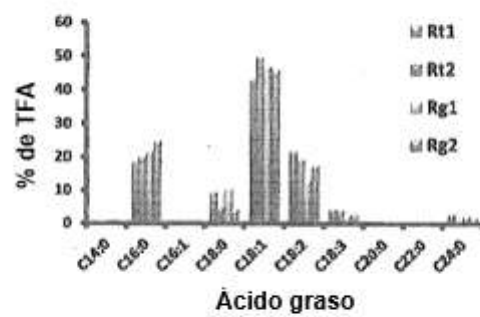


Figura 9

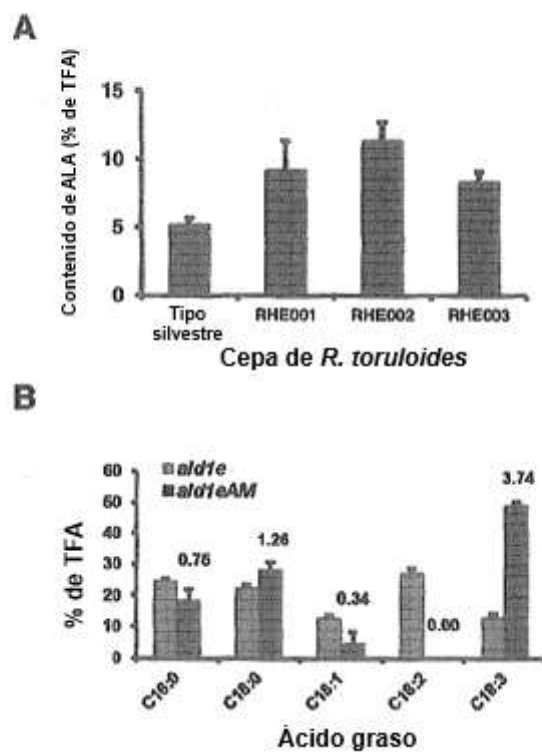


Figura 10

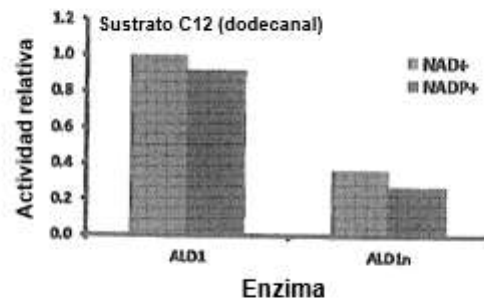


Figura 11

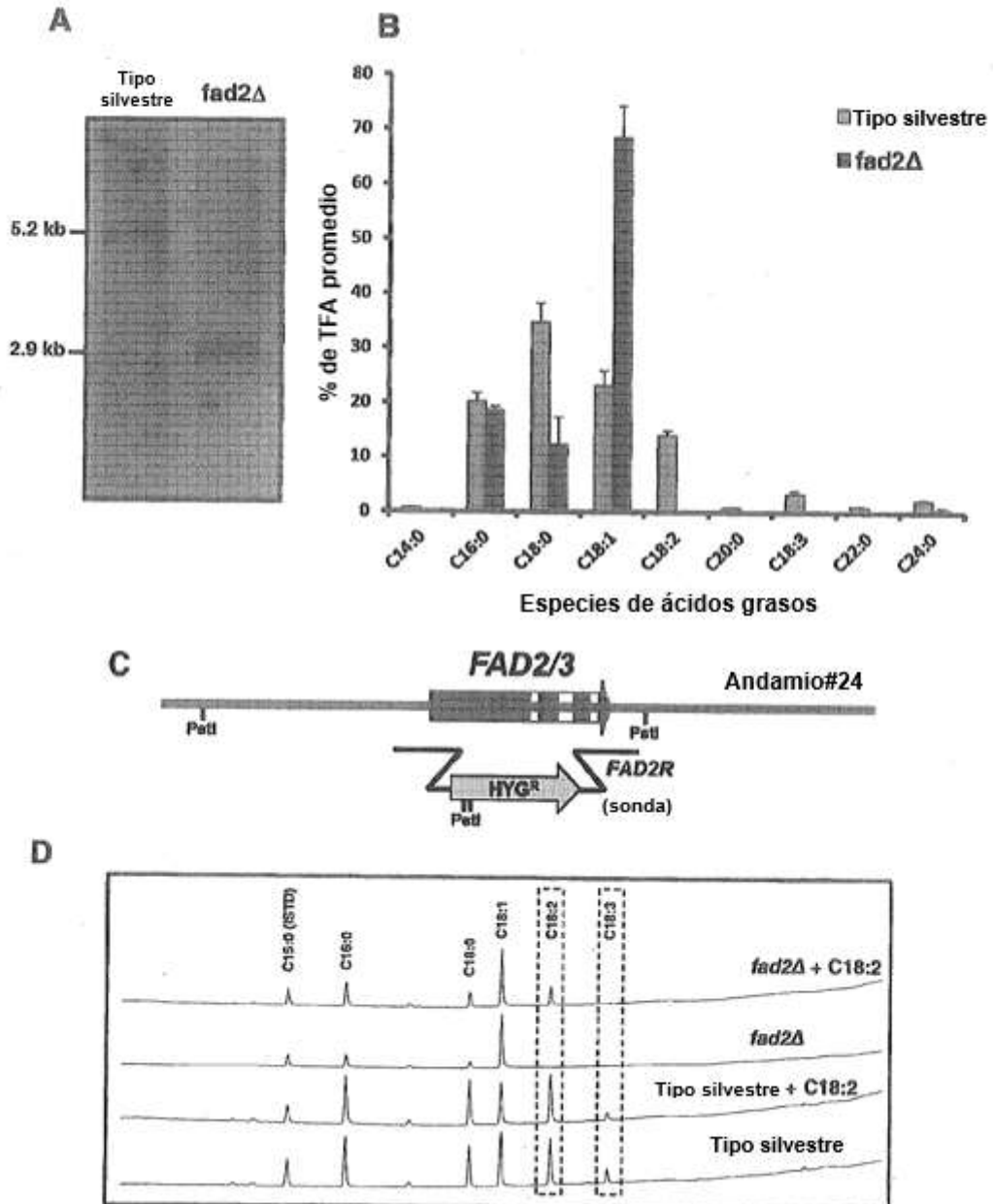


Figura 12

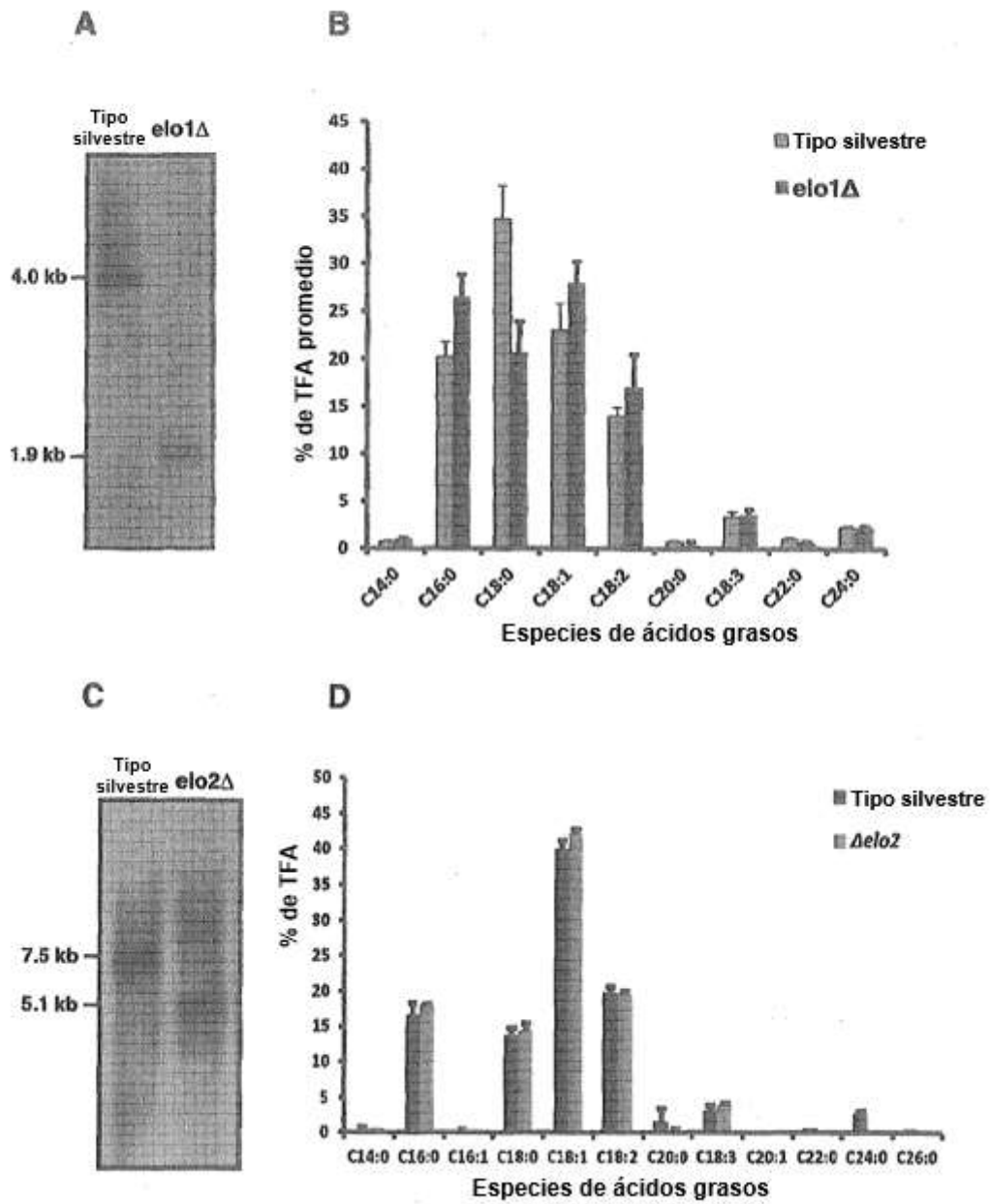


Figura 13

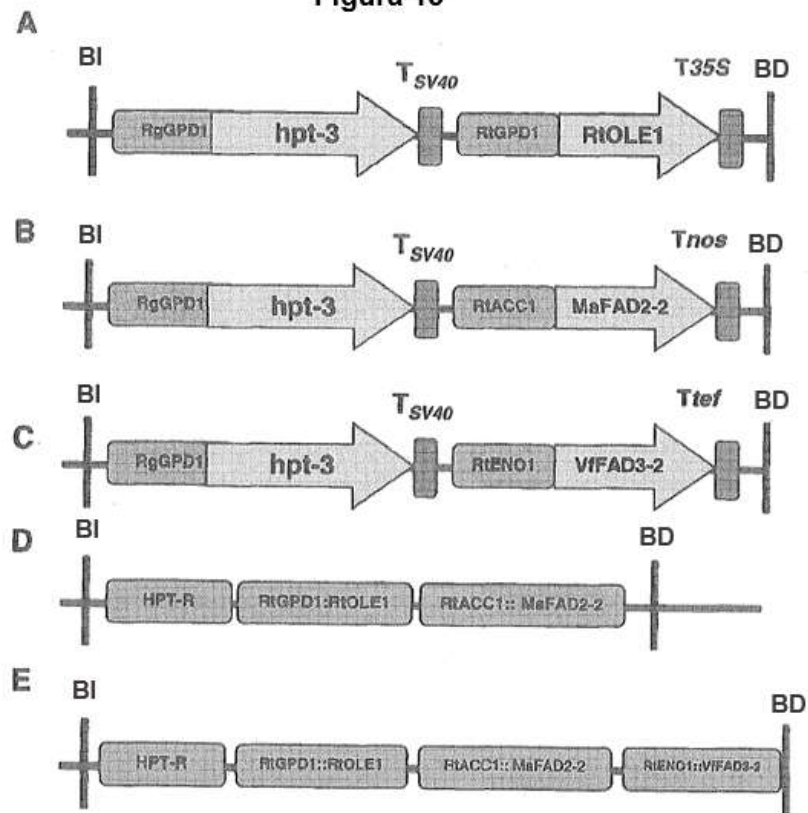


Figura 14

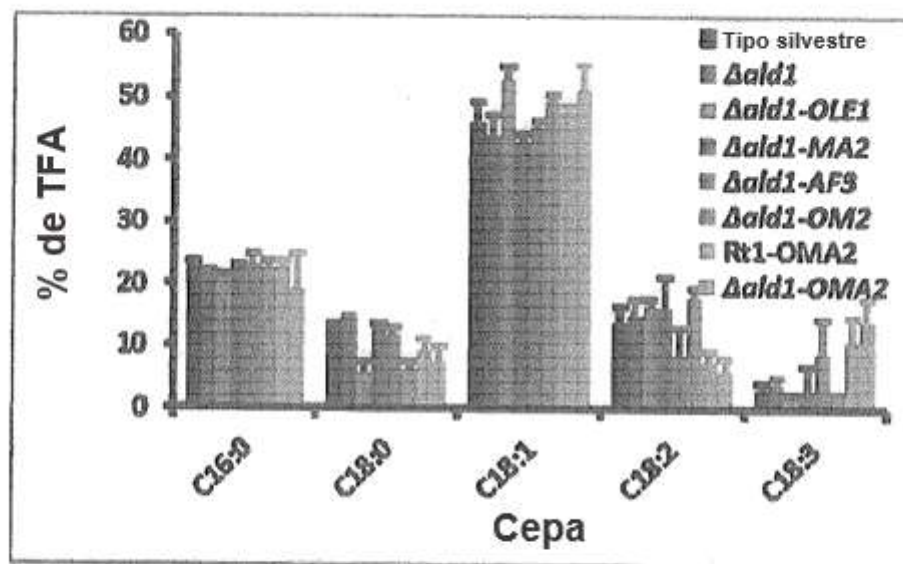


Figura 15

