(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 714 096**

(51) Int. CI.:

C12P 21/06 (2006.01)
C12P 7/20 (2006.01)
C12P 7/64 (2006.01)
C12N 9/00 (2006.01)
C12N 9/16 (2006.01)
C12N 15/79 (2006.01)
C10L 1/02 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **30.11.2009 E 16166059 (2)**

(97) Fecha y número de publicación de la concesión europea: **19.12.2018 EP 3098321**

(54) Título: **Producción de aceites adaptados en microorganismos heterotróficos**

(30) Prioridad:

28.11.2008 US 118590 P
01.12.2008 US 118994 P
30.04.2009 US 174357 P
23.06.2009 US 219525 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
27.05.2019

(73) Titular/es:

CORBION BIOTECH, INC (100.0%)
One Tower Place, Suite 600
South San Francisco, CA 94080, US

(72) Inventor/es:

FRANKLIN, SCOTT;
SOMANCHI, ARAVIND;
ESPINA, KAREN;
RUDENKO, GEORGE y
CHUA, PENELOPE

(74) Agente/Representante:

IZQUIERDO BLANCO, María Alicia

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).

DESCRIPCION

Producción de aceites adaptados en microorganismos heterotróficos

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

Esta solicitud reivindica los beneficios de conformidad con 35 U.S.C. 119(e) de la solicitud provisional de patente de los Estados Unidos núm. 61/118,590, presentada el 28 de noviembre de 2008, solicitud provisional de patente de los Estados Unidos núm. 61/118,994, presentada el 1 de diciembre de 2008, solicitud provisional de patente de los Estados Unidos núm. 61/174,357, presentada el 30 de abril de 2009, y la solicitud provisional de patente de los Estados Unidos núm. 61/219,525, presentada el 23 de junio de 2009.

REFERENCIA AL LISTADO DE SECUENCIA

Esta solicitud incluye un listado de secuencias mostrado en las páginas 1-72 anexas en la presente.

CAMPO DE LA INVENCIÓN

La presente divulgación se refiere a la producción de aceites, combustibles, y oleoquímicos elaborados a partir de microorganismos. En particular, la descripción se relaciona con microalgas oleaginosas, con los métodos de cultivo de las mismas para la producción de compuestos útiles, incluyendo lípidos, ésteres de ácidos grasos, ácidos grasos, aldehídos, alcoholes, y alcanos, así como también con los métodos y reactivos para modificarlas genéticamente con el objetivo de mejorar la eficiencia productiva y modificar el tipo y la composición de los aceites que ellas producen.

ANTECEDENTES DE LA INVENCIÓN

El combustible fósil es un término general para los depósitos geológicos de combustible enterrado de materiales orgánicos, formados a partir de plantas y animales descompuestos, que han sido transformados en aceite crudo, hulla, gas natural o aceites pesados por exposición al calor y presión en la corteza terrestre durante cientos de millones de años. Los combustibles fósiles son una fuente finita, no renovable. El incremento de la demanda energética por la globalización económica también ha ejercido presión en el aumento del costo de los hidrocarburos. Además de la energía, muchas industrias, incluyendo los productores de plástico y sustancias químicas, dependen en gran medida de la disponibilidad de hidrocarburos como materia prima para sus procesos de fabricación. Las alternativas rentables a las fuentes actuales de suministro podrían contribuir a atenuar la presión ascendente en la energía y los costos de estas materias primas.

La publicación PCT núm. 2008/151149 describe los métodos y materiales para cultivar microalgas para la producción de aceite y particularmente ejemplifica la producción de combustible diesel a partir del aceite producido por las microalgas *Chlorella protothecoides.* Existe aún una necesidad de métodos mejorados para producir aceite a partir de microalgas, particularmente de métodos que produzcan aceites de longitud de cadena más corta y con mayor grado de saturación y sin pigmentos, con mayor rendimiento y eficiencia. La presente invención satisface esta necesidad.

SUMARIO DE LA INVENCIÓN

El objeto de la presente invención se describe en las reivindicaciones. La divulgación proporciona células del género *Prototheca* que contienen un gen exógeno y en algunas realizaciones la célula es una variedad de las especies *Prototheca moriformis, Prototheca krugani, Prototheca stagnora* o *Prototheca zopfii,* y en otra realización la célula tiene una secuencia ARNr 23S con al menos 70, 75, 80, 85, o 95 % de identidad de nucleótidos con una o más de las SEQ ID NO: 11-19. En algunas células, el gen exógeno es una secuencia codificante que está en unión operativa con un promotor, y en algunas realizaciones el promotor es de un gen endógeno para una especie del género *Prototheca.* En otras realizaciones la secuencia codificante codifica para una proteína seleccionada del grupo consistente de: sacarosa invertasa, una acil graso-ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, una aldehído graso decarbonilasa, una proteína portadora de acilo, y una proteína que imparte resistencia a un antibiótico. Algunas realizaciones de la acil graso-ACP tioesterasa que tiene actividad hidrolítica hacia uno o más sustratos acil graso-ACP de longitud de cadena C8, C10, C12 o C14, incluyendo las acil-ACP tioesterasas con al menos 50, 60, 70, 80, o 90% de identidad de aminoácidos con una o mas secuencias seleccionadas del grupo consistente de las SEQ ID NO: 59, 61, 63 y 138-140. En otras realizaciones la secuencia codificante comprende una secuencia dirigida al plastidio de microalgas, y en algunas realizaciones la microalga es una especie del género *Prototheca* o *Chlorella* así como de otros géneros de la familia *Chlorellaceae*. En algunas realizaciones, la secuencia dirigida al plastidio tiene al menos 20, 25, 35, 45, o 55% de identidad de secuencia de aminoácidos con una o más de las SEQ ID NO: 127-133 y es capaz de dirigir una proteína codificada por un gen exógeno que no se encuentra en el genoma plastídico del plastidio. En otras realizaciones, el promotor se regula de manera ascendente en respuesta a la reducción o eliminación del nitrógeno

en el medio de cultivo de la célula, tal como al menos una regulación ascendente de tres veces determinada por la abundancia del transcrito en una célula del género *Prototheca* cuando el ambiente extracelular cambia de contener al menos 10 mM o 5 mM de nitrógeno a no contener nada de nitrógeno. En otras realizaciones, el promotor comprende un segmento de 50 nucleótidos o más de una de las SEQ ID NO: 91-102. En otras realizaciones, la célula tiene una secuencia de ARNr 23S tiene al menos 70, 75, 80, 85 o 95% de identidad de nucleótidos con una o más de las SEQ ID NO: 11-19. En otras realizaciones, el gen exógeno se integra a un cromosoma de la célula.

En realizaciones adicionales de las células de la invención, la célula es del género *Prototheca* y contiene un gen exógeno de la acil graso-ACP tioesterasa y un perfil lipídico de al menos 4% C8-C14 de los lípidos totales de la célula, una cantidad de C8 que es al menos 0.3% de los lípidos totales de la célula, una cantidad de C10 que es al menos 2% de los lípidos totales de la célula, una cantidad de C12 que es al menos 2% de los lípidos totales de la célula, una cantidad de C14 que es al menos 4% de los lípidos totales de la célula, y una cantidad de C8-C14 que es 10-30%, 20-30%, o al menos 10, 20 o 30% de los lípidos totales de la célula. En algunas realizaciones la célula contiene además un gen exógeno sacarosa invertasa. En algunas realizaciones, la célula es una cepa de las especies *Prototheca moriformis, Prototheca krugani, Prototheca stagnora* o *Prototheca zopfii*, y en otra realización la célula tiene una secuencia ARNr 23S con al menos 70, 75, 80, 85 o 95% de identidad de nucleótidos con una o más de las SEQ ID NO: 11-19. En otras realizaciones, el gen exógeno acil graso-ACP tioesterasa se integra al cromosoma de la célula. Otras realizaciones de la divulgación incluyen métodos para preparar composiciones de triglicéridos con un perfil lipídico de al menos 4% C8-C14 p/p o porcentaje de área de la composición de triglicéridos, una cantidad de C8 de al menos 0.3% p/p o porcentaje de área, una cantidad de C10 de al menos 2% p/p o porcentaje de área, una cantidad de C12 de al menos 2% p/p o porcentaje de área, una cantidad de C14 de al menos 4% p/p o porcentaje de área, y una cantidad de C8-C14 que es el 10-30%, 20-30%, o al menos 10, 20 o 30% p/p o porcentaje de área. La divulgación también abarca los métodos para preparar composiciones de triglicéridos que incluye cultivar las células ya mencionadas, en donde las células comprenden además un gen exógeno que codifica para una sacarosa invertasa y la sacarosa se suministra como fuente de carbono. En algunas realizaciones la sacarosa invertasa tiene al menos 50, 60, 70, 80, o 90% de identidad de aminoácidos con una o más de las SEQ ID NO: 3, 20-29 y 90.

Las realizaciones de la divulgación incluyen composiciones de aceite de triglicéridos, así como las células que contienen las composiciones de aceite de triglicéridos con un perfil lipídico de al menos 4% C8-C14 y uno o más de los siguientes atributos: 0.1-0.4 microgramos/ml de carotenoides totales, menos de 0.4 microgramos/ml de carotenoides totales, menos de 0.001 microgramos/ml de licopeno, menos de 0.02 microgramos/ml de beta caroteno, menos de 0.02 miligramos de clorofila por kilogramo de aceite; 0.40-0.60 de gamma-tocoferol por 100 gramos de aceite; 0.2-0.5 miligramos de tocotrienoles totales por gramo de aceite, menos de 0.4 miligramos de tocotrienoles totales por gramo de aceite, 4-8 mg por 100 gramos de aceite de campesterol, y 40-60 mg por 100 gramos de aceite de estigmasterol. En algunas realizaciones de la divulgación, las composiciones de aceite de triglicéridos tienen un perfil lipídico de al menos 4% C8-C14 p/p o porcentaje de área de la composición de triglicéridos, una cantidad de C8 de al menos 0.3% p/p o porcentaje de área, una cantidad de C10 de al menos 2% p/p o porcentaje de área, una cantidad de C12 de al menos 2% p/p o porcentaje de área, una cantidad de C14 de al menos 4% p/p o porcentaje de área, y una cantidad de C8-C14 que es 10-30%, 20-30%, o al menos 10, 20 o 30% p/p o porcentaje de área. En otras realizaciones, la composición de aceite de triglicéridos se mezcla con al menos otra composición seleccionada del grupo consistente de: soya, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de vacuno, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuate, grasa blanca selecta, manteca de cerdo, *Camelina* sativa, semillas de mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, linaza (lino), avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez lisa, jojoba, jatrofa, macadamia, nueces del Brasil, aguacate, petróleo, o una fracción destilada de cualquiera de los aceites anteriores.

Los métodos de la divulgación también incluyen procesar los aceites antes mencionados realizando una o más de las reacciones químicas de la lista consistente de: transesterificación, hidrogenación, hidrocraqueo, desoxigenación, isomerización, interesterificación, hidroxilación, hidrólisis para producir ácidos grasos libres, y saponificación. La divulgación incluye, además, los combustibles hidrocarburos obtenidos a partir de la hidrogenación e isomerización de los aceites mencionados anteriormente y ésteres de alquilo de ácidos grasos obtenidos a partir de la transesterificación de los aceites mencionados anteriormente. En algunas realizaciones, el combustible hidrocarburo se obtiene a partir de triglicéridos aislados de células del género *Prototheca* en donde el intervalo de destilación según ASTM D86 T10-T90 es de al menos 25ºC. En otras realizaciones, el combustible éster de alquilo de ácidos grasos se obtiene a partir de triglicéridos aislados de células del género *Prototheca* en donde la composición tiene un tiempo de maceración en frío de menos de 120 segundos según ASTM D6751 A1).

La divulgación también incluye una composición que comprende (a) un polisacárido que comprende uno o más monosacáridos de la lista que consiste en 20-30 porciento molar de galactosa; 55-65 porciento molar de glucosa; y 5-15 porciento molar de manosa; (b) proteína; y (c) ADN que contiene una secuencia de ARNr 23S con identidad nucleotídica de al menos 70, 75, 80, 85 o 95% de una o más de las SEQ ID NO: 11-19; y (d) un gen exógeno. En algunas realizaciones el gen exógeno se selecciona de la sacarosa invertasa y de la acil graso-ACP

tioesterasa, y en otras realizaciones la composición comprende, además, lípidos con un perfil lipídico de al menos 4% de C8-C14. En otras realizaciones, la composición se formula para consumo como alimento animal.

La divulgación incluye ácidos nucleicos recombinantes que codifican promotores que se regulan de manera ascendente en respuesta a la reducción o eliminación del nitrógeno en el medio de cultivo de la célula del género *Prototheca*, tal como al menos una regulación ascendente de 3 veces determinada por la abundancia del transcripto cuando el ambiente extracelular cambia de contener al menos 10 mM o 5 mM de nitrógeno a no contener nada de nitrógeno. En algunas realizaciones, el ácido nucleico recombinante contiene un segmento de 50 nucleótidos o más de una de las SEQ ID NO: 91-102. La divulgación también incluye vectores de ácidos nucleicos que contienen un casete de expresión que comprende (a) un promotor que es activo en una célula del género *Prototheca*; y (b) una secuencia codificante en unión operativa con el promotor en donde la secuencia codificante contiene los codones preferidos o los segundos más preferidos de la Tabla 1 para al menos 20, 30, 40, 50, 60, o 80% de los codones de la secuencia codificante. En algunos vectores, la secuencia codificante comprende una secuencia dirigida al plastidio, dentro en el marco con una acil graso-ACP tioesterasa, incluyendo la tioesterasa que tiene actividad hidrolítica hacia uno o más sustratos acil graso-ACP de longitud de cadena C8, C10, C12 o C14. Algunos vectores incluyen secuencias dirigidas al plastidio que codifican péptidos capaces de dirigir una proteína al plastidio de una célula del género *Prototheca*, incluyendo los de microalgas y aquellos en donde la secuencia dirigida al plastidio tienen al menos 20, 25, 35, 45, o 55% de identidad de secuencia de aminoácidos con una o más de las SEQ ID NO: 127-133 y es capaz de dirigir una proteína al plastidio de una célula del género *Prototheca*. Otros vectores adicionales de la invención comprenden secuencias endógenas de ácidos nucleicos al genoma nuclear de una célula del género *Prototheca*, en donde la secuencia es de al menos 200 nucleótidos de longitud, y algunos vectores comprenden la primera y segunda secuencias endógenas de ácidos nucleicos del genoma nuclear de la célula del género *Prototheca*, en donde la primera y segunda secuencias (a) son cada una de al menos 200 nucleótidos de longitud; (b) flanquean el casete de expresión; y (c) se localizan en el mismo cromosoma *Prototheca* a una distancia de no más de 5, 10, 15, 20, y 50 kB.

La divulgación también incluye un ácido nucleico recombinante que tiene al menos 80, 90, 95 o 98% de identidad de nucleótidos con una o ambas de las SEQ ID NO: 134-135 y un ácido nucleico recombinante que codifica para una proteína con tiene al menos 80, 90, 95 o 98% de identidad de aminoácidos con una o ambas de las SEQ ID NO: 136-137.

La divulgación también incluye métodos para producir composiciones de triglicéridos que comprende (a) cultivar una población de células del género *Prototheca* en presencia de una fuente fija de carbono, en donde: (i) las células contienen un gen exógeno; (ii) las células acumulan al menos 10, 20, 30, 40, 60, o 70% de su peso seco de las células en forma de lípidos; y (iii) la fuente fija de carbono se selecciona de un grupo consistente de sorgo y material celulósico despolimerizado; y (b) aislar los componentes lipídicos de los microorganismos cultivados. En algunas realizaciones, la fuente fija de carbono es material celulósico despolimerizado seleccionado del grupo consistente de rastrojo de maíz, *Miscanthus*, sorgo forrajero, pulpa de remolacha y bagazo de caña de azúcar, que se sometieron opcionalmente a un proceso de lavado con agua antes de la etapa de cultivo. En algunos métodos, la fuente fija de carbono es material celulósico despolimerizado y el nivel de glucosa de este material celulósico despolimerizado se concentra a un nivel de al menos 300g/litro, al menos 400 g/litro, al menos 500 g/litro o al menos 600g/litro antes de la etapa de cultivo y se alimenta al cultivo todo el tiempo mientras las células crecen y acumulan lípidos. En algunos métodos el gen exógeno codifica una acil graso-ACP tioesterasa que tiene actividad hidrolítica hacia uno o más sustratos acil graso-ACP de longitud de cadena C8, C10, C12 o C14, y en algunos métodos los triglicéridos tienen un perfil lipídico de al menos 4% C8-C14 y uno o más de los siguientes atributos: 0.1-0.4 microgramos/ml de carotenoides totales; menos de 0.02 miligramos de clorofila por kilogramo de aceite; 0.40-0.60 miligramos de gamma-tocoferol por 100 gramos de aceite; 0.2-0.5 miligramos de tocotrienoles totales por gramo de aceite, 4-8 mg por 100 gramos de aceite de campesterol, y 40-60 mg por 100 gramos de aceite de estigmasterol.

Otros métodos de la divulgación incluyen producir una composición de triglicéridos que comprende: (a) cultivar una población de microorganismos en presencia de material celulósico despolimerizado, en donde: (i) el material celulósico despolimerizado se somete a un proceso de lavado con agua antes de la etapa de cultivo; (ii) las células acumulan al menos 10, 20, 30, 40, 60, o 70% de su peso seco de las células en forma de lípidos; y (iii) el material celulósico despolimerizado se concentra a al menos 300, 400, 500, ó 600 g/litro de glucosa antes de la etapa de cultivo; (iv) los microorganismos se cultivan en una reacción de lote alimentado, en la cual el material celulósico despolimerizado de al menos 300, 400, 500, o 600 g/litro de glucosa se alimenta a los microorganismos; y (b) aislar los componentes lipídicos de los microorganismos cultivados. En algunas realizaciones, la fuente fija de carbono es material celulósico despolimerizado seleccionado del grupo consistente de rastrojo de maíz, *Miscanthus*, sorgo forrajero, pulpa de remolacha y bagazo de caña de azúcar. En otras realizaciones, los microorganismos son una especie del género *Prototheca* y contienen un gen exógeno que incluye un acil graso-ACP tioesterasa que tiene actividad hidrolítica hacia uno o más sustratos acil graso-ACP de longitud de cadena C8, C10, C12 o C14. Otro método de la invención abarca fabricar aceite de triglicéridos cultivando una célula que tiene una secuencia ARNr 23S con al menos 90 o 96% de identidad de nucleótidos de la SEQ ID NO: 30 en presencia de sacarosa como fuente de carbono.

La divulgación también incluye los métodos para fabricar un producto químico que comprende llevar a cabo una o más reacciones químicas de la lista consistente de transesterificación, hidrogenación, hidrocraqueo, desoxigenación, isomerización, interesterificación, hidroxilación, hidrólisis, y saponificación de un aceite de triglicéridos, en donde el aceite tiene un perfil lipídico de al menos 4% C8-C14 y uno o más de los siguientes atributos: 0.1-0.4 microgramos/ml de carotenoides totales; menos de 0.02 miligramos de clorofila por kilogramo de aceite; 0.10-0.60 miligramos de gamma-tocoferol por 100 gramos de aceite; 0.1-0.5 miligramos de tocotrienoles totales por gramo de aceite, 1-8 mg por 100 gramos de aceite de campesterol, y 10-60 mg por 100 gramos de aceite de estigmasterol. Algunos métodos se realizan al fabricar el aceite mediante el cultivo de células del género *Prototheca* que comprende un gen exógeno acil graso-ACP tioesterasa que codifica una acil graso-ACP tioesterasa que tiene actividad hidrolítica hacia uno o más sustratos acil graso-ACP de longitud de cadena C8, C10, C12 o C14. En algunos métodos, la reacción de hidrólisis se seleccionada del grupo consistente de saponificación, hidrólisis ácida, hidrólisis alcalina, hidrólisis enzimática, hidrólisis catalítica, e hidrólisis con agua caliente comprimida, incluyendo una reacción de hidrólisis catalítica en donde el aceite se divide en glicerol y ácidos grasos. En otros métodos, los ácidos grasos se someten a una reacción de aminación para producir compuestos grasos nitrogenados o una reacción de ozonólisis para producir ácidos mono y dibásicos. En algunas realizaciones, el aceite se somete a un método de desdoblamiento de triglicéridos, seleccionado del grupo consistente de división enzimática y división por presión. En algunos métodos, la reacción de hidrólisis es seguida por una reacción de condensación. Otros métodos incluyen una reacción de hidroprocesamiento del aceite, en donde opcionalmente el producto de la reacción de hidroprocesamiento se somete a una reacción de desoxigenación o condensación antes o simultáneamente a la reacción de hidroprocesamiento. Otros métodos incluyen adicionalmente una reacción de eliminación del gas. Los métodos adicionales incluyen procesar los aceites mencionados anteriormente mediante una reacción de desoxigenación seleccionada del grupo consistente de: una reacción de hidrogenólisis, hidrogenación, una reacción de hidrogenación-hidrogenólisis consecutiva, una reacción de hidrogenólisis-hidrogenación consecutiva, y una reacción de hidrogenación-hidrogenólisis combinada. En algunos métodos, una reacción de desoxigenación es seguida por la reacción de condensación. Otros métodos incluyen realizar una reacción de esterificación en los aceites ante mencionados, opcionalmente una reacción de interesterificación o una reacción de transesterificación. Otros métodos incluyen realizar una reacción de hidroxilación sobre los aceites ante mencionados, en donde, opcionalmente, una reacción de hidroxilación es seguida por la reacción de condensación.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Las Figuras 1 y 2 ilustran las curvas de crecimiento de las especies *Prototheca* y *Chlorella luteoviridis* cepa SAG 2214 cultivadas en sorgo como fuente de carbono.

La Figura 3 muestra el crecimiento en el tiempo de SAG 2214 en glucosa y sacarosa.

La Figura 4 muestra los mapas de los casetes usados en las transformaciones de *Prototheca*, como se describe en el Ejemplo 3.

La Figura 5 muestra los resultados del análisis de transferencia de tipo Southern de tres transformantes de la cepa UTEX 1435, como se describe en el Ejemplo 3.

La Figura 6 muestra un esquema de la construcción transgénica del codón optimizado y del codón no optimizado *suc2* (sacarosa invertasa de levadura (yInv)). Se indican los sitios de restricción relevantes del clonaje y las flechas indican la dirección de la transcripción.

La Figura 7a muestra los resultados del crecimiento de *Prototheca moriformis* en azúcares derivados de celulosa (rastrojo de maíz, pulpa de remolacha, caña de sorgo, *Miscanthus* y control de glucosa). El crecimiento se expresa en medidas de densidad óptica (lecturas A750).

La Figura 7b muestra los resultados de los experimentos de crecimiento usando *Prototheca moriformis* usando diferentes niveles de azúcares celulósicos derivados del rastrojo de maíz en comparación con el control glucosa/xilosa.

La Figura 7c muestra el impacto de la xilosa sobre la producción de lípidos en cultivos de *Prototheca*.

La Figura 7c muestra el impacto de la concentración de sal ((Na_2SO_4) y del antiespumante en el crecimiento (en peso seco de células (PSC)) de *Prototheca*.

La Figura 8 muestra el impacto del tratamiento hidrotérmico de varios materiales celulósicos (bagazo de caña de azúcar, caña de sorgo, *Miscanthus* y pulpa de remolacha) y la corriente de azúcar resultante en el crecimiento de *Prototheca*.

La Figura 9 muestra la disminución de los niveles de hidroximetil furfurales (HMF) y furfurales en la biomasa celulósica (bagazo de caña de azúcar, caña de sorgo, *Miscanthus* y pulpa de remolacha) después de ciclos repetidos de tratamiento hidrotérmico.

La Figura 10 muestra un esquema del proceso de sacarificación de los materiales celulósicos para generar las corrientes de azúcar adecuadas para el uso en la producción heterotrófica de aceite en un fermentador.

La Figura 11 muestra la disminución de los niveles de HMF y de los furfurales en bagazo de caña de azúcar explotado después de ciclos repetidos de tratamiento hidrotérmico.

La Figura 12 muestra un esquema de las construcciones de tioesterasa usadas en las transformaciones de *Prototheca*. Los promotores heterólogos beta-tubulina (determinante Neo^R) y glutamato deshidrogenasa se derivan de *Chlamydomonas reinhardtii* y *Chlorella sorokiniana*, respectivamente. La nitrato reductasa 3´UTR se derivó de *Chlorella vulgaris*. Se indican los sitios de restricción relevantes del clonaje y las flechas indican la dirección en que ocurre la transcripción.

La Figura 13 muestra un cromatograma del diesel renovable producido a partir del aceite de triglicérido de *Prototheca.*

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

La presente invención surge a partir del descubrimiento de que la *Prototheca* y ciertos microorganismos relacionados tienen propiedades inesperadamente ventajosas para la producción económica y en grandes cantidades de aceites, combustibles, y otras composiciones de hidrocarburos o lipídicas, así como del descubrimiento de métodos y reactivos para alterar genéticamente estos microorganismos para mejorar estas propiedades. Los aceites producidos por estos microorganismos se pueden usar en el combustible para la transportación, en la industria petroquímica y/o alimenticia y cosmética, entre otras aplicaciones. La transesterificación de los lípidos produce ésteres de ácidos grasos de cadena larga útiles como biodiesel. Otros procesos enzimáticos y químicos pueden modificarse para producir ácidos grasos, aldehídos, alcoholes, alcanos y alquenos. En algunas aplicaciones, se produce diesel renovable, combustible de reactor, u otros compuestos hidrocarburos. La presente divulgación también proporciona métodos para cultivar microalgas con una productividad elevada y un alto rendimiento lipídico, y/o para la producción más rentable de las composiciones descritas en la presente.

Esta descripción detallada de la divulgación se divide en secciones para conveniencia del lector. La Sección I proporciona las definiciones de los términos usados en la presente. La Sección 2 proporciona una descripción de las condiciones de cultivo útiles en los métodos de la invención. La Sección 3 proporciona una descripción de los materiales y métodos de ingeniería genética. La Sección 4 proporciona una descripción de la ingeniería genética de *Prototheca* para permitir la utilización de la sacarosa. La Sección 5 proporciona una descripción de la ingeniería genética de *Prototheca* para modificar la biosíntesis de los lípidos. La Sección 6 describe los métodos para producir combustibles y sustancias químicas. La Sección 7 describe los ejemplos y realizaciones de la invención. La descripción detallada de la invención es seguida por los ejemplos que ilustran varios aspectos y realizaciones de la invención.

I. DEFINICIONES

A menos que se defina de otra manera, todos los términos técnicos y científicos usados en este documento tienen el significado comúnmente entendido por una persona experta en la materia a la que pertenece esta divulgación. Las siguientes referencias proporcionan al experto una definición general de muchos de los términos usados en esta divulgación: Singleton *y otros*, *Dictionary of Microbiology and Molecular Biology* (2da edición, 1994); *The Cambridge Dictionary of Science and Technology* (Ediciones Walker, 1988); *The Glossary of Genetics*, 5ta Edición, R. Rieger *y otros* (editores), Springer Verlag (1991); y Hale & Marham, *The Harper Collins Dictionary of Biology* (1991). Como se usan en este documento, los siguientes términos tienen los significados atribuidos a ellos, a menos que se especifique de otra manera.

“Activo en microalgas” se refiere a un ácido nucleico que es funcional en las microalgas. Por ejemplo, un promotor que se ha usado para manejar un gen de resistencia a los antibióticos para impartir resistencia antibiótica a una microalga transgénica está activo en las microalgas.

“Proteína portadora de acilos” o “ACP”, es una proteína que se une a una cadena acilo en crecimiento durante la síntesis de ácidos grasos como un tiol éster al tiol distal de la porción de 4-fosfopantoteína y comprende un componente del complejo de la ácido graso sintasa.

“Molécula de Acil-CoA” o “acil-CoA”, es una molécula que comprende una porción acilo covalentemente unida a la coenzima A a través de un enlace tiol éster en el tiol distal de a porción de 4-fosfopantoteína de la coenzima A.

“Porcentaje de área”, se refiere al área de picos observados usando los métodos de detección FAME GC/FID, en los que cada ácido graso en la muestra se convierte en un metil éster de ácido graso (FAME) antes de la detección. Por ejemplo, un pico separado se observa para un ácido graso de 14 átomos del carbono sin insaturaciones (C14:0) comparado con cualquier otro ácido graso, por ejemplo C14:1. El área del pico para cada clase de FAME es directamente proporcional a su porcentaje de la composición en la mezcla y se calcula en base a la suma de todos los picos presentes en la muestra (es decir [área bajo el pico específico/área total de todos los picos medidos] X 100). Con referencia a los perfiles lipídicos de los aceites y las células de la invención, "al menos 4% de C8-C14" significa que al menos el 4% de los ácidos grasos totales en la célula o en la composición del glicerolípidos extraídos, tiene una longitud de cadena que incluye 8, 10, 12 ó 14 átomos de carbono.

“Axénico”, es un cultivo de un organismo libre de contaminación por otros organismos vivos.

“Biodiesel”, es un éster de alquilo de ácidos grasos producido biológicamente adecuado para su uso como combustible en un motor de diesel.

“Biomasa”, es el material producido por el crecimiento y/o propagación de células. La biomasa puede contener células y/o contenido intracelular, así como material extracelular, incluye pero no se limita a, los compuestos secretados por una célula.

“Biorreactor”, es un recinto total o parcial donde se cultivan las células, opcionalmente en suspensión.

"Catalizador" es un agente, tal como una molécula o un complejo macromolecular, capaz de facilitar o promover una reacción química de un reactante a su producto sin convertirse en una parte del producto. Un catalizador aumenta la velocidad de una reacción, después de lo cual, el catalizador puede actuar sobre otro reactante para formar el producto. Un catalizador generalmente disminuye la energía de activación total requerida para la reacción, de modo que esta ocurre más rápidamente o a una temperatura más baja. Así, un equilibrio de la reacción puede lograrse más rápidamente. Los ejemplos de catalizadores incluyen enzimas, que son catalizadores biológicos; calor, que es un catalizador no biológico; y metales usados en los procesos de refinamiento de los aceites fósiles.

“Material celulósico”, es el producto de la digestión de la celulosa, incluyendo glucosa y xilosa, y opcionalmente compuestos adicionales como disacáridos, oligosacáridos, lignina, furfurales y otros compuestos. Los ejemplos no limitantes de fuentes de material celulósico incluyen bagazo de caña de azúcar, pulpa de azúcar de remolacha, rastrojos de maíz, astillas de madera, aserrín y pasto varilla.

“Co-cultivo”, y variantes de éste como “co-cultivar” y “co-fermentar”, se refieren a la presencia de dos o más tipos de células en el mismo biorreactor. Los dos o más tipos de células pueden ser ambos microorganismos, como las microalgas, o pueden ser una célula de microalgas cultivada con un tipo celular diferente. Las condiciones del cultivo pueden ser aquellas que fomentan el crecimiento y/o la propagación de los dos o más tipos celulares o aquellas que facilitan el crecimiento y/o proliferación de una, o un subconjunto de las dos o más células, mientras se mantiene el crecimiento celular para el resto.

“Cofactor”, es cualquier molécula, distinta del sustrato, necesaria para que una enzima pueda realizar su actividad enzimática.

“ADN complementario” o “ADNc”, define a una copia de ADN del ARNm, normalmente obtenida por transcripción inversa del ARN mensajero (ARNm) o por amplificación (por ejemplo, vía reacción en cadena de la polimerasa (“PCR”)).

“Cultivado”, y variantes de este como “cultivo” y "fermentado", se refieren al fomento intencional del crecimiento (incremento del tamaño celular, contenido celular, y/o actividad celular) y/o propagación (incremento en el número de células vía mitosis) de una o varias células mediante el uso de condiciones de cultivo seleccionadas y/o controladas. La combinación de ambos, crecimiento y propagación, puede ser nombrada como proliferación. Los ejemplos de condiciones seleccionadas y/o controladas incluyen el uso de un medio definido (con características conocidas como pH, fuerza iónica, y fuente de carbono), y condiciones especificas de temperatura, tensión de oxígeno, niveles de dióxido de carbono, y crecimiento en un biorreactor. Cultivar no se refiere al crecimiento o propagación de microorganismos en la naturaleza o en otra parte sin la intervención humana; por ejemplo, no es un cultivo el crecimiento natural de un organismo que finalmente se convierte en un fósil para producir el petróleo.

“Citólisis”, es la lisis de células en un entorno hipotónico. La citólisis es provocada por la ósmosis excesiva, o por el movimiento del agua, hacia el interior de una célula (hiperhidratación). La célula no puede resistir la presión osmótica del agua en su interior, y explota.

“Comida deslipidada” y “biomasa microbiana deslipidada”, es la biomasa microbiana después que el aceite (incluyendo los lípidos) se extraen o aíslan de ella, ya sea mediante el uso de extracción mecánica (es decir, ejercida por una prensa) o extracción por solvente, o ambos. La comida deslipidada tiene una cantidad reducida de aceites/lípidos comparado con la biomasa microbiana antes de la extracción o aislamiento de los aceites/lípidos, pero contiene alguna cantidad residual de aceites/lípidos.

“Vector de expresión” o “construcción de expresión” o “plásmido” o “construcción de ADN recombinante”, se refieren a un ácido nucleico generado mediante la intervención humana, incluyendo medios recombinantes o de síntesis química directa, con una serie de ácidos nucleicos específicos que permiten la transcripción y/o traducción de un ácido nucleico particular en una célula huésped. El vector de expresión puede ser parte de un plásmido, virus, o fragmento de ácido nucleico. Típicamente, el vector de expresión incluye un ácido nucleico para ser transcrito operativamente unido a un promotor.

“Gen exógeno”, es un ácido nucleico que codifica para la expresión de un ARN y/o proteína que se ha introducido (“transformado”) en una célula. Una célula transformada puede definirse como una célula recombinante, dentro de la cual se puede introducir un gen(es) exógeno(s) adicional(es). El gen exógeno puede ser de una especie diferente (heterólogo), o de la misma especie (homólogo) con respecto a la célula que se transforma. Así, un gen

exógeno puede incluir un gen homólogo que ocupa una localización diferente en el genoma de la célula o está bajo un control diferente, con respecto a la copia endógena del gen. Un gen exógeno puede estar presente en más de una copia en la célula. Un gen exógeno puede mantenerse en una célula como una inserción en el genoma o como una molécula episomal.

"Proporcionada exógenamente", se refiere a una molécula proporcionada al medio de cultivo de un cultivo celular.

"Prensado con expulsor", es un método mecánico para extraer el aceite de materias primas como la soya y las semillas de colza. Una prensa expulsora es una máquina de tornillo que presiona el material a través de una cavidad enjaulada en forma de barril. La materia prima entra por un lado de la prensa y la torta sale por el otro lado, mientras el aceite se filtra entre las barras de la jaula y se recoge. La máquina usa fricción y presión continua de las unidades de tornillo para mover y comprimir la materia prima. El aceite sale a través de pequeñas aberturas que no permiten que pasen los sólidos. Como la materia prima se prensa, típicamente la fricción provoca calentamiento.

La "acil graso-ACP tioesterasa", es una enzima que cataliza la escisión de un ácido graso de una proteína portadora de acilos (ACP) durante la síntesis de los lípidos.

La "acil graso-CoA/aldehído reductasa", es una enzima que cataliza la reducción de una molécula de acil-CoA a un alcohol primario.

La "acil graso-CoA reductasa", es una enzima que cataliza la reducción de una molécula de acil-CoA a un aldehído.

La "aldehído graso decarbonilasa", es una enzima que cataliza la conversión de un aldehído graso a un alcano.

La "aldehído graso reductasa", es una enzima que cataliza la reducción de un aldehído a un alcohol primario.

"Fuente fija de carbono", es una molécula o moléculas que contienen carbono, típicamente una molécula orgánica que está presente a temperatura y presión ambiente en forma sólida o líquida, que se puede utilizar por un microorganismo cultivado allí.

"Homogenato", es la biomasa que se ha roto físicamente.

"Hidrocarburo", es una molécula que contiene sólo átomos de hidrógeno y carbono, donde los átomos de carbono son covalentemente unidos para formar una cadena principal lineal, ramificada, cíclica, o parcialmente cíclica a la que se unen los átomos de hidrógeno. La estructura molecular de los compuestos hidrocarburos varía de la más simple, en forma de metano (CH_4), que es un constituyente del gas natural, a una muy fuerte y compleja, tales como algunas moléculas como los asfaltenos encontrados en el aceite crudo, petróleo, y betún. Los hidrocarburos pueden encontrarse en forma gaseosa, líquida, o sólida, o en cualquier combinación de estas formas, y pueden tener uno o más dobles o triples enlaces entre los átomos del carbono adyacentes a la cadena principal. Por lo tanto, el término incluye alcanos, alquenos, lípidos y parafinas lineales, ramificadas, cíclicas o parcialmente cíclicas. Los ejemplos incluyen propano, butano, pentano, hexano, octano, y escualeno.

"Relación hidrógeno:carbono", es la relación de átomos de hidrógeno con respecto a los átomos de carbono en una molécula sobre una base de átomo a átomo. La relación puede usarse para referirse al número de átomos de carbono y de hidrógeno en una molécula de hidrocarburo. Por ejemplo, el hidrocarburo con la relación más alta es el metano CH_4 (4:1).

"Fracción hidrofóbica", es la porción, o fracción de un material que es más soluble en una fase hidrófoba comparado con una fase acuosa. Un fragmento hidrófobo es substancialmente insoluble en agua y normalmente es no polar.

"Incremento del rendimiento lipídico", se refiere al aumento en la productividad de un cultivo microbiano, por ejemplo, incrementando el peso seco de células por litro de cultivo, aumentando el porcentaje de células que constituyen los lípidos, o incrementando la cantidad total de lípidos por litro de volumen de cultivo por unidad de tiempo.

El "promotor inducible", es un promotor que media la transcripción de los genes operativamente unidos en respuesta a un estímulo particular.

"En unión operativa", es la unión funcional entre dos secuencias de ácidos nucleicos, como una secuencia control (típicamente un promotor) y la secuencia unida (típicamente una secuencia que codifica a una proteína,

también llamada secuencia codificante). Un promotor está en una unión operativa con un gen exógeno si puede mediar la transcripción del gen.

"In situ", significa "en el lugar" o "en su posición original".

"Concentración limitante de un nutriente", es la concentración de un compuesto en un cultivo que limita la propagación de un organismo en el cultivo. Una concentración "no limitante de un nutriente" es una concentración que ayuda a la propagación máxima durante un período dado del cultivo. Así, el número de células producidas durante un período dado del cultivo es más bajo en presencia de una concentración limitante de un nutriente que cuando el nutriente es no limitante. Se dice que un nutriente está "en exceso" en un cultivo, cuando el nutriente está presente en una concentración mayor que la necesaria para la propagación máxima.

"Lipasa", es una enzima soluble en agua que cataliza la hidrólisis de ésteres unidos a sustratos lipídicos insolubles en agua. Las lipasas catalizan la hidrólisis de lípidos en gliceroles y ácidos grasos.

"Enzima modificante de lípidos", se refiere a una enzima que altera la estructura covalente de un lípido. Los ejemplos de enzimas modificantes de lípidos incluyen una lipasa, una acil graso-ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, y una aldehído graso decarbonilasa.

"Enzima de la vía de los lípidos", es cualquier enzima que juega un papel en el metabolismo de los lípidos, es decir, en la síntesis de los lípidos, modificación, o degradación, y cualquier proteína que modifica químicamente los lípidos, así como las proteínas portadoras.

Los "lípidos", son una clase de moléculas que son solubles en solventes no polares (como el éter y el cloroformo) y son relativa o completamente insolubles en agua. Las moléculas de lípidos tienen estas propiedades porque están constituidas por largas colas de hidrocarburos que son de naturaleza hidrofóbica. Los ejemplos de lípidos incluyen los ácidos grasos (saturados e insaturados); glicéridos o glicerolípidos (como monoglicéridos, diglicéridos, triglicéridos o grasas neutras, y fosfoglicéridos o glicerofosfolípidos); no glicéridos (esfingolípidos, esterol lípidos incluyendo colesterol y hormonas esteroideas, lípidos de prenol incluyendo terpenoides, alcoholes grasos, ceras, y poliquétidos); y derivados de lípidos complejos (lípidos unidos a azúcares, o glicolípidos, y lípidos unidos a proteínas). Las "grasas" son un subgrupo de lípidos denominados "triacilglicéridos".

"Lisado", es una solución que contiene el contenido de células lisadas.

"Lisis", es la rotura de la membrana plasmática y opcionalmente de la pared celular de un organismo biológico suficiente para liberar al menos algún contenido intracelular, a menudo por mecanismos mecánicos, virales u osmóticos que comprometen su integridad.

"Lisar", es la perturbación de la membrana celular y opcionalmente la pared celular de un organismo biológico o célula suficiente para liberar al menos algún contenido intracelular.

"Microalga", es un organismo microbiano eucariota que contiene un cloroplasto o plastidio, y opcionalmente es capaz de realizar la fotosíntesis, o un organismo microbiano procariota que es capaz de realizar la fotosíntesis. Las microalgas incluyen los fotoautótrofos obligados, que no pueden metabolizar una fuente fija de carbono en energía, como lo hacen los heterótrofos, los cuales pueden vivir únicamente a costa de una fuente fija de carbono. Las microalgas incluyen organismos unicelulares que se separan de células hermanas poco tiempo después de la división celular, tales como las *Chlamydomonas*, así como microbios, tales como, por ejemplo, *Volvox*, que es un microbio multicelular fotosintético simple de dos tipos de células diferentes. Las microalgas incluyen células tales como *Chlorella*, *Dunaliella*, y *Prototheca*. Las microalgas también incluyen otros organismos fotosintéticos microbianos que exhiben adhesión célula-célula, tales como *Agmenellum, Anabaena* y *Pyrobotrys*. Las microalgas también incluyen microorganismos heterótrofos obligados que perdieron la capacidad de realizar la fotosíntesis, como ciertas especies de algas dinoflageladas y especies del género *Prototheca*.

"Microorganismo" y "microbio", son organismos microscópicos unicelulares.

"Co-expresados naturalmente", con referencia a dos proteínas o genes significa que las proteínas o sus genes se co-expresan de forma natural en un tejido u organismo del que se derivan, por ejemplo, debido a que los genes que codifican a las dos proteínas están bajo el control de una secuencia reguladora común o debido a que se expresan en respuesta al mismo estímulo.

"Choque osmótico", es la ruptura de las células en una solución después de una reducción repentina de la presión osmótica. El choque osmótico se induce a veces para liberar los componentes celulares de estas células a una solución.

"Enzima de degradación de los polisacáridos", es cualquier enzima capaz de catalizar la hidrólisis o

sacarificación, de cualquier polisacárido. Por ejemplo, las celulasas catalizan la hidrólisis de celulosa.

"Polisacáridos" o "glicanos", son los carbohidratos hechos de monosacáridos unidos por enlaces glucosídicos. La celulosa es un polisacárido que forma parte de ciertas paredes celulares de las plantas. La celulosa puede ser despolimerizada por enzimas para producir monosacáridos como la glucosa y la xilosa, así como disacáridos más grandes y oligosacáridos.

"Promotor", es una secuencia control de ácidos nucleicos que dirige la transcripción de un ácido nucleico. Como se usa en la presente, un promotor incluye necesariamente secuencias de ácidos nucleicos cercanas al sitio de inicio de la transcripción, como, por ejemplo, en el caso del promotor de la polimerasa de tipo II, un elemento TATA. El promotor también incluye, opcionalmente, potenciadores distales o elementos represores, que se pueden encontrar a varios miles de pares de bases del sitio de inicio de la transcripción.

"Recombinante", es una célula, ácido nucleico, proteína o vector, que se modificó debido a la introducción de un ácido nucleico exógeno o la alteración de un ácido nucleico nativo. Así, por ejemplo, las células recombinantes expresan genes que no se encuentran en las células nativas (no recombinantes) o expresan genes nativos de manera diferente a como estos genes son expresados por una célula no recombinante. Un "ácido nucleico recombinante" es un ácido nucleico formado originalmente in vitro, en general, por la manipulación de ácidos nucleicos, por ejemplo, usando polimerasas y endonucleasas, o de otra manera están en una forma que no se encuentra normalmente en la naturaleza. Los ácidos nucleicos recombinantes se pueden producir, por ejemplo, para colocar dos o más ácidos nucleicos en unión operativa. Así, un ácido nucleico aislado o un vector de expresión formado in vitro por la ligadura de moléculas de ADN que normalmente no se unen en la naturaleza, se consideran recombinantes para los propósitos de la presente invención. Una vez que se obtiene un ácido nucleico recombinante y se introduce en una célula o en un organismo huésped, este se puede replicar usando la maquinaria celular in vivo de la célula huésped; sin embargo, estos ácidos nucleicos, una vez producidos por métodos recombinantes, aunque posteriormente se repliquen intracelularmente, siguen considerándose recombinantes para los propósitos de la esta invención. Del mismo modo, una "proteína recombinante" es una proteína producida usando técnicas de recombinación, es decir, a través de la expresión de un ácido nucleico recombinante.

"Diesel renovable", es una mezcla de alcanos (como C10:0, C12:0, C14:0, C16:0 y C18:0) producidos mediante la hidrogenación y desoxigenación de los lípidos.

"Sacarificación", es un proceso de conversión de biomasa, por lo general biomasa celulósica o lignocelulósica, en azúcares monoméricos, como glucosa y xilosa. El material celulósico o biomasa "sacarificada" o "despolimerizada" se refiere al material celulósico o a la biomasa que se convirtió en azúcares monoméricos a través de la sacarificación.

"Sonicación", es el proceso de ruptura de materiales biológicos, tales como una célula, mediante el uso de la energía de las ondas de sonido.

"Especies de furfural", es 2-furancarboxaldehído o un derivados que conserva las mismas características estructurales básicas.

"Rastrojo", son los tallos y hojas secas de un cultivo que quedaron después de que el grano se cosecha.

"Gen de utilización de la sacarosa", es un gen que, cuando se expresa, ayuda a la capacidad de una célula para utilizar la sacarosa como fuente de energía. Las proteínas codificadas por un gen de utilización de la sacarosa se refieren en la presente como "enzimas de utilización de la sacarosa" e incluyen los transportadores de sacarosa, sacarosa invertasas, y hexoquinasas como las glucoquinasas y las fructoquinasas.

II. CULTIVO

La presente invención se relaciona, en general, con el cultivo de variedades de *Prototheca*, en particular con variedades de *Prototheca* recombinantes, para la producción de lípidos. Por conveniencia para el lector, esta sección se subdivide en subsecciones. La Subsección 1 describe las especies *Prototheca* y sus cepas, y cómo identificar nuevas especies de *Prototheca* y sus cepas y microalgas relacionadas por comparación de ADN genómico. La Subsección 2 describe los biorreactores útiles para el cultivo. La Subsección 3 describe los medios de cultivo. La Subsección 4 describe la producción de aceite de acuerdo con los métodos de cultivo ilustrativos de la invención.

1. <u>Especies y cepas de Prototheca</u>

La *Prototheca* es un microorganismo notable para su uso en la producción de lípidos, debido a puede producir altos niveles de lípidos, particularmente lípidos adecuados para la producción de combustible. Los lípidos producidos por *Prototheca* tienen cadenas de hidrocarburos de menor longitud y un mayor grado de saturación que

los producidos por otras microalgas. Además, los lípidos de *Prototheca* están generalmente libres de pigmentos (niveles bajos a indetectables de clorofila y ciertos carotenoides) y en todo caso contienen mucho menos pigmentos que lípidos de otras microalgas. Además, las células de *Prototheca* recombinantes proporcionadas por la invención se pueden usar para producir lípidos con mayor rendimiento y eficiencia, y a un costo reducido, en relación con la producción de lípidos a partir de otros microorganismos. Las cepas ilustrativas de Prototheca para su uso en los métodos de la invención incluyen Prototheca moriformis, Prototheca krugani, Prototheca stagnora o Prototheca zopfii. Además, estas microalgas crecen heterotróficamente y pueden ser modificadas mediante ingeniería genética como *Prototheca wickerhamii, Prototheca stagnora (incluyendo UTEX 327), Prototheca portoricensis, Prototheca moriformis (incluyendo cepas de UTEX 1441, 1435),* y *Prototheca zopfii*. Las especies del género *Prototheca* son heterótrofas obligadas.

Las especies de *Prototheca* usadas en la invención se pueden identificar por amplificación de ciertas regiones diana del genoma. Por ejemplo, la identificación de una especie o variedad específica de *Prototheca* puede lograrse por amplificación y secuenciación de ADN nuclear y/o de cloroplastos usando cebadores y metodología que usa cualquier región del genoma, por ejemplo, usando los métodos descritos en Wu *y otros*, *Bot. Bull. Acad. Sin.* (2001) 42:115-121. Identification of *Chlorella spp*. isolates using ribosomal DNA sequences. Los métodos bien establecidos para el análisis filogenético, tales como la amplificación y secuenciación del espaciador transcrito interno ribosomal (ITS1 y ITS2 ADNr), ARNr 23S, ARNr 18S, y otras regiones conservadas del genoma, se pueden usar por aquellos expertos en la materia para identificar especies no solo de *Prototheca* sino también otros organismos productores de lípidos e hidrocarburos con perfiles lipídicos y capacidad de producción similares. Para ejemplos de métodos de identificación y clasificación de algas ver además, por ejemplo, *Genetics*, agosto de 2005; 170(4): 1601-10 y *RNA*, abril de 2005;11(4):361-4.

Así, la comparación de ADN genómico puede usarse para identificar especies adecuadas de microalgas que se usan en la presente invención. Las regiones de ADN genómico conservadas, tales como pero sin limitarse al ADN que codifica para ARNr 23S, se puede amplificar a partir de especies de microalgas y comparar con secuencias consenso para seleccionar especies de microalgas que estén taxonómicamente relacionadas con las microalgas preferidas usadas en la presente invención. Abajo se muestran los ejemplos de la comparación de las secuencias de ADN para especies del género *Prototheca*. La comparación del ADN genómico también puede ser útil para identificar especies de microalgas que han sido mal identificadas en un colección de cepas. A menudo, una colección de cepas identificará especies de microalgas en base a sus características fenotípicas y morfológicas. El uso de estas características puede conllevar a la categorización incorrecta de las especie o del género de una microalga. El uso de la comparación del ADN genómico puede ser un mejor método de categorización de especies de microalgas basado en su relación filogenética.

Las microalgas para usar en la presente invención tienen, típicamente, secuencias de ADN genómico que codifican para ARNr 23S con al menos 99%, al menos 95%, al menos 90%, o al menos 85% de identidad de nucleótidos con al menos una de las secuencias enumeradas en las SEQ ID NO: 11-19.

Para la comparación de secuencia para determinar el porcentaje de identidad de nucleótidos o de aminoácidos, típicamente una secuencia actúa como secuencia de referencia, contra la cual se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y de referencia se introducen en un ordenador, si es necesario se designan coordenadas posteriores, y se designan los parámetros del programa del algoritmo de secuencia. Después, el algoritmo de comparación de secuencias calcula el porcentaje de identidad de secuencia para la(s) secuencia(s) de prueba(s) con relación a la secuencia de referencia, basado en los parámetros de programa designados.

La alineación óptima de las secuencias para la comparación puede realizarse, *por ejemplo,* por el algoritmo de homología local de Smith & Waterman, *Adv. Sol. Math.* 2:482 (1981), por el algoritmo de alineamiento de homología de Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), por la búsqueda por el método de similitud de Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988), por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA del paquete de software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o por inspección visual (véase generalmente Ausubel y otros*, supra*).

Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad y similitud de secuencia es el algoritmo BLAST, el cual se describe en Altschul y otros*, J. Mol. Biol*. 215:403-410 (1990). El software para ejecutar los análisis BLAST está disponible públicamente a través del Centro Nacional para la Información Biotecnológica (en el sitio web www.ncbi.nlm.nih.gov). Este algoritmo implica, en primer lugar, identificar pares de secuencias de alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia problema, las cuales combinan o satisfacen alguna puntuación umbral mayor que el parámetro T, cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. El parámetro T se refiere como el umbral de puntuación de la palabra vecina (Altschul *y otros*, *supra*). Los aciertos de las palabras vecinas se comportan como semillas para iniciar las búsquedas y encontrar HSP más largos que las contengan. Después, los aciertos de las palabras vecinas se extienden en ambas direcciones a lo largo de cada secuencia hasta que se incremente la puntuación de alineamiento acumulativa. Para las secuencias nucleotídicas, las puntuaciones de

alineamiento acumulativas se calculan usando los parámetros M (puntuación recompensa para un par de residuos que combinan; siempre > 0) y N (puntuación castigo para residuos que no combinan; siempre < 0). Para las secuencias de aminoácidos, la puntuación acumulativa se calcula usando una matriz de puntuación. Las extensiones de los aciertos de las palabras vecinas en cada dirección se detienen cuando: la puntuación de alineamiento acumulativa disminuye en una cantidad X desde su máximo valor alcanzado; la puntuación acumulativa va a cero o por debajo de cero debido a la acumulación de uno o más alineamientos de residuos con puntuación negativa; o se alcanza el extremo de cualquiera de las secuencias. Para identificar si el ácido nucleico o el polipéptido está dentro del alcance de la invención se adecúan los parámetros definidos en los programas BLAST. El programa BLASTN (para las secuencias nucleotídicas) usa como defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4, y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP usa como defecto una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62. El programa TBLATN (usando secuencias de proteínas para secuencia de nucleótidos) usa como defecto una longitud de palabra (W) de 3, una expectativa (E) de 10 y la matriz de puntuación BLOSUM 62 (véase Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

Además de calcular el porcentaje de identidad de la secuencia, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, *por ejemplo*, Karlin & Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-5787 (1993)). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de la suma más pequeña (P(N)), que proporciona una indicación de la probabilidad con que dos secuencias de nucleótidos o de aminoácidos coincidan al azar. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de la suma más pequeña en una comparación entre el ácido nucleico del ensayo y uno de referencia es menor de aproximadamente 0.1, con más preferencia si es menor de aproximadamente 0.01 y con mayor preferencia si es menor de aproximadamente 0.001.

Otras consideraciones que afectan la selección de los microorganismos para su uso en la invención incluyen, además de la producción de lípidos o hidrocarburos adecuados para la producción de aceites, los combustibles y oleoquímicos: (1) alto contenido de lípidos como porcentaje del peso celular; (2) facilidad de crecimiento; (3) facilidad de manipulación por ingeniería genética; y (4) facilidad de procesamiento de la biomasa. En realizaciones particulares, el microorganismo salvaje o modificado mediante ingeniería genética produce células que contienen al menos 40%, al menos 45%, al menos 50%, al menos 55%, al menos 60%, al menos 65%, o al menos 70% o más de lípidos. Los organismos preferidos crecen heterotróficamente (en azúcar en ausencia de luz).

2. <u>Biorreactor</u>

Los microorganismos se cultivan con dos propósitos, el de realizar manipulaciones genéticas y para la producción de hidrocarburos (por ejemplo, lípidos, ácidos grasos, aldehídos, alcoholes y alcanos). El primer tipo de cultivo se realiza a pequeña escala e inicialmente, al menos, bajo condiciones en las que el microorganismo inicial pueda crecer. Los cultivos para la producción de hidrocarburos se realizan usualmente a gran escala (por ejemplo, 10,000 l, 40,000 l, 100,000 l o biorreactores más grandes) en un biorreactor. La *Prototheca* se cultiva, típicamente, en los métodos de la invención en medio líquido en un biorreactor. Típicamente, el biorreactor no permite la entrada de la luz.

El biorreactor o fermentador se usa para cultivar células de microalgas durante las varias fases de su ciclo fisiológico. Los biorreactores ofrecen muchas ventajas para su uso en métodos de crecimiento y propagación heterotrófica. Para producir biomasa para su uso en alimentos, las microalgas se fermentan preferiblemente en grandes cantidades de líquido, tales como cultivos en suspensión como se muestra en un ejemplo. Los biorreactores tales como los fermentadores de acero pueden alojar volúmenes muy grandes de cultivo (se usan biorreactores de 40,000 litros y más de capacidad en varias realizaciones de la invención). Los biorreactores típicamente permiten, además, el control de las condiciones de cultivo tales como temperatura, pH, tensión de oxígeno, y niveles de dióxido de carbono. Por ejemplo, los biorreactores se configuran típicamente, por ejemplo usando puertos unidos a tuberías, para permitir que los componentes gaseosos como el oxígeno o el nitrógeno sean burbujeados a través del líquido de cultivo. Otros parámetros del cultivo, tales como el pH del medio de cultivo, la identidad y concentración de los oligoelementos, y otros constituyentes del medio, se pueden manipular fácilmente usando el biorreactor.
Los biorreactores se pueden configurar para que el medio de cultivo fluya a través del biorreactor durante todo el período de tiempo en que las microalgas se reproducen e incrementan su número. En algunas realizaciones, por ejemplo, el medio puede ser infundido dentro del biorreactor después de la inoculación, pero antes que las células alcancen la densidad deseada. En otros casos, el biorreactor se llena con el medio de cultivo al inicio de un cultivo, y no se infunde más medio de cultivo después que el cultivo es inoculado. En otras palabras, la biomasa de microalgas se cultiva en un medio acuoso durante un período de tiempo en el que las microalgas se reproducen e incrementan su número; sin embargo, durante este período de tiempo las cantidades del medio de cultivo acuoso no fluyen a través del biorreactor. Así, en algunas realizaciones el medio de cultivo acuoso no fluye a través del biorreactor después de la inoculación.

Los biorreactores equipados con dispositivos tales como aspas de agitación e impulsores, mecanismos de balanceo, barras de agitación, medios para infundir gas presurizado, se pueden usar para mezclar los cultivos de

microalgas. La mezcla puede ser continua o intermitente. Por ejemplo, en algunas realizaciones, un régimen de flujo turbulento de entrada de gas y entrada del medio no se mantiene para la reproducción de las microalgas hasta que se haya alcanzado el incremento deseado en número de dichas microalgas.

Los puertos de los biorreactores se pueden usar para introducir, o extraer, gases, sólidos, semisólidos y líquidos, de la cámara del reactor que contiene las microalgas . Aunque muchos biorreactores tienen más de un puerto (por ejemplo, uno para la entrada del medio y otro para el muestreo), no es necesario que solamente una sustancia entre o salga por el puerto. Por ejemplo, un puerto se puede usar para hacer fluir el medio de cultivo dentro del reactor y posteriormente para el muestreo, entrada de gas, salida de gas, u otros propósitos. Preferiblemente, un puerto de muestreo se puede usar repetidamente sin que se altere de forma comprometedora la naturaleza axénica del cultivo. Un puerto de muestreo puede configurarse con una válvula u otro dispositivo que permita detener o iniciar la circulación de la muestra o proporcionar un medio de muestreo continuo. Típicamente, los biorreactores tienen al menos un puerto que permite la inoculación de un cultivo y dicho puerto puede usarse, además, con otros fines tales como la entrada de medio o gas.

Los puertos de los biorreactores permiten manipular el contenido gaseoso de los cultivos de microalgas. Para ilustrar, parte del volumen de un biorreactor puede ser gas en lugar de líquido, y las entradas de gas del biorreactor permiten el bombeo de gases hacia dentro del biorreactor. Los gases que se pueden bombear beneficiosamente dentro de un biorreactor incluyen aire, mezclas de aire/CO_2, gases nobles, tales como argón y otros gases. Los biorreactores están típicamente equipados para permitir que el usuario controle la velocidad de entrada del gas en el biorreactor. Como se denotó anteriormente, al aumentar el flujo de gas dentro del biorreactor aumenta la mezcla del cultivo.

El aumento del flujo gaseoso afecta también la turbidez del cultivo. La turbulencia se puede lograr colocando un puerto de entrada de gas debajo del nivel del medio de cultivo acuoso, de manera tal que el gas que entre al biorreactor burbujee hacia la superficie del cultivo. Uno o mas puertos de salida de gas permiten que el gas escape, impidiendo de este modo que la presión aumente gradualmente en el biorreactor. Preferiblemente, el puerto de salida de gas conduce hacia una válvula “en una sola dirección” que impide que microorganismos contaminantes entren al biorreactor.

3. <u>Medio</u>

El medio de cultivo de microalgas contiene, típicamente, una fuente fija de nitrógeno, una fuente fija de carbono, oligoelementos, opcionalmente un tampón para mantener el pH, y fosfato (típicamente proporcionado como una forma de sal fosfato). Otros componentes pueden incluir sales tales como cloruro de sodio, particularmente para microalgas marinas. Las fuentes de nitrógeno incluyen fuentes de nitrógeno orgánico e inorgánico, incluyendo, por ejemplo, sin limitación, nitrógeno molecular, nitrato, sales de nitrato, amoniaco (puro o en forma de sal, como $(NH_4)_2SO_4$ y NH_4OH), proteínas, harina de soya, licor de maíz fermentado, y extracto de levadura. Los ejemplos de oligoelementos incluyen zinc, boro, cobalto, cobre, manganeso, y molibdeno en, por ejemplo, las formas respectivas de $ZnCl_2$, H_3BO_3, $CoCl_2 \cdot 6H_2O$, $CuCl_2 \cdot 2H_2O$, $MnCl_2 \cdot 4H_2O$ y $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$.

De acuerdo con los métodos de la presente invención los microorganismos útiles se encuentran en varios lugares y ambientes a través del mundo. Como consecuencia de su aislamiento de otras especies y su divergencia evolutiva resultante, el medio de crecimiento particular para un crecimiento óptimo y generación de constituyentes de lípidos y/o hidrocarburos puede ser difícil de predecir. En algunos casos, ciertas cepas de microorganismos pueden ser incapaces de crecer en un medio particular de crecimiento debido a la presencia de algún componente inhibitorio o a la ausencia de un requerimiento nutricional esencial necesario por una cepa particular del microorganismo.

Los medios de crecimiento sólidos y líquidos generalmente están disponibles de una amplia variedad de fuentes, y las instrucciones para la preparación de un medio en particular que sea adecuado para una amplia gama de cepas de microorganismos, pueden encontrarse, por ejemplo, en línea en el sitio web http://www.utex.org/, un sitio mantenido por la Universidad de Texas en Austin, 1 University Station A6700, Austin, Texas, 78712-0183, para su colección de cultivo de algas (UTEX). Por ejemplo, varios medios de agua dulce y agua salada incluyen los descritos en la publicación PCT núm. 2008/151149, incorporada en la presente como referencia.

En un ejemplo particular, el medio Proteosa es adecuado para cultivos axénicos, y un volumen de 1l del medio (pH~6.8) puede prepararse por adición de 1g de proteosa peptona a 1 litro de Medio Bristol. El medio Bristol comprende 2.94 mM de $NaNO_3$, 0.17 mM de $CaCl_2 \cdot 2H_2O$, 0.3 mM de $MgSO_4 \cdot 7H_2O$, 0.43 mM, 1.29 mM de KH_2PO_4, y 1.43 mM de NaCl en una solución acuosa. Para 1.5% de medio agar, se pueden adicionar 15 g de agar a 1 l de la solución. La solución se cubre y se somete a autoclave, y después se almacena a temperatura refrigerada antes del uso. Otro ejemplo es el medio de aislamiento de *Prototheca* (PIM), que comprende 10g/l de hidrógeno ftalato de potasio (KHP), 0.9g/l de hidróxido de sodio, 0.1g/l de sulfato de magnesio, 0.2g/l de hidrógeno fosfato de potasio, 0.3g/l de cloruro de amonio, 10g/l de glucosa, 0.001g/l de hidrocloruro de tiamina, 20g/l de agar, 0.25g/l de 5-fluorocitosina, en un intervalo de pH de 5.0 a 5.2 (véase Pore, 1973, App. Microbiology, 26: 648-649). Otros medios adecuados para su uso con los métodos de la invención pueden ser fácilmente identificados consultando la URL

identificada anteriormente, o consultando otras organizaciones que mantienen cultivos de microorganismos, tales como SAG, CCAP, o CCALA. SAG se refiere a la Colección de Cultivo de Algas en la Universidad de Göttingen (Göttingen, Alemania), CCAP se refiere a la colección de cultivo de algas y protozoos dirigida por la Asociación Escocesa para las Ciencias Marinas (Escocia, Reino Unido), y CCALA se refiere al laboratorio de colección de cultivo de algas en el Instituto de Botánica (Třeboň, República Checa). Adicionalmente, la patente de Estados Unidos núm. 5,900,370 describe formulaciones de medios y condiciones adecuadas para la fermentación heterótrofa de las especies de *Prototheca*.

Para la producción de aceite, es importante la selección de una fuente fija de carbono, ya que como el costo de la fuente fija de carbono debe ser lo suficientemente bajo para que la producción de aceite resulte económica. Así, aunque las fuentes de carbono adecuadas incluyen, por ejemplo, acetato, floridosida, fructosa, galactosa, ácido glucurónico, glucosa, glicerol, lactosa, manosa, N-acetilglucosamina, ramnosa, sacarosa, y/o xilosa, la selección de materias primas que contengan dichos compuestos es un aspecto importante de los métodos de la invención. Las materias primas adecuadas, útiles de acuerdo con los métodos de la invención incluyen, por ejemplo, licor negro, almidón de maíz, material celulósico despolimerizado, suero de leche, melaza, papa, sorgo, sacarosa, azúcar de remolacha, caña de azúcar, arroz, y trigo. Las fuentes de carbono pueden proporcionarse además como una mezcla de sacarosa y pulpa de azúcar de remolacha despolimerizada. Una o más fuentes de carbono pueden proporcionarse a una concentración de al menos aproximadamente 50 □M, al menos aproximadamente 100 □M, al menos aproximadamente 500 □M, al menos aproximadamente 5mM, al menos aproximadamente 50 mM, y al menos aproximadamente 500 mM de una o más fuentes fijas de carbono proporcionadas exógenamente. Las fuentes de carbono de particular interés para los propósitos de la presente invención incluyen, celulosa (en forma despolimerizada), glicerol, sacarosa, y sorgo, cada uno éstos se discute abajo con más detalle.

De acuerdo con la presente invención, los microorganismos se pueden cultivar usando biomasa celulósica despolimerizada como una materia prima. La biomasa celulósica (por ejemplo, rastrojo como rastrojo de maíz) es barata y fácilmente disponible; sin embargo los intentos de usar este material como una materia prima para la levadura han fracasado. En particular, se ha encontrado que estas materias primas tienen un efecto inhibitorio en el crecimiento de la levadura, y la levadura no puede usar los azúcares de 5 carbonos producida a partir de materiales celulósicos (por ejemplo, xilosa a partir de hemicelulosa). En contraste, las microalgas pueden crecer en material celulósico procesado. Los materiales celulósicos incluyen, por lo general, aproximadamente 40-60% de celulosa, aproximadamente 20-40% de hemicelulosa; y 10-30% de lignina.

Los materiales celulósicos adecuados incluyen los residuos de cosechas energéticas herbáceas y leñosas, así como cultivos agrícolas, es decir, partes de plantas, principalmente tallos y hojas no eliminadas de los campos con el alimento principal o el producto de fibra. Los ejemplos incluyen desechos agrícolas tales como el bagazo de caña de azúcar, cascarilla de arroz, fibra de maíz (incluyendo tallos, hojas, cáscaras y mazorcas), paja del trigo, paja del arroz, pulpa del azúcar de remolacha, pulpa de cítricos, cáscaras de cítricos, desechos forestales tales como madera y entresacas de madera blanda, madera dura y blanda y desechos de las operaciones de la madera, desechos de madera tales como desechos de las fábricas (astillas de madera, aserrín) y desechos de plantas de celulosa, desechos urbanos, tales como pedazos de papel de desechos sólidos urbanos, desechos urbanos de madera y desechos urbanos verdes tales como los recortes de césped municipales y desechos de construcción de madera. Los materiales celulósicos adicionales incluyen cosechas celulósicas dedicadas tales como pasto varilla, madera híbrida de álamo, miscantus, fibras de caña, y fibras de sorgo. Los azúcares de 5 carbonos que se producen a partir de estos materiales incluyen la xilosa.

Los materiales celulósicos se tratan para aumentar la eficiencia con la que el microbio puede utilizar el(los) azúcar(es) contenidos en los materiales. La invención proporciona nuevos métodos para el tratamiento de materiales celulósicos después de la explosión con ácidos para que los materiales sean adecuados para usar en un cultivo heterótrofo de microbios (por ejemplo, microalgas y levaduras oleaginosas). Como se discutió anteriormente, la biomasa lignocelulósica está compuesta por varias fracciones, incluyendo celulosa, un polímero cristalino de enlaces beta 1,4-glucosídicos (un azúcar de 6 carbonos), hemicelulosa, un polímero más estrechamente asociado compuesto predominantemente de xilosa (un azúcar de 5 carbonos), y en menor cantidad manosa, galactosa, arabinosa, lignina, un polímero aromático complejo formado por alcohol sinapílico y sus derivados, y pectinas, las cuales son cadenas lineales de un ácido alfa poligalacturónico 1,4 enlazado. Debido a la estructura polimérica de la celulosa y la hemicelulosa, los azúcares (por ejemplo, glucosa y xilosa monomérica) en ellos no se encuentran en una forma tal que se puedan usar (metabolizar) de manera eficiente por muchos microbios. Para estos microbios, el procesamiento posterior de la biomasa celulósica para generar los azúcares monoméricos que componen los polímeros, puede ser de gran ayuda para asegurar que los materiales celulósicos se usen de manera eficiente como materia prima (fuente de carbono).

La biomasa celulósica o de celulosa se somete a un proceso, denominado “explosión”, en el cual la biomasa se trata con ácido sulfúrico (u otro) diluido a presión y temperatura elevadas. Este proceso condiciona la biomasa de manera que se pueda someter de manera eficiente a la hidrólisis enzimática de las fracciones celulósicas y hemicelulósicas en monómeros de glucosa y xilosa. Los azúcares monoméricos que se obtienen se denominan azúcares celulósicos. Los azúcares celulósicos pueden usarse posteriormente por los microorganismos

para producir varios metabolitos (por ejemplo, lípidos). La etapa de explosión con ácido da como resultado la hidrólisis parcial de la fracción de hemicelulosa para obtener monosacáridos. Estos azúcares se liberan completamente de la biomasa mediante un tratamiento posterior. En algunas realizaciones, este tratamiento consiste en un tratamiento hidrotérmico que incluye el lavado del material explotado con agua caliente, lo cual elimina contaminantes tales como las sales. Esta etapa no es necesaria para las fermentaciones celulósicas de etanol debido a que en esos procesos se usan concentraciones de azúcares más diluidas. En otras realizaciones, el tratamiento posterior es un tratamiento adicional con ácido. En aún otras realizaciones, el tratamiento posterior es la hidrólisis enzimática del material explotado. Estos tratamientos se pueden usar además en cualquier combinación. El tipo de tratamiento puede afectar el tipo de azúcar liberado (por ejemplo, azúcares de cinco carbonos versus azúcares de seis carbonos) y la etapa en la cual se liberan en el proceso. Como consecuencia, se pueden crear diferentes corrientes de azúcares, que pueden ser de cinco o de seis átomos de carbono. Estas corrientes enriquecidas de cinco o seis átomos de carbono, se pueden dirigir después a los microorganismos específicos con diferentes capacidades de utilización del carbono.

Los métodos de la presente invención típicamente implican la fermentación a densidades celulares más altas que las que se logran con la fermentación etanólica. Debido a las densidades más altas de los cultivos para la producción de aceites celulósicos heterótrofos, la fuente fija de carbono (por ejemplo, la(s) corriente(s) de azúcar de derivados celulósicos) está, preferiblemente, en forma concentrada. El nivel de glucosa del material celulósico despolimerizado es, preferiblemente, de al menos 300 g/litro, al menos 400 g/litro, al menos 500 g/litro o al menos 600 g/litro antes de la etapa de cultivo, el cual es opcionalmente un cultivo tipo lote alimentado, en el que las células se alimentan con el material mientras crecen y acumulan lípidos. En la producción de etanol celulósico, las corrientes de azucares celulósicos no se usan a ni cerca a estos intervalos de concentración. Así, para generar y mantener las altas densidades de la célula durante la producción del aceite lignocelulósico, la(s) materia(s) prima(s) de carbono se deben suministrar a los cultivos heterótrofos en forma muy concentrada. Sin embargo, cualquier componente en la corriente de alimentación que no sea sustrato y que no se metabolice por el microorganismo oleaginoso se acumulará en el biorreactor, lo que conduce a problemas si el componente es tóxico o tiene actividad inhibitoria sobre la producción del producto final deseado. Aunque la lignina, y los subproductos derivados de la lignina, subproductos derivados de carbohidratos, tales como, furfurales e hidroximetilfurfurales y las sales derivadas de la generación de los materiales celulósicos (durante el proceso de explosión y el proceso de neutralización posterior), e incluso los azúcares no metabolizados pentosa/hexosa, pueden presentar problemas en las fermentaciones etanólicas, estos efectos se amplifican significativamente en los procesos en los cuales su concentración en la materia prima inicial es alta. Para lograr concentraciones de azúcar en el intervalo de 300g/l (o superior) para los azúcares de seis carbonos que se pueden usar en producciones a gran escala de los aceites lignocelulósicos descritos en la presente invención, la concentración de estos materiales tóxicos puede ser 20 veces mayor que la concentración típicamente presente en las fermentaciones etanólicas de la biomasa celulósica.

El tratamiento con el proceso de explosión del material celulósico utiliza cantidades significativas de ácido sulfúrico, calor y presión, liberando de esta forma subproductos de carbohidratos, específicamente furfurales e hidroximetilfurfurales. Los furfurales e hidroximetilfurfurales se producen durante la hidrólisis de la hemicelulosa mediante la deshidratación de la xilosa en furfural y agua. En algunas realizaciones de la presente invención, estos subproductos (por ejemplo, furfurales e hidroximetilfurfurales) se eliminan del material lignocelulósico sacarificado antes de la introducción en el biorreactor. En algunas realizaciones de la presente invención, el proceso de eliminación de los subproductos de carbohidratos se realiza mediante tratamiento hidrotérmico de los materiales celulósicos explotados. Además, la presente invención proporciona métodos en los cuales las cepas capaces de tolerar compuestos tales como los furfurales e hidroximetilfurfurales se usan para la producción de aceites lignocelulósicos. En otra realización, la presente invención también proporciona métodos y microorganismos que no solo son capaces de tolerar los furfurales en el medio de fermentación, sino que son realmente capaces de metabolizar estos subproductos durante la producción del aceite lignocelulósico.

El proceso de explosión también genera niveles significativos de sales. Por ejemplo, las condiciones típicas de la explosión pueden dar como resultado conductividades superiores a los 5 mS/cm cuando la biomasa celulósica explotada se resuspende en una relación de 10:1 agua:sólidos (peso seco). En ciertas realizaciones de la presente invención, la biomasa explotada diluida se somete a sacarificación enzimática, y el sobrenadante obtenido se concentra hasta 25 veces para su uso en el biorreactor. El nivel de sal (medido por conductividad) en la(s) corriente(s) de azúcar concentrada, puede ser inaceptablemente alto (hasta 1.5 M Na^+ equivalentes). Las sales adicionales se generan con la neutralización de los materiales explotados para el proceso de sacarificación enzimática posterior. La presente invención proporciona métodos para eliminar estas sales de manera que la(s) corriente(s) de azúcar celulósica concentrada resultante se pueda usar en los procesos heterótrofos para producir aceite lignocelulósico. En algunas realizaciones, el método de eliminación de estas sales es la desionización con resinas, tales como, pero sin limitarse a, DOWEX Marathon MR3. En ciertas realizaciones, la etapa de deionización con resina ocurre antes de la concentración del azúcar o el ajuste del pH y el tratamiento hidrotérmico de la biomasa antes de la sacarificación, o cualquier combinación del procedimiento; en otras realizaciones la etapa es llevada a cabo después de uno o más de estos procesos. En otras realizaciones, el proceso de explosión en sí se cambia para evitar la generación de sales a niveles altos inaceptables. Por ejemplo, una alternativa adecuada para la explosión de la biomasa celulósica con ácido sulfúrico (u otro ácido) es la fabricación de pasta mecánica para hacer que la

biomasa celulósica sea receptiva a la hidrólisis enzimática (sacarificación). En aún otras realizaciones, se usan cepas nativas de microorganismos resistentes a niveles altos de sal o cepas modificadas mediante ingeniería genética con resistencia a niveles altos de sal.

Una realización preferida para el proceso de preparación de la biomasa celulósica explotada para su uso en la producción de aceite lignocelulósico heterótrofo usando microbios oleaginosos se representa en la Figura 10. La etapa I comprende ajustar el pH de la biomasa celulósica explotada resuspendida en un intervalo de 5.0-5.3, seguido del lavado de la biomasa celulósica tres veces. Esta etapa de lavado se puede realizar usando una variedad de medios que incluyen el uso de resinas de desalado y de intercambio iónico, ósmosis inversa, tratamiento hidrotérmico (como el descrito anteriormente), o simplemente por re-suspensión repetida en agua desionizada y centrifugación. Esta etapa de lavado resulta en una corriente celulósica cuya conductividad se encuentra entre 100-300 µS/cm y la eliminación de cantidades significativas de furfurales e hidroximetilfurfurales. Lo decantado de esta etapa de lavado se puede guardar para concentrar los azúcares de cinco carbonos liberados de la fracción de hemicelulosa. La etapa II comprende la sacarificación enzimática de la biomasa celulósica lavada. En una realización preferida se usa Accellerase (Genencor). La etapa III comprende la recuperación de los azúcares mediante centrifugación o decantación y enjuague de la biomasa sacarificada. La biomasa resultante (sólidos) es un componente denso energético, rico en lignina que puede usarse como combustible o eliminar como desecho. La corriente de azúcar recuperada durante el proceso de centrifugación/decantado y enjuague se recoge. La etapa IV comprende la microfiltración para eliminar los contaminantes sólidos y recuperar el permeado. La etapa V comprende una etapa de concentración que se puede realizar usando un evaporador a vacío. Esta etapa puede incluir, opcionalmente, la adición de agente antiespumantes tales como el P'2000 (Sigma/Fluka), el cual se necesita algunas veces debido al contenido proteico de la materia prima de azúcar resultante.

En otra realización de los métodos de la invención, la fuente de carbono es glicerol, incluyendo los subproductos de glicerol acidulado y no acidulado obtenidos de la transesterificación del biodiesel. En una realización, la fuente de carbono incluye el glicerol yal menos otra fuente de carbono. En algunos casos, todo el glicerol y al menos una de las otras fuentes de fijas carbono se suministran al microorganismo al comienzo de la fermentación. En algunos casos, el glicerol y al menos una de las otras fuentes fijas de carbono se suministran al microorganismo simultáneamente en una proporción predeterminada. En algunos casos, el glicerol y al menos una de las otras fuentes fijas de carbono se alimentan a los microbios a una velocidad predeterminada durante el curso de la fermentación.

Algunas microalgas se someten a división celular de manera más rápida en presencia de glicerol que en presencia de glucosa (véase la publicación PCT núm. 2008/151149). En estos casos, los procesos de crecimiento en dos etapas en los que las células se alimentan primeramente con glicerol para aumentar rápidamente la densidad celular, y después se alimentan con glucosa para acumular lípidos, pueden mejorar la eficiencia con la que se producen los lípidos. El uso del subproducto de glicerol del proceso de transesterificación proporciona ventajas económicas significativas al ponerlo nuevamente en el proceso de producción. También se proporcionan otros métodos de alimentación tales como mezclas de glicerol y glucosa. La alimentación de tales mezclas proporciona, además, los mismos beneficios económicos. Además, la invención proporciona métodos para alimentar las microalgas con azúcares alternativos, tales como la sacarosa en varias combinaciones con glicerol.

En otra realización de los métodos de la invención, la fuente de carbono es sacarosa, incluyendo una materia prima compleja que contiene sacarosa, tal como el jugo espeso de la caña del procesamiento de la caña de azúcar. En una realización, el medio de cultivo incluye, además, al menos una enzima de utilización de la sacarosa. En algunos casos, el medio de cultivo incluye una sacarosa invertasa. En una realización, la enzima sacarosa invertasa es una enzima sacarosa invertasa secretable codificada por un gen de sacarosa invertasa exógeno expresado por la población de microorganismos. De esta manera, en algunos casos, como se describe más detalladamente en la Sección IV, abajo, las microalgas se modifican por ingeniería genética para que expresen una enzima de utilización de la sacarosa, tal como, un transportador de sacarosa , una sacarosa invertasa, una hexoquinasa, una glucoquinasa, o una fructoquinasa.

Las materias primas complejas que contienen sacarosa incluyen a los desechos de melaza del procesamiento de la caña de azúcar; el uso de estos productos de desecho del procesamiento de la caña de azúcar de bajo valor proporciona ahorros significativos de los costos de producción de hidrocarburos y otros aceites. El sorgo es otra materia prima compleja que contiene sacarosa que es útil en los métodos de la invención, incluyendo sirope de sorgo y sorgo puro. El sirope de sorgo se produce a partir del jugo de la caña dulce de sorgo. Su perfil de azúcares se compone fundamentalmente de glucosa (dextrosa), fructosa y sacarosa.

4. <u>Producción de aceite</u>

Para la producción de aceite de acuerdo con los métodos de la invención, es preferible cultivar las células en la oscuridad, como es el caso, por ejemplo, cuando se usan fermentadores extremadamente grandes (40.000 litros y más) que no permiten que la luz afecte al cultivo. Las especies de *Prototheca* crecen y se propagan, para la producción de aceite en un medio que contiene una fuente fija de carbono y en ausencia de luz; ese cultivo se

conoce como cultivo heterótrofo.

Como ejemplo, un inóculo de las células de microalgas productoras de lípidos se introduce en el medio; hay un período de latencia (fase lag) antes de que las células comiencen a propagarse. Tras el período de latencia, la velocidad de propagación aumenta a un ritmo constante y entra en la fase log o exponencial. La fase exponencial, a su vez, es seguida por una desaceleración de la propagación debido a la disminución de nutrientes como el nitrógeno, el aumento de sustancias tóxicas, y los mecanismos de detección de quórum. Después de esta desaceleración, se detiene la propagación y las células entran en una fase estacionaria o estado constante, dependiendo del ambiente particular proporcionado a las células. Para la obtención de biomasa rica en lípidos, el cultivo típicamente se cosecha bien después del final de la fase exponencial, lo que puede terminar tempranamente al permitir que el nitrógeno u otro nutriente clave (distinto del carbono) se agote, lo que obliga a las células a convertir en lípidos las fuentes de carbono presentes en exceso. Los parámetros de las condiciones de cultivo pueden manipularse para optimizar la producción total de aceite, la combinación de especies de lípidos producidos y/o la producción de un aceite específico.

Como se mencionó anteriormente, un biorreactor o fermentador se usa para permitir que las células se sometan a las diversas fases de su ciclo de crecimiento. Como ejemplo, se puede introducir un inóculo de las células productoras de lípidos en un medio, seguido por un período de latencia (fase lag) antes que las células comiencen a crecer. Tras el período de latencia, la velocidad de crecimiento aumenta constantemente y entra en la fase log o exponencial. La fase exponencial, a su vez, es seguida por una desaceleración del crecimiento debido a la disminución de nutrientes y/o el aumento de sustancias tóxicas. Después de esta desaceleración, el crecimiento se detiene, y las células entran en una fase estacionaria o estado constante, dependiendo del ambiente particular proporcionado a las células. La producción de lípidos por las células descrita en la presente puede ocurrir durante la fase exponencial o posteriormente, incluyendo la fase estacionaria en donde los nutrientes se suministran, o todavía están disponibles, para permitir la continuación de la producción de lípidos en ausencia de división celular.

Preferiblemente, los microorganismos que crecen usando las condiciones descritas en la presente y que se conocen en la materia, comprenden al menos aproximadamente 20% en peso de lípidos, preferiblemente al menos aproximadamente 40% en peso, con mayor preferencia al menos aproximadamente 50% en peso, y con la máxima preferencia al menos aproximadamente 60% en peso. Las condiciones del proceso se pueden ajustar para aumentar el rendimiento de los lípidos adecuados para un uso particular y/o para reducir los costos de producción. Por ejemplo, en ciertas realizaciones, una microalga se cultiva en presencia de una concentración limitante de uno o más nutrientes, como por ejemplo, nitrógeno, fósforo o azufre, mientras que se proporciona un exceso de energía fija de carbono como la glucosa. La limitación del nitrógeno tiende a aumentar el rendimiento de lípidos microbianos, por encima del rendimiento de lípidos microbianos en un cultivo en el que el nitrógeno se suministra en exceso. En realizaciones particulares, el aumento en el rendimiento de los lípidos es al menos aproximadamente: 10%, 50%, 100%, 200% o 500%. El microbio puede cultivarse en presencia de una cantidad limitante de un nutriente por una parte del período total de cultivo o por todo el período. En realizaciones particulares, la concentración de nutrientes se mueve entre una concentración limitante y una concentración no limitante, al menos dos veces durante el período total de cultivo. El contenido de lípidos de las células se puede aumentar al continuar el cultivo por períodos de tiempo mayores, mientras que se proporcione un exceso de carbono, limitando o no el nitrógeno.

En otra realización, el rendimiento de lípidos aumenta al cultivar un microbio productor de lípidos (por ejemplo, microalgas) en presencia de uno o más cofactor(es) para una enzima de la vía de los lípidos (por ejemplo, una enzima para la síntesis de ácidos grasos). En general, la concentración del(de los) cofactor(es) es suficiente para aumentar el rendimiento de lípidos microbianos (por ejemplo, ácidos grasos) por encima del rendimiento de lípidos microbianos en ausencia del (de los) cofactor(es). En una realización particular, el(los) cofactor(es) se proporcionan al cultivo mediante la inclusión de un microbio en el cultivo (por ejemplo, microalgas) que contiene un gen exógeno que codifica para el(los) cofactor(es) . Como alternativa, el(los) cofactor(es) se pueden proporcionar a un cultivo mediante la inclusión de un microbio (por ejemplo, microalgas) que contiene un gen exógeno que codifica para una proteína que participa en la síntesis del cofactor. En ciertas realizaciones, los cofactores adecuados incluyen cualquier vitamina requerida por una enzima de la vía de los lípidos, tales como, por ejemplo: biotina, pantotenato. Los genes que codifican para los cofactores adecuados para su uso en la invención o que participan en la síntesis de tales cofactores son bien conocidos y se pueden introducir en los microbios (por ejemplo, microalgas), usando construcciones y técnicas como las descritas anteriormente.

Los ejemplos específicos de biorreactores, condiciones de cultivo, y métodos de crecimiento y propagación heterótrofa descritos en la presente, se pueden combinar de manera adecuada para mejorar la eficiencia del crecimiento microbiano y la producción de lípidos y/o proteínas.

La biomasa de microalgas con un alto porcentaje de acumulación de aceites/lípidos por peso seco se ha generado usando diferentes métodos de cultivo, los que se conocen en la materia (véase la publicación PCT núm. 2008/151149). La biomasa de microalgas generada por los métodos de cultivo descritos en la presente y útiles de conformidad con la presente invención, comprende al menos 10% de aceite de microalgas por peso seco. En algunas realizaciones, la biomasa de microalgas comprende al menos 25%, al menos 50%, al menos 55%, o al

menos 60% de aceite de microalgas por peso seco. En algunas realizaciones, la biomasa de microalgas contiene de 10-90% de aceite de microalgas, de 25-75% de aceite de microalgas, de 40-75% de aceite de microalgas, o de 50-70% de aceite de microalgas por peso seco.

El aceite de microalgas de la biomasa descrita en la presente, o extraída de la biomasa para su uso en los métodos y composiciones de la presente invención, puede comprender glicerolípidos con una o más cadenas laterales de ácidos grasos esterificados distintos. Los glicerolípidos comprenden una molécula de glicerol esterificado para uno, dos o tres moléculas de ácidos grasos, que pueden ser de diferentes longitudes y tener diferentes grados de saturación. Las características de la longitud y la saturación de las moléculas de ácidos grasos (y los aceites de microalgas) pueden manipularse para modificar las propiedades o proporciones de las moléculas de ácidos grasos en los aceites de microalgas de la presente invención, a través de las condiciones de cultivo o por vía de la ingeniería genética de los lípidos, como se describe con más detalle a continuación en la Sección IV. Por lo tanto, se pueden preparar mezclas específicas de aceite de algas de una sola especie de algas, por mezcla de la biomasa o aceite de algas a partir de dos o más especies de microalgas, o por una mezcla de aceite de algas de la invención con aceites de otras fuentes tales como soya, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de vacuno, sebo porcino, microalgas, macroalgas, microbios, cufea, lino, cacahuate, grasa blanca selecta, manteca de cerdo, Camelina sativa, semillas de mostaza, nuez de anacardo, avena, lupino, kenaf, caléndula, cáñamo, café, linaza (lino), avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez, jojoba, macadamia, nueces del Brasil, aguacate, petróleo, o una fracción destilada de cualquiera de los aceites anteriores.

La composición del aceite, *es decir*, las propiedades y proporciones de los ácidos grasos que constituyen los glicerolípidos, también se puede manipular mediante la combinación de la biomasa o el aceite de al menos dos especies distintas de microalgas. En algunas realizaciones, al menos dos de las especies de microalgas distintas tienen diferentes perfiles de glicerolípidos. Las distintas especies de microalgas se pueden cultivar juntas o por separado, como se describe en en la presente, de preferencia bajo condiciones heterótrofas, para generar los aceites respectivos. Las diferentes especies de microalgas pueden contener distintos porcentajes de los distintos ácidos grasos que constituyen los glicerolípidos de la célula.

Generalmente, las cepas de *Prototheca* tienen muy poco o no tienen ácidos grasos con longitud de cadena C8-C14. Por ejemplo, *Prototheca moriformis* (UTEX 1435), *Prototheca krugani* (UTEX 329), *Prototheca stagnora* (UTEX 1442) y *Prototheca zopfii* (UTEX 1438) no contienen (o contienen cantidades indetectables) de ácidos grasos C8, entre 0-0.01% de ácidos grasos C10, entre 0.03-2.1% de ácidos grasos C12 y entre 1.0-1.7% de ácidos grasos C14.

En algunos casos, las variedades de *Prototheca* contienen un transgen que codifica para una acil graso-ACP tioesterasa que tiene actividad hacia los sustratos acil graso-ACP de longitud de cadena C8-10, que tienen al menos 0.3%, al menos 0.8%, al menos 1.5% o más ácidos grasos de longitud de cadena C8 y al menos 0.3%, al menos 1.0%, al menos 3.0%, al menos 5% o más ácidos grasos de longitud de cadena C10. En otros casos, las cepas de *Prototheca* contienen un transgen que codifica para una acil graso-ACP tioesterasa que tiene actividad sobre sustratos acil-ACP adiposos de longitud de cadena C12, que tiene al menos 3.0%, al menos 5%, al menos 7%, al menos 10% , al menos 13% o más ácidos grasos de longitud de cadena C12 y al menos 1.5%, al menos 2%, o al menos 3% o más ácidos grasos de longitud de cadena C14. En otros casos, las cepas de *Prototheca* contienen un transgen que codifica para una acil graso-ACP tioesterasa que tiene actividad hacia los sustratos acil graso-ACP de longitud de cadena C14, que tiene al menos 4.0%, al menos 7%, al menos 10%, al menos 15% , al menos 20%, al menos 25% o más ácidos grasos de longitud de cadena C14, y al menos 0.4%, al menos 1%, al menos 1.5%, o más ácidos grasos de longitud de cadena C12.

En los ejemplos no limitantes, las cepas de *Prototheca* contienen un transgen que codifica una acil graso-ACP tioesterasa que tiene actividad hacia los sustratos acil graso ACP de longitud de cadena C8 y C10, que tienen entre 0.3-1.58% de ácidos grasos de longitud de cadena C8 y entre 0.35-6.76% de ácidos grasos de longitud de cadena C10. En otros ejemplos no limitantes, las cepas de *Prototheca* contienen un transgen que codifica una acil graso-ACP tioesterasa que tiene actividad hacia el sustrato acil graso-ACP de longitud de cadena C12, que tiene entre 3.9-14.11% de ácidos grasos de longitud de cadena C12 y entre 1.95-3.05% de ácidos grasos de longitud de la cadena C14. En otros ejemplos no limitantes, las cepas de *Prototheca* contienen un transgen que codifica una acil graso-ACP tioesterasa que tiene actividad hacia los sustratos acil graso ACP de longitud de cadena C14, que tiene entre 4.40-17.35% de ácidos grasos de longitud de cadena C14 y entre 0.4-1.83% de ácidos grasos de longitud de cadena C12. En algunos casos, las cepas de *Prototheca* contienen un transgen que codifica una acil graso-ACP tioestera de longitud de cadena C8 que tiene actividad hacia los sustratos acil graso-ACP de longitud de cadena entre C8 y C14 que tienen entre 3.5-20% de ácidos grasos de cadena media (C8-C14). En algunos casos, mantener las cepas de *Prototheca* transgénicas bajo presión selectiva constante y elevada para conservar los genes exógenos, es ventajoso para aumentar los ácidos grasos deseados de una longitud de cadena específica. En un ejemplo no limitante, el Ejemplo 5 demuestra un incremento del doble de los ácidos grasos de longitud de cadena C14 (más de 30% de ácidos grasos de longitud de cadena C8-C14) cuando se reserva el cultivo de *Prototheca*

moriformis que contiene un gen exógeno de tioesterasa de preferencia C14. Los altos niveles de retención del gen exógeno también se pueden lograr mediante la inserción de genes exógenos en los cromosomas nucleares de las células usando la recombinación homóloga de vectores y los métodos descritos en la presente. Las células recombinantes que contienen genes exógenos integrados a los cromosomas nucleares son un objeto de la invención.

El aceite de microalgas también puede incluir otros constituyentes producidos por las microalgas, o incorporados en el aceite de microalgas a partir del medio de cultivo. Estos otros constituyentes pueden estar presentes en mayor o menor cantidad dependiendo de las condiciones del cultivo usado para cultivar las microalgas, las especies de microalgas, el método de extracción usado para recuperar el aceite de microalgas a partir de la biomasa y otros factores que pueden afectar la composición del aceite de microalgas. Los ejemplos no limitantes de dichos constituyentes incluyen los carotenoides, presentes de 0.1-0.4 microgramos/ml, clorofila presente de 0-0.02 miligramos/kilogramo de aceite, gamma-tocoferol presente de 0.4-0.6 miligramos/100 gramos de aceite, y tocotrienoles totales presentes de 0.2-0.5 miligramos/gramo de aceite.

Los otros constituyentes pueden incluir, sin limitación, fosfolípidos, tocoferoles, tocotrienoles, carotenoides (*por ejemplo*, alfa-caroteno, beta-caroteno, licopeno, etc), xantófilas (*por ejemplo*, luteína, zeaxantina, alfa-criptoxantina y beta-criptoxantina), y varios compuestos orgánicos o inorgánicos.

En algunos casos, el aceite extraído de las especies de *Prototheca* comprende no más de 0.02 mg/Kg de clorofila. En algunos casos, el aceite extraído de las especies de *Prototheca* comprende no más de 0.4 mcg/ml de carotenoides totales. En algunos casos el aceite de *Prototheca* comprende entre 0.40-0.60 miligramos de gamma-tocoferol por cada 100 gramos de aceite. En otros casos, el aceite de *Prototheca* comprende entre 0.2-0.5 miligramos de tocotrienoles totales por gramo de aceite.

III. MÉTODOS Y MATERIALES DE INGENIERÍA GENÉTICA

La presente invención proporciona métodos y materiales para modificar genéticamente células de Prototheca y células huésped recombinantes útiles en los métodos de la presente invención, que incluyen pero sin limitarse a, células huésped de *Prototheca moriformis, Prototheca zopfii, Prototheca krugani,* y *Prototheca stagnora* recombinantes. La descripción de estos métodos y los materiales se divide en subsecciones para la conveniencia del lector. En la subsección 1, se describen los métodos de transformación. En la subsección 2, se describen métodos de ingeniería genética por medio del uso de la recombinación homóloga. En la subsección 3, se describen vectores de expresión y los componentes.

1. Métodos de Ingeniería – Transformación

Las células se pueden transformar por cualquier técnica adecuada que incluyen, por ejemplo, electroporación, biolística, (ver Maruyama y otros (2004), Biotechnology Techniques 8:821-826), transformación por perlas de vidrio y transformación por bigotes de carburo de silicio. Otro método que se puede usar comprende la formación de protoplastos y el uso de $CaCl_2$ y polietilenglicol (PEG) para introducir ADN recombinante en células de microalgas (ver Kim y otros (2002), *Mar. Biotechnol.* 4:63-73, que reporta el uso de este método para la transformación de *Chorella ellipsoidea*). La co-transformación de microalgas se puede usar para introducir dos moléculas diferentes de vectores en una célula simultáneamente (ver, por ejemplo Protist 2004 Dic;155(4):381-93).

Para la transformación de una célula de Prototheca se pueden usar, además, métodos biolísticos (ver, por ejemplo, Sanford, Trends In Biotech. (1988) 6:299 302, patente de Estados Unidos núm. 4,945,050); electroporación (Fromm y otros, Proc. Natl. Acad. Sci. (USA) (1985) 82:5824 5828); el uso de un rayo láser, la microinyección o cualquier otro método capaz de introducir ADN en una microalga.

2. Métodos de Ingeniería - Recombinación Homóloga

La recombinación homóloga es la capacidad de las secuencias de ADN complementarias para alinearse e intercambiar las regiones de homología. El ADN transgénico ("donante") que contiene secuencias homólogas a las secuencias genómicas diana ("patrón") se introduce en el organismo y luego tiene lugar la recombinación dentro del genoma en el sitio de la secuencia genómica homóloga correspondiente. Las etapas mecanísticas de este proceso, en la mayoría de los casos, incluyen: (1) aparear los segmentos de ADN homólogos, (2) la introducción de roturas de doble cadena en la molécula de ADN de los donantes, (3) la invasión de la molécula de DNA patrón por los extremos libres del ADN del donante seguida de la síntesis de ADN; y (4) la resolución de los eventos de reparación de las rupturas de la doble cadena que resultan en los productos finales de la recombinación.

La habilidad para llevar a cabo la recombinación homóloga en un organismo huésped tiene muchas implicaciones prácticas, por lo que se puede llevar a cabo a nivel genético molecular y es útil en la generación de un microbio oleaginoso que puede producir aceites específicos. Por su propia naturaleza la recombinación homóloga es un evento preciso dirigido a los genes, por lo tanto, la mayoría de las líneas transgénicas que se generan con la

misma secuencia de orientación serán esencialmente idénticas en términos de fenotipo, por lo que requieren la evaluación de muchos menos eventos de transformación. La recombinación homóloga también apunta a eventos de inserción de genes en el cromosoma del huésped, lo que resulta en una estabilidad genética excelente, incluso en ausencia de selección genética. Debido a que diferentes locus cromosómicos pudieran impactar en la expresión génica, incluso de promotores heterólogos/UTR, la recombinación homóloga puede ser un método para buscar un locus en un ambiente genómico no familiar y evaluar el impacto de estos ambientes en la expresión génica.

Una aplicación particularmente útil de ingeniería genética que usa la recombinación homóloga es cooptar elementos reguladores específicos del huésped tales como promotores/UTR para dirigir la expresión de genes heterólogos en una forma altamente específica. Por ejemplo, se podría esperar que la ablación precisa de los genes de la estearoil ACP desaturasa endógena con un casete de genes para la FATB (tioesterasa) heteróloga específica para C12:0 y un marcador selectivo adecuado, disminuya drásticamente los niveles endógenos de los ácidos grasos C18:1 concomitante con aumento de los niveles de los ácidos grasos C12:0. El Ejemplo 13 describe la construcción para la recombinación homóloga que es adecuada para la ablación de un gen endógeno de la estearoil ACP desaturasa de *Prototheca moriformis*.

Debido a que la recombinación homóloga es un evento preciso dirigido a los genes, se puede usar para modificar con precisión cualquier nucleótido(s) dentro de un gen o secuencia de interés, siempre y cuando se identifiquen suficientes regiones flancos. Por lo tanto, la recombinación homóloga se puede usar como un medio para modificar secuencias reguladoras que afectan la expresión de genes de ARN y/o proteínas. También se puede usar para modificar las regiones codificantes de proteínas en un esfuerzo por modificar las actividades de enzimas como la especificidad por el sustrato, las afinidades y la Km, y por lo tanto afectar el cambio que se desea en el metabolismo de la célula huésped. La recombinación homóloga proporciona un medio poderoso para manipular el genoma del huésped, lo que da como resultado la mutagénesis dirigida, la conversión de genes, la eliminación de genes, la duplicación de genes, la inversión de genes y el intercambio de elementos de regulación de la expresión génica tales como promotores, potenciadores y 3'UTR.

La recombinación homóloga se puede lograr por medio del uso de construcciones de orientación que contienen pedazos de secuencias endógenas para "orientar" el gen o región de interés en el genoma endógeno de la célula huésped. Tales secuencias de orientación pueden situarse ya sea 5' del gen o región de interés, 3' del gen/región de interés o incluso flanquean el gen/región de interés. Tales construcciones de orientación se pueden transformar dentro de la célula huésped, ya sea como un plásmido de ADN superenrollado con un esqueleto de vector adicional, un producto de PCR sin esqueleto de vector, o como una molécula lineal. En algunos casos, puede ser ventajoso exponer primero las secuencias homólogas dentro del ADN transgénico (ADN donante) con una enzima de restricción. Esta etapa puede incrementar la eficiencia de la recombinación y disminuir la ocurrencia de eventos no deseados. Otros métodos para incrementar la eficiencia de recombinación incluyen el uso de PCR para generar ADN transgénico transformado que contiene extremos lineales homólogos a las secuencias genómicas objetivo.

3. <u>Vectores y componentes de vectores</u>

Los vectores para la transformación de microorganismos, de acuerdo con la presente invención se pueden preparar mediante técnicas familiares conocidas por los expertos en la materia en vista de la descripción de la presente. Un vector contiene, típicamente, uno o más genes en el que cada gen codifica para la expresión de un producto deseado (el producto del gen) y se une operativamente a una o más secuencias control que regulan la expresión de genes o dirigen el producto del gen a una localización en la célula recombinante. Para ayudar al lector, esta subsección se divide en subsecciones. La Subsección A describe las secuencias control que se encuentran, típicamente, en los vectores, así como nuevas secuencias de control que se proporcionan en la presente invención. La Subsección B describe genes que se encuentran, típicamente, en los vectores, así como nuevos métodos de optimización del codón y genes que se preparan por medio del uso de los mismos proporcionados por la invención.

A. Secuencias de control

Las secuencias de control son ácidos nucleicos que regulan la expresión de una secuencia codificante o dirigen un producto del gen a una localización particular dentro o fuera de la célula. Las secuencias de control que regulan la expresión incluyen, por ejemplo, los promotores que regulan la transcripción de una secuencia codificante y terminadores que terminan la transcripción de una secuencia codificante. Otra secuencia de control es una secuencia 3' no traducida que se localiza al final de una secuencia codificante que codifica una señal de poliadenilación. Las secuencias de control que dirigen los productos de genes a una localización particular incluyen las que codifican péptidos señal, los cuales dirigen la proteína a la que están unidos a una localización particular dentro o fuera de la célula.

Por lo tanto, un diseño de vector ejemplar para la expresión de un gen exógeno en una microalga contiene una secuencia que codifica para un producto del gen deseado (por ejemplo, un marcador de selección, una enzima que modifica la vía de los lípidos, o una enzima de utilización de la sacarosa) en una unión operativa con un

promotor activo en microalgas. Alternativamente, si el vector no contiene un promotor en unión operativa con la secuencia codificante de interés, la secuencia codificante se puede transformar dentro de las células de tal manera que se convierte en operativamente unida a un promotor endógeno en el punto de la integración del vector. Se demostró que el método de transformación sin promotor funciona en microalgas (ver por ejemplo Plant Journal 14:4, (1998), pp.441-447).

Muchos promotores están activos en microalgas, incluyendo los promotores que son endógenos a las algas que se transforman, así como promotores que no son endógenos a las algas que se transforman (es decir, promotores de otras algas, promotores de plantas superiores, y los promotores de virus de plantas o virus de algas). Los promotores exógenos o endógenos ilustrativos que son activos en microalgas (así como los genes de resistencia a antibióticos funcionales en microalgas) se describen en la publicación PCT núm. 2008/151149 y las referencias citadas allí).

El promotor que se usa para expresar un gen exógeno puede ser el promotor que se une naturalmente a ese gen o puede ser un gen heterólogo. Algunos promotores están activos en más de una especie de microalgas. Otros promotores son específicos de especie. Los promotores ilustrativos incluyen promotores tales como β-tubulina de *Chlamydomonas reinhardtii*, que se usa en los Ejemplos a continuación y promotores virales, tales como el virus del mosaico de la coliflor (CMV) y virus de la Chlorella, los cuales mostraron ser activos en múltiples especies de microalgas (ver, por ejemplo Plant Cell Rep. 2005 marzo; 23(10-11):727-35; Microbiol J 2005 agosto;43 (4):361-5; Mar Biotechnol (NY) 2002 enero;4(1):63-73). Otro promotor que es adecuado para usar para la expresión de genes exógenos en *Prototheca* es el promotor/5'UTR de la glutamato deshidrogenasa de la *Chlorella sorokiniana* (SEQ ID NO: 69). Opcionalmente, se usan al menos 10, 20, 30, 40, 50, ó 60 nucleótidos o más de estas secuencias que contienen un promotor. Los promotores ilustrativos útiles para la expresión de genes exógenos en *Prototheca* se enumeran en el listado de secuencias de esta solicitud, tal como el promotor del gen HUP1de la *Chlorella* (SEQ ID NO: 1) y el promotor de la nitrato reductasa *Chlorella ellipsoidea* (SEQ ID NO: 2). Los promotores del virus de la Chlorella también se pueden usar para expresar genes en *Prototheca*, tales como SEQ ID NO: 1-7 de la patente de Estados Unidos 6,395,965. Los promotores adicionales activos en *Prototheca* se pueden encontrar, por ejemplo, in Biochem Biophys Res Commun. 1994 octubre 14;204(1):187-94; Plant Mol Biol. 1994 octubre;26(1):85-93; Virology. 2004 agosto 15;326(1):150-9; y Virology. 2004 enero 5;318(1):214-23.

Un promotor se puede caracterizar, generalmente, como constitutivo o inducible. Los promotores constitutivos son generalmente activos o funcionan para guiar la expresión en todo momento (o en ciertos momentos del ciclo de vida de la célula) en el mismo nivel. Los promotores inducibles, por el contrario, se activan (o se vuelven inactivos) o se regulan significativamente ascendente- o descendentemente solamente en respuesta a un estímulo. Ambos tipos de promotores encuentran aplicación en los métodos de la invención. Los promotores inducibles útiles en la invención incluyen aquellos que median la transcripción de un gen unido operativamente en respuesta a un estímulo, tal como una molécula pequeña suministrada exógenamente (por ejemplo, glucosa, como en la SEQ ID NO: 1), temperatura (calor o frío), la falta de nitrógeno en los medios de cultivo, etc. Los promotores adecuados pueden activar la transcripción de un gen esencialmente silente o regular ascendentemente, preferiblemente de forma substancial, la transcripción de un gen unido operativamente que se transcribe a un nivel bajo.

La inclusión de la secuencia de control de la región de terminación es opcional, y si se emplea, entonces la elección será la de conveniencia, ya que la región de terminación es relativamente intercambiable. La región de terminación puede ser nativa a la región de iniciación de la transcripción (el promotor), puede ser nativa de la secuencia de ADN de interés, o se puede obtener de otra fuente. Ver, por ejemplo, Chen y Orozco, Nucleic Acids Res. (1988) 16:8411.

La presente invención también proporciona secuencias de control y genes recombinantes y vectores que las contienen lo que proporciona la expresión compartimentalizada de un gen de interés. Los organelos para la orientación son los cloroplastos, los plastidios, las mitocondrias y el retículo endoplásmico. Además, la presente invención proporciona secuencias de control y los genes recombinantes y vectores que las contienen lo que proporciona la secreción de una proteína fuera de la célula.

Las proteínas que se expresan en el genoma nuclear de *Prototheca* se pueden dirigir a los plastidios mediante señales dirigidas al plastidio. Las secuencias dirigidas a los plastidios endógenas para *Chlorella* se conocen, tales como genes en el genoma nuclear de *Chlorella* que codifican proteínas que se dirigen a los plastidios; ver, por ejemplo los números de acceso AY646197 y AF499684 en GenBank , y en una realización, tales secuencias de control se usan en los vectores de la presente invención para dirigir la expresión de una proteína a un plastidio de *Prototheca*.

Los ejemplos a continuación describen el uso de secuencias dirigidas a los plastidios de algas para dirigir las proteínas heterólogas al compartimiento correcto en la célula huésped. Las bibliotecas de ADNc se realizaron por medio del uso de células de *Prototheca moriformis* y *Chlorella protothecodies* y se describen en los Ejemplos 12 y Ejemplo 11 debajo. Las secuencias se evaluaron por BLAST y se analizaron por homología con proteínas conocidas que transitan a plastidio/cloroplasto. Los ADNc que codifican estas proteínas se clonaron y las secuencias dirigidas

a los plastidios se aislaron de estos ADNc. Las secuencias de aminoácidos de las secuencias dirigidas a los plastidios de algas se identificaron a partir de las bibliotecas de ADNc y las secuencias de aminoácidos de la acil graso ACP tiosterasa de planta que se usan en los ejemplos de expresión heteróloga debajo se enumeran en la SEQ ID NO: 127-133.

En otra realización de la presente invención, la expresión de un polipéptido en *Prototheca* se dirige al retículo endoplásmico. La inclusión de una señal de retención o clasificación apropiada en un vector de expresión garantiza que las proteínas se retengan en el retículo endoplasmático (RE) y no vayan hacia abajo adentro de Golgi. Por ejemplo, el vector IMPACTVECTOR1.3, de Wageningen UR-Plant Research International, incluye la señale KDEL de retención o clasificación bien conocida. Con este vector, la retención en el RE tiene una ventaja práctica ya que se reportó que mejora los niveles de expresión 5 veces o más. La razón principal de esto parece ser que el RE contiene concentraciones más bajas y/o proteasas diferentes responsables de la degradación post-traduccional de proteínas que se expresan que las presentes en el citoplasma. Se conocen las señales de retención del RE funcionales en microalgas verdes. Por ejemplo, ver Proc Natl Acad Sci USA. 26 de abril de 2005, 102(17):6225-30.

En otra realización de la presente invención, un polipéptido se dirige para su secreción fuera de la célula en los medios de cultivo. Ver Hawkins y otros, Current Microbiology Vol. 38 (1999), pp 335-341 para los ejemplos de señales de secreción activa en *Chlorella* que se pueden usar, de acuerdo con los métodos de la invención, en *Prototheca*.

B. Optimización de genes y codón

Por lo general, un gen incluye un promotor, la secuencia codificante, y las secuencias de control de la terminación. Cuando se ensambla por tecnología de ADN recombinante, un gen puede denominarse un casete de expresión y se puede flanquear por sitios de restricción para la inserción adecuada dentro de un vector que se usa para introducir el gen recombinante dentro de una célula huésped. El casete de expresión se puede flanquear por secuencias de ADN del genoma u otros ácidos nucleicos diana para facilitar la integración estable del casete de expresión en el genoma por recombinación homóloga. Alternativamente, el vector y su casete de expresión pueden permanecer sin integrarse, en cuyo caso, el vector incluye, típicamente, un origen de replicación que es capaz de proporcionar la replicación del ADN del vector heterólogo.

Un gen común presente en un vector es un gen que codifica para una proteína, cuya expresión permite a la célula recombinante que contiene la proteína diferenciarse de las células que no expresan la proteína. Tal gen y su producto génico correspondiente, se llama un marcador de selección. Cualquiera de una gran variedad de marcadores de selección se puede emplear en una construcción transgénica útil para transformar *Prototheca*. Los ejemplos de marcadores de selección adecuados incluyen el gen de resistencia G418, el gen de la nitrato reductasa (ver Dawson y otros (1997), Current Microbiology 35:356-362.), el gen de la higromicina fosfotransferasa (HPT; ver Kim y otros (2002), Mar. Biotechnol. 4:63-73), el gen de la neomicina fosfotransferasa y el gen ble, que confiere resistencia a la fleomicina (Huang y otros. (2007), Appl. Microbiol. Biotechnol. 72:197-205). Los métodos para determinar la sensibilidad de las microalgas a los antibióticos son bien conocidos. Por ejemplo, Mol Gen Genet. 16 de octubre de 1996; 252(5):572-9.

Para los propósitos de la presente invención, el vector de expresión que se usa para preparar una célula huésped recombinante de la invención incluirá al menos dos, y a menudo tres, genes, si uno de los genes es un marcador de selección. Por ejemplo, una *Prototheca* genéticamente modificada de la invención se puede hacer por la transformación con vectores de la invención que comprenden, además de un marcador de selección, uno o más genes exógenos, tales como, por ejemplo, el gen de la sacarosa invertasa o el gen de la acil ACP tioesterasa. Uno o ambos genes se pueden expresar por medio del uso de un promotor inducible, el cual permite controlar el tiempo de expresión relativo de estos genes para mejorar el rendimiento de los lípidos y la conversión a ésteres de ácidos grasos. La expresión de los dos o más genes exógenos puede estar bajo el control del mismo promotor inducible o bajo el control de diferentes promotores inducibles (o constitutivos). En la última situación, la expresión de un primer gen exógeno se puede inducir por un primer período de tiempo (durante el cual la expresión de un segundo gen exógeno puede o no inducirse) y la expresión de un segundo gen exógeno se puede inducir por un segundo período de tiempo (durante el cual la expresión del primer gen exógeno puede o no inducirse).

En otras realizaciones, los dos o más genes exógenos (además de cualquier marcador de selección) son: una acil graso ACP tiosterasa y una acil graso-CoA/aldehído reductasa, cuyas acciones combinadas rinden un producto del alcohol. Además se proporcionan otras combinaciones de genes exógenos, que incluyen, sin limitación, una acil graso acp tiosterasa y una acil graso-CoA reductasa para generar aldehídos. En una realización, el vector proporciona la combinación de un una acil graso acp tiosterasa, una acil graso-CoA reductasa y una aldehído graso decarbonilasa para generar alcanos. En cada una de estas realizaciones, uno o más de los genes exógenos se pueden expresar por medio del uso de un promotor inducible.

Otros vectores ilustrativos de la invención que expresan dos o más genes exógenos incluyen aquellos que codifican tanto un transportador de sacarosa y una enzima sacarosa invertasa y aquellos que codifican un marcador

de selección y una sacarosa invertasa secretada. La *Prototheca* recombinante transformada con cualquier tipo de vector produce lípidos a un costo de fabricación más bajo debido a la capacidad generada por ingeniería de usar la caña de azúcar (y azúcares derivados del azúcar de caña) como fuente de carbono. La inserción de los dos genes exógenos que se describieron anteriormente se puede combinar con la interrupción de la biosíntesis de polisacáridos a través de mutagénesis dirigida y/o al azar, la cual que dirige el flujo cada vez mayor de carbono a la producción de lípidos. Individualmente y en combinación, la conversión trófica, las modificaciones para alterar la producción de lípidos y el tratamiento con enzimas exógenos alteran la composición lipídica producida por un microorganismo. La alteración puede ser un cambio en la cantidad de lípidos producidos, la cantidad de una o más especies de hidrocarburos producidos en relación con otros lípidos, y/o los tipos de especies de lípidos que se producen en el microorganismo. Por ejemplo, las microalgas se pueden modificar para producir una mayor cantidad y/o porcentaje de TAG.

Para la expresión óptima de una proteína recombinante, es beneficioso emplear secuencias de codificación que producen ARNm con codones que se usan preferentemente por la célula huésped que se va a transformar. Así, la expresión adecuada de los transgenes puede requerir que el uso del codón del transgén coincida con el codón específico con sesgo en el organismo en el que se expresa el transgén. Los mecanismos exactos subyacentes a este efecto son muchos, pero incluyen el equilibrio adecuado de los almacenes de ARNt aminoacilados disponibles con las proteínas que se sintetizan en la célula, acoplado a una traducción más eficiente de los transgénicos de ARN mensajero (ARNm) cuando se satisface esta necesidad. Cuando el uso de codones en el transgén no es optimizado, los almacenes de ARNt disponibles no son suficientes para permitir la traducción eficiente del ARNm heterólogo lo que resulta en el estancamiento y la terminación ribosomal y la posible inestabilidad del ARNm transgénico.
La presente invención proporciona ácidos nucleicos con codones optimizados útiles para la expresión exitosa de proteínas recombinantes en *Prototheca*. El uso de codones en la especie *Prototheca* se analizó mediante el estudio de secuencias de ADN que se aislaron de *Prototheca moriformis*. Este análisis representa el interrogatorio de más de 24, 000 codones y dio como resultado la Tabla 1 debajo.

Tabla 1. Uso preferido del codón en cepas de *Prototheca*.

Ala	GCG	345 (0.36)	Asn	AAT	8 (0.04)
	GCA	66 (0.07)		AAC	201 (0.96)
	GCT	101 (0.11)			
	GCC	442 (0.46)	Pro	CCG	161 (0.29)
				CCA	49 (0.09)
Cys	TGT	12 (0.10)		CCT	71 (0.13)
	TGC	105 (0.90)		CCC	267 (0.49)
Asp	GAT	43 (0.12)	Gln	CAG	226 (0.82)
	GAC	316 (0.88)		CAA	48 (0.18)
Glu	GAG	377 (0.96)	Arg	AGG	33 (0.06)
	GAA	14 (0.04)		AGA	14 (0.02)
				CGG	102 (0.18)
Phe	TTT	89 (0.29)		CGA	49 (0.08)
	TTC	216 (0.71)		CGT	51 (0.09)
				CGC	331 (0.57)
Gly	GGG	92 (0.12)			
	GGA	56 (0.07)	Ser	AGT	16 (0.03)
	GGT	76 (0.10)		AGC	123 (0.22)
	GGC	559 (0.71)		TCG	152 (0.28)
				TCA	31 (0.06)
His	CAT	42 (0.21)		TCT	55 (0.10)
	CAC	154 (0.79)		TCC	173 (0.31)
Ile	ATA	4 (0.01)	Thr	ACG	184 (0.38)
	ATT	30 (0.08)		ACA	24 (0.05)
	ATC	338 (0.91)		ACT	21 (0.05)
				ACC	249 (0.52)
Lys	AAG	284 (0.98)			
	AAA	7 (0.02)	Val	GTG	308 (0.50)
				GTA	9 (0.01)
Leu	TTG	26 (0.04)		GTT	35 (0.06)
	TTA	3 (0.00)		GTC	262 (0.43)
	CTG	447 (0.61)			
	CTA	20 (0.03)	Trp	TGG	107 (1.00)

(continuación)

CTT		45 (0.06)			
	CTC	190 (0.26)	Tyr	TAT	10 (0.05)
				TAC	180 (0.95)
Met	ATG	191 (1.00)			
			Terminación		TGA/TAG/TAA

En otras realizaciones, se optimizaron los codones en el gen en el vector recombinante con referencia a una cepa de microalgas que no sea una cepa *Prototheca*. Por ejemplo, los métodos de registro para la expresión de genes en las microalgas se describen en la patente de Estados Unidos 7,135,290. Información adicional para la optimización del codón está disponible, por ejemplo, en la base de datos de uso de codones de GenBank.

Mientras los métodos y materiales de la invención permiten la introducción de cualquier gen exógeno en *Prototheca*, los genes relacionados con el uso de la sacarosa y la modificación de la vía de lípidos son de particular interés, como se discute en las secciones siguientes.

IV. UTILIZACIÓN DE LA SACAROSA

En la realización, la célula de *Prototheca* recombinante de la invención contiene, además, uno o varios genes exógenos para la utilización de la sacarosa. En varias realizaciones, el o más genes codifican una o más proteínas seleccionadas del grupo que consiste en una fructocinasa, una glucocinasa, una hexocinasa, una sacarosa invertasa, un transportador de sacarosa. Por ejemplo, la expresión de un transportador de sacarosa y una sacarosa invertasa permite a la *Prototheca* transportar la sacarosa al interior de la célula desde el medio de cultivo e hidroliza la sacarosa para producir glucosa y fructosa. Opcionalmente, una fructocinasa se puede expresar también en los casos donde la actividad de la hexocinasa endógena es insuficiente para la fosforilación máxima de fructosa. Los ejemplos de los transportadores de sacarosa adecuados son los números de acceso en Genbank CAD91334, CAB92307 y CAA53390. Los ejemplos de fructocinasas adecuadas son los números de acceso en Genbank P26984, P26420 y CAA43322.

En una realización, la presente invención proporciona una célula huésped de *Prototheca* que secreta una sacarosa invertasa. La secreción de una sacarosa invertasa evita la necesidad de la expresión de un transportador que pueda transportar la sacarosa hacia el interior de la célula. Esto se debe a que la invertasa secretada cataliza la conversión de una molécula de sacarosa en una molécula de glucosa y una molécula de fructosa, las cuales se pueden transportar y usar por los microbios proporcionados por la invención. Por ejemplo, la expresión de una sacarosa invertasa (tales como SEQ ID NO: 3) con una señal de secreción (tal como SEQ ID NO: 4 (de levadura), SEQ ID NO: 5 (de plantas superiores), SEQ ID NO: 6 (señal consenso de secreción de eucariotas), y SEQ ID NO: 7 (combinación de secuencia señales de plantas superiores y el consenso de eucariota) genera la actividad invertasa fuera de la célula. La expresión de esa proteína, como la que permite la metodología de la ingeniería genética descrita en la presente, permite a las células que ya son capaces de usar la glucosa extracelular como fuente de energía, usar la sacarosa como fuente de energía extracelular.

Las especies de Prototheca que expresan una invertasa en un medio que contiene sacarosa son una especie preferida de microalgas para la producción de aceite. El Ejemplo 3 ilustra cómo los métodos y reactivos de la invención se pueden usar para expresar una invertasa recombinante de levadura y secretarla desde una célula *Prototheca* recombinante. La expresión y direccionalidad extracelular de esta proteína completamente activa permite a las células huésped resultantes crecer en sacarosa, mientras que sus contrapartes no transformadas no pueden. Por lo tanto, la presente invención proporciona células *Prototheca* recombinantes con codón optimizado para el gen de la invertasa, incluye pero no se limita al gen de la invertasa de levadura, integrado en su genoma de forma tal que el gen de la invertasa se expresa según se evaluó por la actividad de la invertasa y la hidrólisis de la sacarosa. La presente invención también proporciona genes de la invertasa útiles como marcadores de selección en las células *Prototheca* recombinantes, ya que estas células son capaces de crecer en sacarosa, mientras que sus contrapartes no transformadas no pueden, y métodos para la selección de las células huésped recombinantes por medio del uso de una invertasa como un marcador de selección poderoso para la genética molecular de algas.

La expresión exitosa de una sacarosa invertasa en *Prototheca* también ilustra otro aspecto de la presente invención, ya que demuestra que las proteínas heterólogas (recombinantes) se pueden expresar en la célula de alga y transitar con éxito fuera de la célula y en el medio de cultivo en una forma completamente activa y funcional. Por lo tanto, la presente invención proporciona métodos y reactivos para expresar una amplia y diversa gama de proteínas heterólogas en microalgas y secretarlas fuera de la célula huésped. Tales proteínas incluyen, por ejemplo, enzimas industriales, tales como, por ejemplo, lipasas, proteasas, celulasas, pectinasas, amilasas, esterasas, oxidorreductasas, transferasas, lactasas, isomerasas e invertasas, así como proteínas terapéuticas tales como, por ejemplo, factores de crecimiento, citocinas, anticuerpos de talla completa que comprenden dos cadenas ligeras y dos pesadas, Fab, scFv (fragmento variable de simple cadena), anticuerpos tipo-camello, fragmentos de anticuerpos, fusiones de fragmentos de anticuerpos, fusiones de anticuerpo-receptor, insulina, interferones, y los

factores de crecimiento tipo insulina.

La expresión exitosa de una sacarosa invertasa en *Prototheca* también ilustra otro aspecto de la presente invención, ya que proporciona métodos y reactivos para el uso de péptidos de tránsito de hongos en las algas para dirigir la secreción de proteínas en *Prototheca*; y métodos y reactivos para determinar si un péptido puede funcionar, y su capacidad de funcionar, como un péptido de tránsito en células de *Prototheca*. Los métodos y reactivos de la invención se pueden usar como una herramienta y plataforma para identificar otros péptidos para transportar con éxito proteínas al exterior de una célula, y que la invertasa de levadura tiene gran utilidad en estos métodos. Como se demuestra en este ejemplo, la eliminación del péptido de tránsito de invertasa endógeno de levadura y su sustitución por otros péptidos de tránsito, ya sean endógenos a las algas huésped o de otras fuentes (eucariotas, procariotas y virus), puede identificar si cualquier péptido de interés puede funcionar como un péptido de tránsito para guiar a las proteínas en su salida de la célula.

Los ejemplos de sacarosa invertasas adecuadas incluyen los identificados por los números de acceso en Genbank CAB95010, NP_012104 y CAA06839. Los ejemplos no limitantes de invertasas adecuadas se enumeran a continuación en la Tabla 2. Las secuencias de aminoácidos para cada invertasa de la lista se incluyen en el listado de secuencias a continuación. En algunos casos, el gen de utilización de la sacarosa exógena adecuado para su uso en los métodos y los vectores de la invención codifica una sacarosa invertasa que tiene al menos 40, 50, 60, 75, o 90% o más identidad de aminoácidos con una sacarosa invertasa que se selecciona de la Tabla 2.

Tabla 2. Sacarosa invertasa.

Descripción	**Organismo**	**Núm. acceso**	**SEQ ID NO:**
Invertasa	*Chicorium intybus*	Y11124	SEQ ID NO:20
Invertasa	*Schizosaccharomyces pombe*	AB011433	SEQ ID NO:21
beta-fructofuranosidasa (invertasa)	*Pichia anomala*	X80640	SEQ ID NO:22
Invertasa	*Debaryomyces occidentalis*	X17604	SEQ ID NO:23
Invertasa	*Oryza sativa*	AF019113	SEQ ID NO:24
Invertasa	*cepa Allium*	AJ006067	SEQ ID NO:25
Invertasa	*Beta vulgaris subsp. Vulgaris*	AJ278531	SEQ ID NO:26
beta-fructofuranosidasa (invertasa)	*Bifidobacterium breve UCC2003*	AAT28190	SEQ ID NO:27
Invertasa	*Saccharomyces cerevisiae*	NP_012104	SEQ ID NO:8 (nucleótido) SEQ ID NO:28 (aminoácido)
Invertasa A	*Zymomonas mobilis*	AAO38865	SEQ ID NO:29

La secreción de una invertasa al medio de cultivo por *Prototheca* permite a las células crecer también en melazas residuales del procesamiento de la caña de azúcar como lo hacen en glucosa pura de grado-reactivo; el uso de este producto de desecho de poco valor del procesamiento de la caña de azúcar puede proporcionar importantes ahorros en los costos en la producción de lípidos y otros aceites. Por lo tanto, la presente invención

proporciona un cultivo de microorganismos que contiene una población de microorganismos *Prototheca*, y un medio de cultivo que comprende (i) sacarosa y (ii) una enzima sacarosa invertasa. En varias realizaciones la sacarosa en el cultivo viene de sorgo, remolacha azucarera, caña de azúcar, melaza, o material celulósico despolimerizado (que opcionalmente puede contener lignina). En otro aspecto, los métodos y reactivos de la invención aumentan significativamente el número y tipos de materias primas que puede usar la *Prototheca* recombinante. Mientras que los microbios que se ejemplifican aquí se alteran de tal forma que pueden usar la sacarosa, los métodos y reactivos de la invención se pueden aplicar de manera que las materias primas, tales como las celulósicas son utilizables por un microbio huésped modificado de la invención con capacidad para secretar celulasas, pectinasas, isomerasas , o similares, de tal manera que los productos de degradación de las reacciones enzimáticas ya no son simplemente tolerados, sino más bien se usan como fuente de carbono por el huésped.

V. INGENIERÍA DE LA VÍA DE LOS LÍPIDOS

Además de alterar la capacidad de *Prototheca* de usar materias primas como las materias primas que contienen sacarosa, la presente invención también proporciona *Prototheca* recombinante que se modificó para alterar las propiedades y/o proporciones de los lípidos producidos. La vía puede además, o alternativamente, modificarse para alterar las propiedades y/o proporciones de varias moléculas de lípidos que se producen a través del procesamiento enzimático de los lípidos y los intermedios en las vías del ácido graso. En varias realizaciones, las células de *Prototheca* recombinante de la invención tienen, en comparación con sus contrapartes no transformadas, un rendimiento de lípidos optimizado por unidad de volumen y/o por unidad de tiempo, la longitud de la cadena de carbono (por ejemplo, para la producción de diesel renovable o para aplicaciones de la industria química que requieren materias primas lipídicas), número reducido de dobles o triples enlaces, opcionalmente a cero, y un aumento de la relación hidrógeno:carbono de una especie particular de lípidos o de una población de lípidos distintos.

En realizaciones particulares, una o más enzimas claves que controlan puntos de ramificación en el metabolismo de la síntesis de ácidos grasos se regularon ascendente o descendentemente para mejorar la producción de lípidos. La regulación ascendente se puede lograr, por ejemplo, mediante la transformación de células con construcciones de expresión en las que un gen que codifica la enzima de interés se expresa, por ejemplo, por medio del uso de un promotor fuerte y/o elementos potenciadores que aumentan la transcripción. Tales construcciones pueden incluir un marcador de selección de manera que los transformantes pueden ser objeto de selección, la cual puede resultar en una amplificación de la construcción y un aumento en el nivel de expresión de la enzima codificada. Los ejemplos de enzimas adecuadas para la regulación ascendente de acuerdo con los métodos de la invención incluyen la piruvato deshidrogenasa, que desempeña un papel en la conversión del piruvato a acetil-CoA (ejemplos, algunas de microalgas, incluyen números de acceso en Genbank NP_415392; AAA53047; Q1XDM1; y CAF05587). La regulación ascendente de la piruvato deshidrogenasa puede aumentar la producción de acetil-CoA, y de ese modo aumentar la síntesis de ácidos grasos. La acetil-CoA carboxilasa cataliza la etapa inicial en la síntesis de ácidos grasos. En consecuencia, esta enzima se puede regular ascendentemente para aumentar la producción de ácidos grasos (ejemplos, algunas de microalgas, incluyen números de acceso de Genbank BAA94752; AAA75528; AAA81471; YP_537052; YP_536879; NP_045833 y BAA57908). La producción de ácidos grasos también se puede incrementar por regulación ascendente de la proteína portadora de acilo (ACP), la cual transporta las cadenas de acilo en crecimiento durante la síntesis de ácidos grasos (ejemplo, algunas de microalgas, incluyen números de acceso en Genbank A0T0F8; P51280; NP_849041; YP_874433). La glicerol-3-fosfato acil transferasa cataliza el paso limitante de la velocidad de la síntesis de ácidos grasos. La regulación ascendente de esta enzima puede aumentar la producción de ácidos grasos (por ejemplo, algunas de microalgas, incluyen números de acceso en Genbank AAA74319; AAA33122; AAA37647; P44857 y ABO94442).

La regulación ascendente o descendente de los genes se puede aplicar a los reguladores globales que controlan la expresión de los genes de las vías biosintéticas de los ácidos grasos. En consecuencia, uno o más reguladores globales de la síntesis de ácidos grasos se regula positiva o negativamente, según proceda, para inhibir o aumentar, respectivamente, la expresión de una pluralidad de genes sintéticos de ácidos grasos y, en última instancia, aumentar la producción de lípidos. Los ejemplos incluyen las proteínas de unión a elementos reguladores del esterol, (SREBP), tales como SREBP-1a y SREBP-1c (para ejemplos ver los números de acceso en Genbank NP_035610 y Q9WTN3).

La presente invención también proporciona células de *Prototheca* recombinantes que se modificaron para contener uno o más genes exógenos que codifican enzimas para la modificación de lípidos tales como, por ejemplo, acil graso-ACP tioesterasas (ver la Tabla 3), acil graso-CoA/aldehído reductasas (ver la Tabla 4), acil graso-CoA reductasas (ver la Tabla 5), aldehídos grasos decarboxilasa (ver la Tabla 6), aldehídos grasos reductasas y escualeno sintasas (ver el número de acceso en Genbank AF205791). En algunas realizaciones, los genes que codifican la acil graso ACP tioesterasa y una proteína portadora de acilo que se co-expresa naturalmente se transforman dentro de una célula *Prototheca*, opcionalmente con uno o más genes que codifican otras enzimas para la modificación de lípidos. En otras realizaciones, la ACP y la acil graso ACP tioesterasa pueden tener una afinidad una por la otra que da una ventaja cuando las dos se usan juntas en los microbios y los métodos de la presente invención, sin tomar en cuenta si ellas se co-expresan naturalmente o no en un tejido u organismo particular. Por lo

tanto, la presente invención contempla tanto pares de estas enzimas que se co-expresan naturalmente como aquellas que comparten una afinidad para interactuar una con la otra para facilitar la escisión de una cadena de carbono de longitud específica a partir de ACP.

En otras realizaciones, un gen exógeno que codifica una desaturasa se transforma dentro de una célula de *Prototheca* en conjunto con uno o más genes que codifican otras enzimas de modificación de lípidos para proporcionar modificaciones con respecto a la saturación de lípidos. La estearoil ACP desaturasa (ver, por ejemplo, números de acceso al GenBank AAF15308; ABM45911; y AAY86086), por ejemplo, cataliza la conversión de estearoil ACP a oleoil ACP. La regulación ascendente de este gen puede aumentar la proporción de ácidos grasos monoinsaturados producida por una célula, mientras que la regulación descendente puede reducir la proporción de monoinsaturados. Del mismo modo, la expresión de una o más desaturasas de glicerolípidos se puede controlar para alterar la relación de ácidos grasos insaturados a saturados tales como ácido graso ω-6 desaturasa, ácidos grasos ω-3 desaturasa, o ω-6-oleato desaturasa. En algunas realizaciones, la desaturasa se puede seleccionar con referencia a una longitud de cadena de carbón deseada, de manera que la desaturasa es capaz de hacer modificaciones en ubicaciones específicas dentro de un sustrato de longitud de carbonos especificado, o sustratos que tienen una longitud de carbonos dentro de un rango especificado.

Así, en realizaciones particulares, los microbios de la presente invención son manipulados genéticamente para expresar uno o más genes exógenos seleccionados de una acil-ACP tioesterasa, una acil-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, una aldehído graso decarbonilasa, o una proteína portadora de acilo naturalmente expresada. Los métodos de expresión adecuados se describen anteriormente en relación con la expresión de un gen de la lipasa, incluyendo, entre otros métodos, la expresión inducible y la expresión compartimentada. Una acil graso ACP tioesterasa escinde un ácido graso de una proteína portadora de acilo (ACP) durante la síntesis de lípidos. Mediante procesamiento enzimático adicional, el ácido graso escindido se combina con una coenzima para producir una molécula de acil-CoA. Este acil-CoA es el sustrato para la actividad enzimática de la acil graso-CoA reductasa para rendir un aldehído, así como para una acil graso-CoA/aldehído reductasa para producir un alcohol. El aldehído que se produce por la acción de la acil graso-CoA reductasa identificada arriba es el sustrato para la actividad enzimática adicional ya sea por una aldehído graso reductasa para producir un alcohol, o una aldehído graso decarbonilasa para producir un alcano o alqueno.

En algunas realizaciones, los ácidos grasos, glicerolípidos, o alcoholes primarios correspondientes, aldehídos, alcanos o alquenos, que se generan por los métodos que se describen en la presente, contienen 8, 10, 12 o 14 átomos de carbono. Los ácidos grasos preferidos para la producción de diesel, biodiesel, diesel renovable, o combustible de reactor, o los correspondientes alcoholes primarios, aldehídos, alcanos y alquenos, para aplicaciones industriales contienen de 8 a 14 átomos de carbono. En ciertas realizaciones, los ácidos grasos anteriores, así como las otras moléculas de hidrocarburo correspondientes, están saturados (sin ningún doble o triple enlace carbono-carbono), mono insaturados (un solo doble enlace); poli insaturado (dos o más dobles enlaces); es lineal (no cíclico) o ramificado. Para la producción de combustible, se prefiere una mayor saturación.

Las enzimas que se describen directamente arriba tienen una especificidad preferencial para la hidrólisis de un sustrato que contiene un número específico de átomos de carbono. Por ejemplo, una acil graso-ACP tioesterasa puede tener preferencia por la escisión de un ácido graso que tiene 12 átomos de carbono de acp. En algunas realizaciones, la ACP y la tioesterasa de longitud específica pueden tener una afinidad una por la otra que las hace especialmente útiles en una combinación (por ejemplo, los genes exógenos de la ACPy la tioesterasa se pueden co-expresar naturalmente en un determinado tejido u organismo del cual se derivan). Por lo tanto, en varias realizaciones, las células de *Prototheca* recombinantes de la invención pueden contener un gen exógeno que codifica una proteína con especificidad para catalizar una actividad enzimática (por ejemplo, la escisión de un ácido graso de una ACP, la reducción de un acil-CoA a un aldehído o un alcohol, o la conversión de un aldehído a un alcano) con respecto al número de átomos de carbono contenidos en el sustrato. La especificidad enzimática puede, en varias realizaciones, ser por un substrato que tiene de 8 a 34 átomos de carbono, preferentemente de 8 a 18 átomos de carbono, y con mayor preferencia de 8 a 14 átomos de carbono. Una especificidad preferida es por un substrato que tiene menos, es decir, 12, mejor que más, es decir, 18 átomos de carbono.

En ejemplos no limitantes sino ilustrativos, la presente invención proporciona vectores y células Prototheca huésped que expresan una tioesterasa exógena y en consecuencia producen lípidos enriquecidos, en relación con el perfil lipídico de las células *Prototheca* no transformadas, en la longitud de la cadena para la cual es específica la tioesterasa. Las tioesterasas ilustradas son (i) *Cinnamomum camphorum* FatB1 (núm. de acceso al GenBank Q39473, la secuencia de aminoácidos está en la SEQ ID NO: 59, la secuencia de aminoácidos sin secuencia dirigida a plastidios (STP) está en la SEQ ID NO: 139, y la secuencia de ADNc con el codón optimizado que se basa en la Tabla 1 está en la SEQ ID NO: 60), la cual tiene una preferencia por el sustrato acil graso ACP, con una longitud de cadena de carbono de 14, (ii) *Cuphea hookeriana* FatB2 (núm. de acceso al GenBank AAC49269, la secuencia de aminoácidos está en SEQ ID NO: 61, la secuencia de aminoácidos sin STP está en la SEQ ID NO: 138, y la secuencia de ADNc con el codón optimizado que basa en la Tabla1 se encuentra en la SEQ ID NO: 62), que tiene una preferencia por un sustrato acil graso ACP con longitud de la cadena de carbono de 8-10, y (iii) *Umbellularia* Fat B1 (núm. de acceso al GenBank Q41635, la secuencia de aminoácidos se incluye en la SEQ ID NO: 63, la

secuencia de aminoácidos sin STP está en la SEQ ID NO: 139, y la secuencia de ADNc con el codón optimizado que basa en la Tabla1 se encuentra en la SEQ ID NO: 64), que tiene una preferencia por un sustrato acil graso ACP con longitud de la cadena de carbono de 12.

Otras acil graso-ACP tioesterasas adecuadas para usar con los microbios y métodos de la invención incluyen, sin limitarse a, las enumeradas en la Tabla 3.

Tabla 3. Acil graso-ACP tioesterasas y números de acceso al GenBank.

Umbellularia californica acil graso-ACP tioesterasa (GenBank #AAC49001)
Cinnamomum camphora acil graso-ACP tioesterasa (GenBank #Q39473)
Umbellularia californica acil graso-ACP tioesterasa (GenBank #Q41635)
Myristica fragrans acil graso-ACP tioesterasa (GenBank #AAB71729)
Myristica fragrans acil graso-ACP tioesterasa (GenBank #AAB71730)
Elaeis guineensis acil graso-ACP tioesterasa (GenBank #ABD83939)
Elaeis guineensis acil graso-ACP tioesterasa (GenBank #AAD42220)
Populus tomentosa acil graso-ACP tioesterasa (GenBank #ABC47311)
Arabidopsis thaliana acil graso-ACP tioesterasa (GenBank #NP_172327)
Arabidopsis thaliana acil graso-ACP tioesterasa (GenBank #CAA85387)
Arabidopsis thaliana acil graso-ACP tioesterasa (GenBank #CAA85388)
Gossypium hirsutum acil graso-ACP tioesterasa (GenBank #Q9SQI3)
Cuphea lanceolata acil graso-ACP tioesterasa (GenBank #CAA54060)
Cuphea hookeriana acil graso-ACP tioesterasa (GenBank #AAC72882)
Cuphea calophylla subsp. mesostemon acil graso-ACP tioesterasa (GenBank #ABB71581)
Cuphea lanceolata acil graso-ACP tioesterasa (GenBank #CAC19933)
Elaeis guineensis acil graso-ACP tioesterasa (GenBank #AAL15645)
Cuphea hookeriana acil graso-ACP tioesterasa (GenBank #Q39513)
Gossypium hirsutum acil graso-ACP tioesterasa (GenBank #AAD01982)
Vitis vinifera acil graso-ACP tioesterasa (GenBank #CAN81819)
Garcinia mangostana acil graso-ACP tioesterasa (GenBank #AAB51525)
Brassica juncea acil graso-ACP tioesterasa (GenBank #ABI18986)
Madhuca longifolia acil graso-ACP tioesterasa (GenBank #AAX51637)
Brassica napus acil graso-ACP tioesterasa (GenBank #ABH11710)
Oryza sativa (cultivares grupo índica) acil graso-ACP tioesterasa (GenBank #EAY86877)
Oryza sativa (cultivares grupo japónica) acil graso-ACP tioesterasa (GenBank #NP_001068400)
Oryza sativa (cultivares grupo índica) acil graso-ACP tioesterasa (GenBank #EAY99617)
Cuphea hookeriana acil graso-ACP tioesterasa (GenBank #AAC49269)
Ulmus Americana acil graso-ACP tioesterasa (GenBank #AAB71731)
Cuphea lanceolata acil graso-ACP tioesterasa (GenBank #CAB60830)
Cuphea palustris acil graso-ACP tioesterasa (GenBank #AAC49180)
Iris germanica acil graso-ACP tioesterasa (GenBank #AAG43858)
Cuphea palustris acil graso-ACP tioesterasa (GenBank #AAC49179)
Myristica fragrans acil graso-ACP tioesterasa (GenBank# AAB71729)
Cuphea hookeriana acil graso-ACP tioesterasa (GenBank #U39834)
Umbelluaria californica acil graso-ACP tioesterasa (GenBank # M94159)
Cinnamomum camphora acil graso-ACP tioesterasa (GenBank #U31813)

Los ejemplos debajo describen el éxito de la orientación y la expresión heteróloga de la acil graso ACP de *Cuphea hookeriana, Umbellularia californica, Cinnamomun camphora* en especies de *Prototheca*. Adicionalmente, las alteraciones en el perfil de ácidos grasos se confirmaron en la expresión de las células huésped de estas acil graso-ACP tioesterasas heterólogas. Estos resultados fueron inesperados dada la falta de identidad de secuencia entre las tioesterasas de algas y de plantas superiores en general, y entre las acil graso-ACP tioesterasas de *Prototheca moriformis* y las acil graso-ACP tioesterasas heterólogas enumeradas arriba. Se aislaron y secuenciaron dos acil graso-ACP tioesterasas de *Prototheca moriformis*. Las secuencias de los dos ADNc mostraron un alto grado de identidad entre ellas, y difieren sólo en 12 posiciones a nivel de nucleótidos y cinco posiciones en el nivel de aminoácidos, cuatro de ellos en el péptido de tránsito a plastidios. Un análisis más detallado de la secuencia genómica de *Prototheca moriformis* confirmó que estos dos ADNc se codificaron de hecho en cóntigos separados, y aunque muy homólogos, se codificaron por dos genes distintos. La secuencia de ADNc y amino ácidos de las dos acil graso-ACP tioesterasas de *Prototheca moriformis P.moriformis* acil graso ACP tioesterasa-1 y *P.moriformis* acil graso ACP tioesterasa -2, se muestran como SEQ ID NO: 134-137.

Cuando las secuencias de aminoácidos de estos dos ADNc se analizaron por BLAST contra la base de datos NCBI, las dos secuencias más homóloga fueron las de acil graso-ACP tioesterasas de *Chlamydomonas reinhardtii* y *Arabidopsis thaliana*. Sorprendentemente, el nivel de identidad de aminoácidos entre las acil graso-ACP tioesterasas de *Prototheca moriformis* y las tioesterasas de plantas superiores fue bastante bajo, con sólo 49 y 37%

de identidad. Además, también hay una sutil diferencia en las secuencias que rodean la porción amino terminal de la tríada catalítica ($NXHX_{36}C$) entre estas acil graso-ACP tioesterasas. Treinta y nueve de cuarenta acil graso-ACP tioesterasas de plantas superiores que se estudiaron mostraron la secuencia LDMNQH que rodea los residuos N y H en el extremo amino terminal de la tríada, mientras que todas las secuencias de algas que se identificaron tenían la secuencia MDMNGH. Dada la baja identidad de secuencia de aminoácidos y las diferencias que rodean la tríada catalítica de las tioesterasas, los resultados exitosos de expresión de acil graso-ACP tioesterasas exógenas que se obtuvieron y describieron en los Ejemplos fueron inesperados, especialmente si se tiene en cuenta el hecho de que la actividad de las acil graso-ACP tioesterasas exógenas dependía de una interacción funcional proteína-proteína con la proteína portadora de acilo endógena de *Prototheca*.

Las acil graso-CoA/aldehído reductasas adecuadas para su uso con los microbios y métodos de la invención incluyen, sin limitarse a, los enumerados en la Tabla 4.

Tabla 4. Acil graso-CoA/aldehído reductasas enumeradas por los números de acceso al GenBank.

AAC45217, YP_047869, BAB85476, YP_001086217, YP_580344, YP_001280274, YP_264583, YP_436109, YP_959769, ZP_01736962, ZP_01900335, ZP_01892096, ZP_01103974, ZP_01915077, YP_924106, YP_130411, ZP_01222731, YP_550815, YP_983712, YP_001019688, YP_524762, YP_856798, ZP_01115500, YP_001141848, NP_336047, NP_216059, YP_882409, YP_706156, YP_001136150, YP_952365, ZP_01221833, YP_130076, NP_567936, AAR88762, ABK28586, NP_197634, CAD30694, NP_001063962, BAD46254, NP_001030809, EAZ10132, EAZ43639, EAZ07989, NP_001062488, CAB88537, NP_001052541, CAH66597, CAE02214, CAH66590, CAB88538, EAZ39844, AAZ06658, CAA68190, CAA52019, y BAC84377

Las acil graso-CoA reductasas adecuadas para su uso con los microbios y métodos de la invención incluyen, sin limitarse a, las enumeradas en la Tabla 5.

Tabla 5. Acil graso-CoA reductasas enumeradas por los números de acceso al GenBank.

NP_187805, ABO14927, NP_001049083, CAN83375, NP_191229, EAZ42242, EAZ06453, CAD30696, BAD31814, NP_190040, AAD38039, CAD30692, CAN81280, NP_197642, NP_190041, AAL15288, y NP_190042

Las aldehído graso decarbonilasas adecuadas para su uso con los microbios y métodos de la invención incluyen, sin limitarse a, las enumeradas en la Tabla 6.

Tabla 6. Aldehído graso decarbonilasas enumeradas por los números de acceso al GenBank.

NP_850932, ABN07985, CAN60676, AAC23640, CAA65199, AAC24373, CAE03390, ABD28319, NP_181306, EAZ31322, CAN63491, EAY94825, EAY86731, CAL55686, XP_001420263, EAZ23849, NP_200588, NP_001063227, CAN83072, AAR90847, y AAR97643

Las combinaciones de acil graso-ACP tioesterasas que se co-expresan de manera natural y las proteínas portadoras de acilo son adecuadas para su uso con los microbios y métodos de la invención.

Los ejemplos adicionales de enzimas de modificación de lípido o hidrocarburo incluyen secuencias de aminoácidos contenidas en, referenciadas en, o codificada por secuencias de ácido nucleico contenidas o referenciadas en, cualquiera de las siguientes patentes de Estados Unidos: 6,610,527; 6,451,576; 6,429,014; 6,342,380; 6,265,639; 6,194,185; 6,114,160; 6,083,731; 6,043,072 ; 5,994,114; 5,891,697; 5,871,988; 6,265,639, y descritas en el GenBank con números de acceso: AAO18435; ZP_00513891; Q38710; AAK60613; AAK60610; AAK60611; NP_113747; CAB75874; AAK60612; AAF20201; BAA11024; AF205791; y CAA03710.

Otras enzimas adecuadas para su uso con los microbios y los métodos de la invención incluyen aquellas que tienen al menos el 70% de identidad de aminoácidos con una de las proteínas que se listan en las Tablas 3-6, y que exhiben la correspondiente actividad enzimática deseada (por ejemplo, la escisión de un ácido graso de una proteína portadora de acilo, la reducción de un acil-CoA a un aldehído o un alcohol, o la conversión de un aldehído a un alcano). En realizaciones adicionales, la actividad enzimática está presente en una secuencia que tiene al menos aproximadamente 75%, al menos aproximadamente 80%, al menos aproximadamente 85%, al menos aproximadamente 90%, al menos aproximadamente 95%, o al menos aproximadamente 99% de identidad con una de las secuencias descritas anteriormente, las que se incorporan como referencia en la presente como se exponen completamente.

Al seleccionar la combinación deseada de genes exógenos a expresarse, se puede adaptar el producto generado por el microbio, el cual se puede extraer de la biomasa acuosa. Por ejemplo, el microbio puede contener: (i) un gen exógeno que codifica una acil graso-ACP tioesterasa; y, opcionalmente, (ii) una proteína portadora de acilo

co-expresada u otra proteína portadora de acilo que tiene afinidad por la acil graso-ACP tioesterasa (o viceversa), y, opcionalmente, (iii) un gen exógeno que codifica una acil graso-CoA/aldehído reductasa o una acil graso ACP reductasa y, opcionalmente, (iv) un gen exógeno que codifica una aldehído graso reductasa o una aldehído graso decarbonilasa. El microbio, bajo las condiciones de cultivo que se describen en la presente, sintetiza un ácido graso unido a una ACPp y la acil graso-ACP tioesterasa cataliza la escisión del ácido graso de la ACP para producir, mediante procesamiento enzimático adicional, una molécula de acil-CoA. Cuando está presente, el acil graso-CoA/aldehído reductasa cataliza la reducción de la acil-CoA a un alcohol. Del mismo modo, la acil graso-CoA reductasa, cuando está presente, cataliza la reducción de la acil-CoA a un aldehído. En aquellas realizaciones en las que un gen exógeno que codifica una acil graso-CoA reductasa está presente y se expresa para rendir un producto aldehído, una aldehído graso reductasa, codificada por el tercer gen exógeno, cataliza la reducción del aldehído a un alcohol. Del mismo modo, un aldehído graso decarbonilasa cataliza la conversión del aldehído a un alcano o un alqueno, cuando está presente.

Los genes que codifican tales enzimas se pueden obtener a partir de células que ya se conoce que presentan una producción de lípidos significativa, tales como *Chlorella protothecoides*. Los genes que ya se conoce que tienen un papel en la producción de lípidos, por ejemplo, un gen que codifica una enzima que satura los dobles enlaces, se puede transformar de forma individual dentro de células receptoras. Sin embargo, en la práctica de la invención no es necesario hacer suposiciones *a priori* en cuanto a cuales genes son necesarios. Los métodos para la identificación de genes que pueden alterar (mejorar) la producción de lípidos en las microalgas se describen en la publicación PCT núm. 2008/151149.

Por lo tanto, la presente invención proporciona una célula *Prototheca* que se modificó genéticamente para expresar una enzima de la vía de lípidos a un nivel alterado en comparación con una célula de tipo salvaje de la misma especie. En algunos casos, la célula produce más lípidos en comparación con la célula de tipo salvaje cuando ambas células se cultivan en las mismas condiciones. En algunos casos, la célula se modifica genéticamente y/o se selecciona para expresar una enzima de la vía de lípidos en un nivel más alto que el de las células de tipo salvaje. En algunos casos, la enzima de la vía de lípidos se selecciona del grupo que forman la piruvato deshidrogenasa, la acetil-CoA carboxilasa, la proteína portadora de acilo, y la glicerol-3 fosfato aciltransferasa. En algunos casos, la célula se modifica genéticamente y/o se selecciona para expresar una enzima de la vía de lípidos en un nivel inferior a la célula de tipo salvaje. En al menos una realización en la cual la célula expresa la enzima de la vía de lípidos en un nivel inferior, la enzima de la vía de lípidos comprende la citrato sintasa.

En algunas realizaciones, la célula se modifica genéticamente y/o se selecciona para expresar un regulador global de la síntesis de ácidos grasos en un nivel alterado en comparación con la célula de tipo salvaje, por lo que los niveles de expresión de una pluralidad de genes sintéticos de ácidos grasos están alterados en comparación con la célula de tipo salvaje. En algunos casos, la enzima de la vía de lípidos comprende una enzima que modifica un ácido graso. En algunos casos, la enzima de la vía de lípidos se selecciona de una estearoil-ACP desaturasa y una glicerolípido desaturasa.

En otras realizaciones, la presente invención se dirige a un microbio productor de aceite que contiene uno o más genes exógenos, en donde los genes exógenos codifican la(s) proteína(s) seleccionadas del grupo consistente de una acil graso-ACP tioesterasa, una acil-CoA reductasa, una aldehído graso reductasa, una acil graso-CoA/aldehído reductasa, una aldehído graso decarbonilasa, y una proteína portadora de acilo. En una realización, el gen exógeno se encuentra en unión operativa con un promotor, que se induce o reprime en respuesta a un estímulo. En algunos casos, el estímulo se selecciona del grupo que forman una molécula pequeña que se suministra exógenamente, calor, frío y nitrógeno limitado o nulo en el medio de cultivo. En algunos casos, el gen exógeno se expresa en un compartimento celular. En algunas realizaciones, el compartimento celular se selecciona del grupo que consiste en un cloroplasto, un plastidio y una mitocondria. En algunas realizaciones el microbio es *Prototheca moriformis, Prototheca krugani, Prototheca stagnora* o *Prototheca zopfii.*

En una realización, el gen exógeno codifica una ácido graso acil-ACP tioesterasa. En algunos casos, la tioesterasa codificada por el gen exógeno cataliza la escisión de un ácido graso de 8 a 18 carbonos de una proteína portadora de acilo (ACP). En algunos casos, la tioesterasa codificada por el gen exógeno cataliza la escisión de un ácido graso de 10 a 14 carbonos de una ACP. En una realización, la tioesterasa codificada por el gen exógeno cataliza la escisión de un ácido graso de 12 carbonos de una ACP.

En una realización, el gen exógeno codifica una acil graso-CoA/aldehído reductasa. En algunos casos, la reductasa codificada por el gen exógeno cataliza la reducción de un acil graso-CoA de 8 a 18 carbonos al alcohol primario correspondiente. En algunos casos, la reductasa codificada por el gen exógeno cataliza la reducción de un acil graso-CoA de 10 a 14 carbonos al alcohol primario correspondiente. En una realización, la reductasa codificada por el gen exógeno cataliza la reducción de un acil graso-CoA de 12 carbonos a dodecanol.

La presente invención también proporciona una célula de *Prototheca* recombinante que contiene dos genes exógenos, en donde el primer gen exógeno codifica una acil-ACP tioesterasa y el segundo gen exógeno codifica una proteína que se selecciona del grupo que consiste de una acil-CoA reductasa, una acil graso-CoA/aldehído

reductasa, y una proteína portadora de acilo. En algunos casos, los dos genes exógenos están cada uno en unión operativa con un promotor, el cual es inducible en respuesta a un estímulo. En algunos casos, cada promotor es inducible en respuesta a un estímulo idéntico, tal como nitrógeno limitado o nulo en el medio de cultivo. La limitación o ausencia total de nitrógeno en el medio de cultivo estimula la producción de aceite en algunos microorganismos, tales como la especie Prototheca, y se puede usar como un disparador para inducir la producción de aceite a altos niveles. Cuando se usa en combinación con los métodos de ingeniería genética descritos en la presente, el lípido como un porcentaje del peso seco de las células se puede llevar a altos niveles tales como al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70% y al menos 75%; los métodos descritos en la presente proporcionan células con esos niveles de lípidos, en donde el lípido es al menos 4% C8-C14, al menos 0.3% C8, al menos 2% C10, al menos 2% C12, y al menos 2% C14. En algunas realizaciones las células están por encima de 25% lípido por peso seco de las células y contienen lípido que es al menos 10% C8-C14, al menos 20% C8-C14, al menos 30% C8-C14, 10-30% C8-C14 y 20-30% C8-C14.

Los novedosos aceites que se describen en la presente son distintos de otros aceites de origen natural con alto contenido de ácidos grasos de micro-cadena, tales como el aceite de palma, aceite de almendra de palma y aceite de coco. Por ejemplo, los niveles de contaminantes tales como los carotenoides son mucho más altos en el aceite de palma y el aceite de almendra de palma que en los aceites de la invención. Los aceites de palma y de almendra de palma, en particular, contienen alfa y beta carotenos y licopeno en cantidades mucho más altas que las que hay en los aceites de la invención. Además, se encuentran más de 20 carotenoides diferentes en el aceite de almendra de palma y palma, mientras que los Ejemplos demuestran que los aceites de la invención contienen muy pocas especies de carotenoides y niveles muy bajos. Adicionalmente, los niveles de los compuestos de vitamina E tales como tocotrienoles son mucho más altos en los aceites de palma, almendra de palma y de coco que en los aceites de la invención.

En una realización, la tioesterasa codificada por el primer gen exógeno cataliza la escisión de un ácido graso de 8 a 18 carbonos de una ACP. En algunas realizaciones, el segundo gen exógeno codifica una acil graso-CoA/aldehído reductasa que cataliza la reducción de un acil graso-CoA de 8 a 18 carbonos a un alcohol primario correspondiente. En algunos casos, la tioesterasa codificada por el primer gen exógeno cataliza la escisión de un ácido graso de 10 a 14 de carbono de una ACP, y la reductasa codificada por el segundo gen exógeno cataliza la reducción de un acil graso-CoA de 10 a 14 carbono al correspondiente alcohol primario, en donde la tioesterasa y la reductasa actúan sobre la misma longitud de cadena de carbono. En una realización, la tioesterasa codificada por el primer gen exógeno cataliza la escisión de un ácido graso de 12 carbonos de una ACP, y la reductasa codificada por el segundo gen exógeno cataliza la reducción de un acil graso-CoA de 12 carbonos a dodecanol. En algunas realizaciones, el segundo gen exógeno codifica para una acil graso-CoA reductasa la cual cataliza la reducción de un acil graso-CoA de 8 a 18 carbonos al aldehído correspondiente. En algunas realizaciones, el segundo gen exógeno codifica una proteína portadora de acilo que se co-expresa naturalmente con la acil- ACP tioesterasa.

En algunas realizaciones, el segundo gen exógeno codifica para una acil-CoA reductasa, y el microbio contiene además un tercer gen exógeno que codifica una aldehído graso decarbonilasa. En algunos casos, la tioesterasa codificada por el primer gen exógeno cataliza la escisión de un ácido graso de 8 a 18 carbonos de una ACP, la reductasa codificada por el segundo gen exógeno cataliza la reducción de un acil graso-CoA de 8 a 18 carbonos a un aldehído graso correspondiente, y la decarbonilasa codificada por el tercer gen exógeno cataliza la conversión de un aldehído graso de 8 a18 carbonos a un alcano correspondiente, en donde la tioesterasa, la reductasa, y la decarbonilasa actúan sobre la misma longitud de cadena de carbono.

En algunas realizaciones, el segundo gen exógeno codifica una proteína portadora de acilo, y el microbio contiene además un tercer gen exógeno que codifica una proteína seleccionada del grupo que consiste en una acil-CoA reductasa y una acil graso-CoA/aldehído reductasa. En algunos casos, el tercer gen exógeno codifica una acil graso-CoA reductasa, y el microbio contiene además un cuarto gen exógeno adicional que codifica una aldehído graso decarbonilasa.

La presente invención también proporciona métodos para producir un alcohol que comprende el cultivo de una población de células *Prototheca* recombinantes en un medio de cultivo, en donde las células contienen (i) un primer gen exógeno que codifica una acil graso-ACP tioesterasa, y (ii) un segundo gen exógeno que codifica una acil graso-CoA/aldehído reductasa, y en el cual los microbios sintetizan un ácido graso vinculado a una proteína portadora de acilo (acp), la acil-ACP tioesterasa cataliza la escisión del ácido graso de la ACP para producir, a través de la transformación posterior, una acil-CoA reductasa y la acil graso-CoA/aldehído reductasa cataliza la reducción de la acil-CoA a un alcohol.

La presente invención también proporciona métodos para producir una molécula de lípido en una célula de *Prototheca*. En una realización, el método comprende cultivar una población de células *Prototheca* en un medio de cultivo, en donde las células contienen (i) un primer gen exógeno codifica una acil-ACP tioesterasa, y (ii) un segundo gen exógeno que codifica un acilo graso-CoA reductasa, y en donde los microbios sintetizan un ácido graso enlazado a una proteína portadora de acilo (ACP), la acil-ACP tioesterasa cataliza la escisión del ácido graso de la ACP para producir, a través de la transformación posterior, acil graso-CoA reductasa, y la acil graso-CoA reductasa

cataliza la reducción de la acil-CoA a un aldehído.

La presente invención también proporciona métodos para producir una molécula de ácido graso que tiene una cadena de carbono de longitud especificada en una célula de *Prototheca*. En una realización, el método comprende cultivar una población de células *Prototheca* que produce lípidos en un medio de cultivo, en donde los microbios contienen un gen exógeno que codifica una acil graso-ACP tioesterasa, que tiene una actividad específica o preferencial por cierta longitud de cadena de carbono, tal como 8, 10, 12 o 14 átomos de carbono, y en donde los microbios sintetizan un ácido graso unido a una proteína portadora de acilo (ACP) y la tioesterasa cataliza la escisión del ácido graso de la acp cuando el ácido graso se sintetizó a la longitud específica de la cadena de carbono.

En las varias realizaciones descritas anteriormente, la célula *Prototheca* puede contener al menos un gen exógeno que codifica una enzima de la vía de los lípidos. En algunos casos, la enzima de la vía de los lípidos se selecciona del grupo que consiste en una estearoil ACP desaturasa, una glicerolípido desaturasa, una piruvato deshidrogenasa, una acetil-CoA carboxilasa, una proteína portadora de acilo, y una glicerol-3 fosfato aciltransferasa. En otros casos, la célula *Prototheca* contiene una enzima de modificación de lípidos que se seleccionó del grupo que consiste en una acil-ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, una aldehído graso decarbonilasa, y/o una proteína portadora de acilo

VI. COMBUSTIBLES Y PRODUCCIÓN DE PRODUCTOS QUÍMICOS

Para la producción de combustible de acuerdo con los métodos de la invención, se recolectan los lípidos producidos por las células de la invención, o se recogen de otra manera, por cualquier medio conveniente. Los lípidos se pueden aislar por extracción de las células completas. Primero las células se rompen, y entonces lípidos intracelulares y asociados a la membrana/pared celulares así como los hidrocarburos extracelulares, se pueden separar de la masa celular, tal como mediante el uso de centrifugación como se describe anteriormente. Los lípidos intracelulares producidos en los microorganismos, en algunas realizaciones, se extraen después de las lisis de las células del microorganismo. Una vez extraídos, los lípidos se refinan posteriormente para producir aceites, combustibles o productos oleoquímicos.

Después del completamiento del cultivo, los microorganismos se pueden separar del caldo de fermentación. Opcionalmente, la separación se efectúa por centrifugación para generar una pasta concentrada. La centrifugación no elimina las cantidades significativas de agua intracelular proveniente de los microorganismos y no es una etapa de secado. La biomasa se puede lavar opcionalmente con una solución de lavado (por ejemplo, agua DI) para deshacerse del caldo de fermentación y los residuos celulares. Opcionalmente, la biomasa microbiana lavada también se puede secar (secado en el horno, liofilizada, etc.) antes de la ruptura celular. Por otra parte, las células se pueden lisar sin separación de parte o todo el caldo de fermentación, cuando se completa la fermentación. Por ejemplo, las células pueden estar en una proporción de menos de 1:1, v:v de células con respecto al líquido extracelular cuando ellas se lisan.

Los microorganismos que contiene un lípido se pueden lisar para producir un lisado. Como se detalla en la presente, la etapa de lisar un microorganismo (también conocida como lisis celular) se puede lograr por cualquier medio conveniente, incluyendo lisis inducida por calor, adición de una base, adición de un ácido, uso de enzimas como las proteasas y enzimas de degradación de polisacáridos tales como las amilasas, usando ultrasonido, lisis mecánica, usando choque osmótico, infección con un virus lítico, y/o expresión de uno o más genes líticos. La lisis se realiza para liberar las moléculas intracelulares producidas por el microorganismo. Cada uno de estos métodos para la lisis de un microorganismo se puede usar como un método único o en combinación simultánea o secuencial. La magnitud de la ruptura celular se puede observar por análisis microscópico. Usando uno o varios de los métodos descritos en la presente, se observa típicamente más del 70% de ruptura celular. Preferentemente la ruptura celular es más del 80%, con más preferencia más del 90% y con la máxima preferencia aproximadamente 100%.

En realizaciones particulares, el microorganismo se lisa después del crecimiento, por ejemplo, para incrementar la exposición del lípido celular y/o el hidrocarburo a la extracción o procesamiento posterior. El momento de la expresión de la lipasa (por ejemplo, a través de un promotor inducible) o de la lisis celular, se pueden ajustar para optimizar el rendimiento de los lípidos y/o hidrocarburos. A continuación se describen una serie de técnicas de lisis. Estas técnicas se pueden usar individualmente o en combinación.

En una realización de la presente invención, la etapa de lisis de un microorganismo comprende el calentamiento de una suspensión celular que contiene el microorganismo. En esta realización, el caldo de fermentación que contiene los microorganismos (o una suspensión de microorganismos aislados a partir del caldo de fermentación) se calienta hasta que los microorganismos, es decir, las paredes celulares y las membranas de los microorganismos se degraden o se rompan. Típicamente, las temperaturas que se aplican son al menos de 50˚C. Temperaturas más altas, tales como, al menos 30˚C, al menos 60˚C, al menos 70˚C, al menos 80˚C, al menos 90˚C, al menos 100˚C, al menos 110˚C, al menos 120˚C, al menos 130˚C o superiores, se usan para la lisis celular más eficiente. El lisado de las células mediante tratamiento térmico se puede realizar hirviendo al microorganismo. Alternativamente, el tratamiento térmico (sin hervir) se puede realizar en una autoclave. El lisado tratado

térmicamente se puede enfriar para su posterior tratamiento. La ruptura celular también se puede realizar por tratamiento con vapor, es decir, mediante la adición de vapor a presión. El tratamiento con vapor de las microalgas para la ruptura celular se describe, por ejemplo, en la patente de los Estados Unidos núm. 6,750,048. En algunas realizaciones, el tratamiento con vapor se puede lograr por aspersión de vapor en el fermentador y mantenimiento del caldo de cultivo a una temperatura deseada por menos de aproximadamente 90 minutos, preferentemente menos de aproximadamente 60 minutos, y con más preferencia menos de aproximadamente 30 minutos.

En otra realización de la presente invención, la etapa de lisis de un microorganismo comprende la adición de una base a una suspensión celular que contiene el microorganismo. La base debe ser lo suficientemente fuerte para hidrolizar al menos una porción de los compuestos proteicos de los microorganismos usados. Las bases que son útiles para solubilizar las proteínas se conocen en la materia de la química. Las bases ejemplares que son útiles en los métodos de la presente invención incluyen, pero sin limitarse a, hidróxidos, carbonatos y bicarbonatos de litio, sodio, potasio, calcio y mezclas de los mismos. Una base preferida es el KOH. El tratamiento con base de las microalgas para la ruptura celular se describe, por ejemplo, en la patente de los Estados Unidos núm. 6,750,048.

En otra realización de la presente invención, la etapa de lisis de un microorganismo comprende la adición de un ácido a una suspensión celular que contiene el microorganismo. La lisis ácida se puede llevar a cabo usando un ácido a una concentración de 10 a 500 mN o preferentemente de 40 a 160 nM. La lisis ácida se realiza preferentemente a temperatura por encima de la temperatura ambiente (por ejemplo, 40-160°, y preferentemente una temperatura de 50-130°). Para temperaturas moderadas (por ejemplo, temperatura ambiente a 100°C y particularmente de temperatura ambiente a 65°C), el tratamiento con ácido puede ser útil en combinación con sonicación u otros métodos de o ruptura celular.

En otra realización de la presente invención, la etapa de lisis de un microorganismo comprende la lisis del microorganismo mediante el uso de una enzima. Las enzimas de preferencia para la lisis de un microorganismo son las proteasas y las enzimas que degradan polisacáridos tales como hemicelulasa (por ejemplo, la hemicelulasa de *Aspergillus niger*; Sigma Aldrich, St. Louis, MO; # H2125), pectinasa (por ejemplo, la pectinasa de *Rhizopus sp*.; Sigma Aldrich, St. Louis, MO; # P2401), Mannaway 4.0 L (Novozymes), celulasa (por ejemplo, la celulasa de Trichoderma viride, Sigma Aldrich, St. Louis, MO; # C9422) y driselasa (por ejemplo, la driselasa de *Basidiomycetes sp*.; Sigma Aldrich, St. Louis, MO; # D9515).

En otras realizaciones de la invención, la lisis se lleva a cabo usando una enzima, tal como, por ejemplo, una celulasa tal como una enzima de degradación de polisacáridos, opcionalmente proveniente de *Chlorella* o un virus de *Chlorella*, o una proteasa, tal como la proteasa de *Streptomyces griseus*, quimotripsina, proteinasa K, las proteasas que se listan en Degradation of Polylactide by Commercial Proteases, Oda Y. y otros., Journal of Polymers and the Environment, Volumen 8, número 1, enero 2000, pp. 29-32(4), Alcalase 2.4 FG (Novozymes), y Flavourzyme 100 L (Novozymes). Cualquier combinación de una proteasa y una enzima de degradación de polisacáridos también se puede usar, incluyendo cualquier combinación de las proteasas y las enzimas que degradan polisacáridos anteriores.

En otra realización, la lisis se puede realizar usando una prensa expulsora. En este proceso, la biomasa se fuerza a pasar a través de un dispositivo de tipo rosca a alta presión, lisando las células y causando que los lípidos intracelulares se liberen y se separen de la proteína y la fibra (y los otros componentes) en la célula.

En otra realización de la presente invención, la etapa de lisis de un microorganismo se realiza mediante ultrasonido, es decir, sonicación. De esta manera, las células pueden también lisarse con sonido de alta frecuencia. El sonido se puede producir por vía electrónica y se transporta a través de una punta metálica a una suspensión celular apropiadamente concentrada. Esta sonicación (o ultrasonicación) altera la integridad celular basada en la formación de cavidades en la suspensión celular.

En otra realización de la presente invención, la etapa lisis de un microorganismo se lleva a cabo por lisis mecánica. Las células se pueden lisar mecánicamente y, opcionalmente, homogeneizarse para facilitar la recolección de los hidrocarburos (por ejemplo, los lípidos). Por ejemplo, un disruptor de presión se puede utilizar para bombear una mezcla que contiene las células a través de una válvula de orificio restringido. Se aplica alta presión (hasta 1500 bar), seguido por una expansión instantánea a través de una boquilla de salida. La ruptura celular se lleva a cabo por tres mecanismos diferentes: impacto en la válvula, alta cizalla del líquido en el orificio, y la repentina caída de presión después de la descarga, causando una explosión de la célula. El método libera las moléculas intracelulares. Alternativamente se puede usar un molino de bolas. En un molino de bolas, las células se agitan en suspensión con pequeñas partículas abrasivas, tales como gránulos. Las células se rompen debido a las fuerzas de cizalla, el molido entre los gránulos, y las colisiones con los gránulos. Los gránulos rompen las células para liberar los contenidos celulares. Las células también se pueden romper por las fuerzas de cizalla, tal como con el uso de mezclado (tal como con una licuadora de alta velocidad o Waring como ejemplos), prensa francesa, o incluso centrifugación en el caso de paredes celulares débiles, para romper las células.

En otra realización de la presente invención, la etapa de lisis de un microorganismo se realiza mediante la

aplicación de un choque osmótico.

En otra realización de la presente invención, la etapa de lisis de un microorganismo comprende la infección del microorganismo con un virus lítico. Se conoce una amplia variedad de virus para lisar los microorganismos adecuados para usar en la presente invención, y la selección y el uso de un virus lítico particular para un determinado microorganismo se encuentra dentro del nivel de experiencia en la materia. Por ejemplo, el virus de *paramecium bursaria chlorella* (PBCV-1) es el prototipo de un grupo (familia Phycodnaviridae, género Chlorovirus) de virus de ADN de doble cadena de gran tamaño, icosaédricos, que forman placas, se replican en, y lisan, ciertas algas verdes unicelulares eucariotas similares a Chlorella. En consecuencia, cualquier microalgas sensible se puede lisar mediante infección del cultivo con un virus de chlorella adecuado. Se conocen los métodos de infectar las especies de *Chlorella* con un virus de chlorella. Ver, por ejemplo, *Adv. Virus Res.* 2006; 66: 293-336; *Virology*, 25 de abril de 1999; 257(1):15-23; *Virology*, 5 de enero de 2004; 318(1): 214-23; *Nucleic Acids Symp. Ser.* 2000; (44): 161-2; *J. Virol*. Marzo de 2006 ; 80(5): 2437-44; y *Annu. Rev. Microbiol.* 1999; 53: 447-94.

En otra realización de la presente invención, la etapa de lisis de un microorganismo comprende la autolisis. En esta realización, un microorganismo de la invención se modifica por ingeniería genética para producir una proteína lítica que lisa al microorganismo. Este gen lítico se puede expresar con un promotor inducible, por lo que las células primero se pueden crecer a una densidad deseable en un fermentador, seguido por la inducción del promotor para expresar el gen lítico para lisar las células. En una realización, el gen codifica para una enzima lítica que degrada los polisacáridos. En ciertas otras realizaciones, el gen lítico es un gen de un virus lítico. De esta manera, por ejemplo, un gen lítico proveniente de un virus de Chlorella se puede expresar en una célula de alga, ver *Virology* 260, 308-315 (1999); *FEMS Microbiology Letters* 180 (1999) 45-53; *Virology* 263, 376-387 (1999); y *Virology* 230, 361–368 (1997). La expresión de los genes líticos se hace preferentemente usando un promotor inducible, tal como un promotor activo en microalgas que se induce mediante un estímulo, tal como la presencia de una molécula pequeña, la luz, el calor y otros estímulos.

Existen varios métodos para separar los lípidos a partir de los lisados celulares producidos por los métodos anteriores. Por ejemplo, los lípidos y los derivados de los lípidos, tales como aldehídos grasos, alcoholes grasos, y los hidrocarburos, tales como los alcanos, se pueden extraer con un solvente hidrofóbico como el hexano (ver Frenz y otros, 1989, Enzyme Microb. Technol., 11:717). Los derivados de los lípidos y los lípidos también se pueden extraer usando licuefacción (ver, por ejemplo, Sawayama y otros., 1999, Biomass and Bioenergy 17:33-39 e Inoe y otros, 1993, Biomass Bioenergy 6 (4): 269-274); licuefacción de petróleo (ver, por ejemplo, Minowa y otros, 1995, Fuel 74 (12): 1735-1738); y extracción de CO_2 supercrítico (ver, por ejemplo, Mendes y otros, 2003, Inorganica Chimica Acta 356: 328-334). Miao y Wu describen un protocolo de recuperación de lípidos de microalgas a partir de un cultivo de *Chlorella prototheocoides,* en cual las células se recolectaron por centrifugación, se lavaron con agua destilada y se secaron mediante liofilización. El polvo de las células resultante se pulverizó en un mortero y entonces se extrajeron con *n*-hexano (Miao y Wu, Biosource Technology (2006) 97: 841-846).

De esta manera, los lípidos, derivados de lípidos e hidrocarburos generados por los microorganismos de la presente invención se pueden recuperar mediante extracción con un solvente orgánico. En algunos casos, el solvente orgánico preferido es el hexano. Típicamente, el solvente orgánico se añade directamente al lisado sin separación previa de los componentes del lisado. En una realización, el lisado generado por uno o varios de los métodos descritos anteriormente se pone en contacto con un solvente orgánico por un período de tiempo suficiente para permitir que los lípidos y/o los componentes de los hidrocarburos formen una solución con el solvente orgánico. En algunos casos, la solución puede refinarse adicionalmente para recuperar los lípidos o los componentes de los hidrocarburos específicos deseados. Los métodos de extracción con hexano son bien conocidos en la materia.

Los lípidos y derivados de lípidos, tales como aldehídos grasos, alcoholes grasos, y los hidrocarburos, tales como los alcanos producidos por las células, como se describe en la presente, pueden modificarse mediante el uso de una o más enzimas, incluyendo una lipasa, como se describió anteriormente. Cuando los hidrocarburos se encuentran en el ambiente extracelular de las células, una o más enzimas se pueden añadir a ese ambiente bajo condiciones en las cuales la enzima modifica el hidrocarburo o completa su síntesis a partir de un precursor del hidrocarburo. Alternativamente, los hidrocarburos se pueden aislar parcialmente o completamente del material celular antes de la adición de uno o más catalizadores tales como las enzimas. Estos catalizadores se agregan exógenamente, y su actividad ocurre fuera de la célula o *in vitro*.

De esta manera, los lípidos y los hidrocarburos producidos por las células *in vivo* o modificados enzimáticamente *in vitro*, tal como se describe en la presente, pueden opcionalmente procesarse por medios convencionales. El procesamiento puede incluir "craqueo" (descomposición molecular) para reducir el tamaño, y por lo tanto incrementar la proporción de hidrógeno:carbono de las moléculas de hidrocarburos. Los métodos de craqueo catalítico y térmico se usan rutinariamente en el procesamiento de los hidrocarburos y de los aceites de triglicéridos. Los métodos catalíticos involucran el uso de un catalizador, tal como un catalizador ácido sólido. El catalizador puede ser de sílice-alúmina o una zeolita, los cuales resultan en la ruptura heterolítica o asimétrica, de un enlace carbono-carbono para resultar en un carbocatión y un anión hidruro. Estos intermediarios reactivos entonces sufren ya bien o un reordenamiento o la transferencia del hidruro con otro hidrocarburo. Las reacciones, por lo tanto,

pueden regenerar los intermediarios para dar lugar a un mecanismo de cadena auto-multiplicativa. Los hidrocarburos también pueden procesarse para reducir, opcionalmente a cero, el número de dobles enlaces o triples carbono-carbono de los mismos. Los hidrocarburos también pueden procesarse para quitar o eliminar un anillo o estructura cíclica en ellos. Los hidrocarburos también pueden procesarse para aumentar la proporción de hidrógeno:carbono. Esto puede incluir la adición de hidrógeno (hidrogenación) y/o el "craqueo" de los hidrocarburos en otros hidrocarburos más pequeños.

Los métodos térmicos involucran el uso de temperatura y presión elevadas para reducir el tamaño del hidrocarburo. Se puede usar una temperatura elevada de aproximadamente 800°C y una presión de aproximadamente 700kPa. Estas condiciones generan "luz", un término que se usa a veces para referirse a las moléculas de hidrocarburos ricas en hidrógeno (a diferencia de las del flujo de fotones), mientras generan además, por condensación, moléculas de hidrocarburos más pesadas las cuales carecen relativamente de hidrógeno. La metodología proporciona la ruptura homolítica o simétrica, y produce alquenos que pueden saturarse enzimáticamente de forma opcional, como se describe anteriormente.

Los métodos catalíticos y térmicos son estándares en las plantas de procesamiento de hidrocarburos y refinación de petróleo. De esta manera, los hidrocarburos que se producen por las células, como se describe en la presente, pueden recolectarse y procesarse o refinarse a través de medios convencionales. Ver Hillen y otros (Biotechnology and Bioengineering, Vol. XXIV: 193-205 (1982)), una publicación sobre hidrocraqueo de hidrocarburos producidos por microalgas. En realizaciones alternativas, la fracción se trata con otro catalizador, tal como un compuesto orgánico, calor y/o un compuesto inorgánico. Para el procesamiento de los lípidos en biodiesel, se usa un proceso de transesterificación como se describe en la Sección IV del presente documento.

Los hidrocarburos producidos a través de los métodos de la presente invención, son útiles en una variedad de aplicaciones industriales. Por ejemplo, la producción de sulfonato de alquilbenceno lineal (LAS), un surfactante aniónico usado en casi todos los tipos de detergentes y productos de limpieza, utiliza los hidrocarburos que comprenden generalmente una cadena de 10-14 átomos de carbono. Ver, por ejemplo, las patentes de los Estados Unidos núms. 6,946,430; 5,506,201; 6,692,730; 6,268,517; 6,020,509; 6,140,302; 5,080,848 y 5,567,359. Los surfactantes, tales como LAS, se pueden usar en la fabricación de composiciones para el cuidado personal y detergentes, tales como los descritos en las patentes de los Estados Unidos núms. 5,942,479; 6,086,903; 5,833,999; 6,468,955 y 6,407,044.

Un interés creciente se dirige al uso de los componentes de los hidrocarburos de origen biológico en los combustibles, tales como biodiesel, diesel renovable y combustible de reactor, ya que las materias primas biológicas renovables pueden sustituir a las materias primas derivadas de los combustibles fósiles que están disponibles, y su uso, por lo tanto, es deseable. Existe una necesidad urgente de métodos para la producción de componentes de hidrocarburos a partir de materiales biológicos. La presente invención satisface esta necesidad al proporcionar métodos para la producción de biodiesel, diesel renovable y combustible de reactor usando los lípidos generados por los métodos descritos en la presente como un material biológico para producir biodiesel, diesel renovable y combustible de reactor.

Los combustibles diesel tradicionales son destilados del petróleo ricos en hidrocarburos parafínicos. Ellos tienen intervalos de ebullición tan amplios como de 370° a 780°F, los cuales son adecuados para la combustión en un motor de ignición por compresión, tal como un vehículo de motor diesel. La American Society of Testing and Materials (ASTM), establece el grado de diesel de acuerdo con el intervalo de ebullición, junto con los intervalos permisibles de otras propiedades del combustibles, tales como el número de cetano, punto de nube, punto de inflamación, viscosidad, punto de anilina, contenido de azufre, contenido de agua, contenido de cenizas, corrosión de la tira de cobre, y residuos de carbono. Técnicamente, cualquier material destilado de hidrocarburos derivados de biomasa o de otro tipo que cumpla con la especificación apropiada de la ASTM, se puede definir como combustible diesel (ASTM D975), combustible de reactor (ASTM D1655), o como biodiesel o si es un éster metílico de ácido graso (ASTM D6751).

Después de la extracción, los lípidos y/o los componentes de los hidrocarburos recuperados a partir de la biomasa microbiana descrita en la presente, se pueden someter a tratamiento químico para fabricar un combustible para su uso en vehículos a diesel y motores a reacción.

El biodiesel es un líquido que varía en color – entre dorado y marrón oscuro – en dependencia de la materia prima usada en la producción. Es prácticamente inmiscible en agua, tiene un alto punto de ebullición y baja presión de vapor. El biodiesel se refiere a un combustible procesado equivalente al diesel para su uso en vehículos con motor diesel. El biodiesel es biodegradable y no tóxico. Un beneficio adicional del biodiesel con respecto a los combustibles diesel convencionales, es el menor desgaste del motor. Típicamente, el biodiesel comprende ésteres de alquilo de C14-C18. Varios procesos convierten la biomasa o un lípido producido y aislado según lo descrito en la presente, en los combustibles diesel. Un método preferido para producir biodiesel es la transesterificación de un lípido como se describe en la presente. Un éster de alquilo preferido para su uso como biodiesel es un éster metílico o éster etílico.

El biodiesel producido por un método descrito en la presente se puede usar solo o mezclado con combustible diesel convencional en cualquier concentración en la mayoría de los vehículos modernos de motores diesel. Cuando se mezcla con el combustible diesel convencional (diesel), el biodiesel puede estar presente en aproximadamente de 0.% a aproximadamente 99.9%. Gran parte del mundo usa un sistema conocido como el factor "B", para indicar la cantidad de biodiesel en cualquier mezcla de combustible. Por ejemplo, el combustible que contiene 20% de biodiesel se etiqueta como B20. El biodiesel puro se refiere como B100.

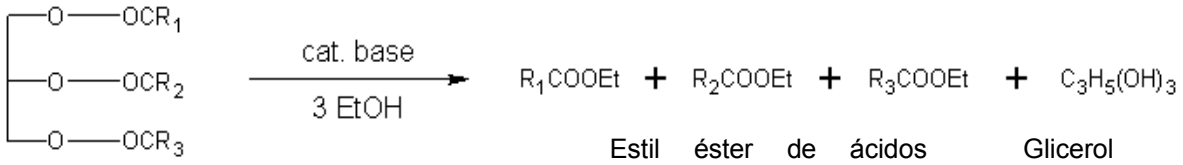
El biodiesel también puede usarse como combustible de calefacción en calderas domésticas y comerciales. Las calderas de aceite que existen pueden contener partes de caucho, y pueden requerir conversión para funcionar con biodiesel. El proceso de conversión es relativamente sencillo, el cual involucra el intercambio de las partes de caucho por partes sintéticas debido a que el biodiesel es un fuerte solvente. Debido a su fuerte poder solvente, el quemado del biodiesel incrementará la eficiencia de las calderas. El biodiesel puede usarse como un aditivo en las formulaciones de diesel para incrementar la lubricidad del combustible Ultra-Low Sulfur Diesel (ULSD) puro, lo cual es ventajoso, ya que éste prácticamente no tiene contenido de azufre. El biodiesel es un mejor solvente que el petrodiesel y se puede usar para separar los depósitos de los residuos en las líneas de combustible de los vehículos que previamente han funcionado con el petrodiesel.

El biodiesel puede producirse por transesterificación de los triglicéridos contenidos en la biomasa rica en aceites. Por lo tanto, en otro aspecto de la presente invención se proporciona un método para producir biodiesel. En una realización preferida, el método para producir biodiesel comprende las etapas de (a) cultivar un microorganismo que contiene lípidos usando los métodos descritos en la presente, (b) lisar un microorganismo que contiene lípidos para producir un lisado, (c) aislar los lípidos a partir del microorganismo lisado y (d) transesterificar la composición lipídica, mediante lo cual se produce el biodiesel. Los métodos para el crecimiento de un microorganismo, lisar un microorganismo para producir un lisado, tratar el lisado en un medio que comprende un solvente orgánico para formar una mezcla heterogénea y separar el lisado tratado en una composición lipídica, se han descrito anteriormente y también se pueden usar en el método de producción de biodiesel.

El perfil lipídico del biodiesel usualmente es muy similar al perfil de lípidos del aceite como materia prima. Los demás aceites proporcionados por los métodos y las composiciones de la invención se pueden someter a transesterificación para producir biodiesel con los perfiles de lípidos, que incluyen (a) al menos 4% de C8-C14, (b) al menos 0,3% de C8; (c) al menos 2% de C10, (d) al menos 2% de C12, y (e) al menos 30% de C8-C14.

Las composiciones de lípidos se pueden someter a transesterificación para producir ésteres de ácidos grasos de cadena larga, útiles como biodiesel. Las reacciones de transesterificación preferidas se describen a continuación e incluyen transesterificación catalizada por bases y transesterificación usando lipasas recombinantes. En un proceso de transesterificación catalizada por bases, los triglicéridos se hacen reaccionar con un alcohol, tal como metanol o etanol, en presencia de un catalizador alcalino, típicamente hidróxido de potasio. Esta reacción forma ésteres metílicos o etílicos y glicerina (glicerol) como un subproducto.

Los aceites animales y vegetales se hacen típicamente de triglicéridos que son ésteres de ácidos grasos libres con el alcohol trihídrico, glicerol. En la transesterificación, el glicerol en un triacilglicérido (TAG) se sustituye por un alcohol de cadena corta tal como el metanol o el etanol. Un esquema de la reacción típica es el siguiente:



Triglicérido Estil éster de ácidos Glicerol

En esta reacción, el alcohol se desprotona con una base para hacerlo un nucleófilo más fuerte. Comúnmente, el etanol o metanol se usan en gran exceso (hasta 50 veces). Normalmente, esta reacción se llevará a cabo de forma excesivamente lenta o no ocurrirá en lo absoluto. El calor, así como un ácido o una base, se pueden usar para ayudar a la reacción a que se lleve a cabo más rápidamente. El ácido o la base no se consumen por la reacción de transesterificación, por lo que no son reactivos sino catalizadores. Casi todo el biodiesel se produce usando la técnica de catalización con base, ya que ésta sólo requiere bajas temperaturas y presiones y produce por encima del 98% de rendimiento de conversión (siempre y cuando el aceite de partida sea bajo en humedad y libre de ácidos grasos).

La transesterificación también se lleva a cabo, como se mencionó anteriormente, usando una enzima, tal como una lipasa en lugar de una base. La transesterificación catalizada por la lipasa se puede llevar a cabo, por ejemplo, a una temperatura entre la temperatura ambiente y 80□C, y una relación molar del TAG con respecto al alcohol inferior mayor de 1:1, preferentemente aproximadamente 3:1. Las lipasas adecuadas para su uso en la

transesterificación incluyen, pero no se limitan, a las que se enumeran en la Tabla 7. Otros ejemplos de lipasas útiles para la transesterificación se encuentran en, por ejemplo, las patentes de los Estados Unidos núms. 4,798,793; 4,940,845 5,156,963; 5,342,768; 5,776,741 y WO89/01032. Tales lipasas incluyen, pero sin limitarse a, lipasas producidas por microorganismos de Rhizopus, Aspergillus, Candida, Mucor, Pseudomonas, Rhizomucor, Candida, y Humicola y lipasa del páncreas.

Tabla 7. Lipasas adecuadas para su uso en la transesterificación.

Lipasa de *Aspergillus niger* ABG73614, lipasa B de*Candida antarctica* (novozym-435) CAA83122, lipase de *Candida cylindracea* AAR24090, lipasa de *Candida lipolytica* (Lipasa L; Amano Pharmaceutical Co., Ltd.), lipasa de *Candida rugosa* (por ejemplo, Lipasa-OF; Meito Sangyo Co., Ltd.), lipasa de *Mucor miehei* (Lipozyme IM 20), lipasa de *Pseudomonas fluorescens* AAA25882, lipasa de *Rhizopus japonicas* (Lilipase A-10FG) Q7M4U7_1, lipasa de*Rhizomucor miehei* B34959, lipasa de *Rhizopus oryzae* (Lipasa F) AAF32408, lipasa de *Serratia marcescens* (Enzima SM) ABI13521, lipasa de *Thermomyces lanuginosa* CAB58509, Lipasa P (Nagase ChemteX Corporation), y Lipasa QLM (Meito Sangyo Co., Ltd., Nagoya, Japón)

Un reto para usar una lipasa para la producción de ésteres de ácido graso adecuados para biodiesel es que el precio de la lipasa es mucho mayor que le precio del hidróxido sódico (NaOH) usado por los procesos con base fuerte. Este reto se abordó usando una lipasa inmovilizada, que se puede reciclar. Sin embargo, la actividad de la lipasa inmovilizada debe mantenerse después de reciclarse por un número mínimo de ciclos para permitir un proceso basado en lipasa que compita con el proceso de base fuerte en términos de costo de producción. Las lipasas inmovilizadas se someten a envenenamiento por los alcoholes inferiores típicamente usados en la transesterificación. La patente de los Estados Unidos núm. 6,398,707 (concedida el 04 de junio 2002 a Wu y otros) describe los métodos para incrementar la actividad de las lipasas inmovilizadas y la regeneración de las lipasas inmovilizadas con actividad reducida. Algunos métodos adecuados incluyen sumergir una lipasa inmovilizada en un alcohol que tiene un número de átomos de carbono no menor que 3 durante un período de tiempo, preferentemente de 0.5-48 horas, y con la máxima preferencia de 0.5-1.5 horas. Algunos métodos adecuados también incluyen el lavado de una lipasa inmovilizada desactivada con un alcohol que tiene un número de átomos de carbono no menor que 3, y luego sumergir la lipasa inmovilizada desactivada en un aceite vegetal durante 0.5-48 horas.

En realizaciones particulares, una lipasa recombinante se expresa en los mismos microorganismos que producen los lípidos sobre los cuales actúa la lipasa. Las lipasas recombinante adecuadas incluyen las que se listan más arriba en la Tabla 7 y/o que tienen los números de acceso en el GenBank que se enumeran en la Tabla 7 anterior, o un polipéptido que tiene al menos el 70% de identidad de aminoácidos con una de las lipasas enumeradas anteriormente en la Tabla 7 y que exhibe actividad de lipasa. En las realizaciones adicionales, la actividad enzimática está presente en una secuencia que tiene al menos aproximadamente 75%, al menos aproximadamente 80%, al menos aproximadamente 85%, al menos aproximadamente 90%, al menos aproximadamente 95%, o al menos aproximadamente 99% de identidad con una de las secuencias descritas anteriormente, todas las cuales se incorporan en la presente mediante referencia, como si estuvieran completamente definidas. El ADN que codifica para la lipasa y el marcador seleccionable es preferentemente ADNc de codón optimizado. Los métodos de recodificar los genes para la expresión en microalgas se describen en la patente de los Estados Unidos núm. 7,135,290.

La norma internacional común para el biodiesel es EN 14214. ASTM D6751, la norma de biodiesel más común que de referencia en los Estados Unidos y Canadá. Alemania usa la norma DIN EN 14214 y el Reino Unido exige el cumplimiento de la norma BS EN 14214. Las pruebas industriales básicas para determinar si los productos se ajustan a estas normas típicamente incluyen cromatografía de gases, HPLC, y otras. El biodiesel que cumple las normas de calidad no es muy tóxico, con un índice de toxicidad (DL_{50}) mayor de 50 ml/kg.

Aunque el biodiesel que cumpla con las normas ASTM tiene que ser no tóxico, puede haber contaminantes que tiendan a cristalizar y/o precipitar y se desprendan de la solución como sedimentos. La formación de sedimentos es particularmente un problema cuando el biodiesel se usa a temperaturas más bajas. El sedimento o los precipitados pueden causar problemas tales como la disminución del flujo del combustible, la obstrucción de las tuberías de combustible, la obstrucción de filtros, etc. Existen procesos bien conocidos en la materia que tratan específicamente de la eliminación de estos contaminantes y sedimentos en el biodiesel, con el objetivo de producir un producto de mayor calidad. Los ejemplos de estos procesos incluyen, pero sin limitarse al tratamiento previo del aceite para eliminar contaminantes tales como fosfolípidos y ácidos grasos libres (*por ejemplo*, desgomado, refinado cáustico y filtración con adsorbente de sílice) y filtración en frío. La filtración en frío es un proceso que se desarrolló específicamente para eliminar cualquiera de las partículas y sedimentos que están presentes en el biodiesel después de la producción. Este proceso enfría al biodiesel y filtra cualquiera de los sedimentos o precipitados que puedan formarse cuando el combustible se usa en una temperatura más baja. Este proceso es bien conocido en la materia y se describe en la publicación de la solicitud de patente núm. 2007-0175091. Los métodos adecuados pueden incluir el enfriamiento del biodiesel a una temperatura inferior a aproximadamente 38°C, de manera que las impurezas y los contaminantes precipiten como partículas en el líquido del biodiesel. La tierra de diatomeas u otro material de filtrado se pueden entonces añadir al biodiesel enfriado para formar una pasta, que luego puede filtrarse

a través de una hoja de presión u otro tipo de filtro para eliminar las partículas. El biodiesel filtrado, a continuación se puede pasar a través de un filtro de pulido para eliminar los sedimentos restantes y la tierra de diatomeas, con el fin de producir el producto de biodiesel final.

El Ejemplo 14 describe la producción de biodiesel usando aceite de triglicéridos preveniente de *Prototheca moriformis*. La Cold Soak Filterability según el método de ASTM D6751 A1 del biodiesel producido en el Ejemplo 14 fue de 120 segundos para un volumen de 300 ml. Esta prueba involucra la filtración de 300 ml de B100, enfriado a 40°F durante 16 horas, se le deja calentar a temperatura ambiente, y se filtra al vacío usando fibra de vidrio de 0.7 micras con soporte de acero inoxidable. Los aceites de la invención pueden transesterificarse para generar biodiesel con un tiempo de maceración en frío de menos de 120 segundos, menos de 100 y menos de 90 segundos.

Procesos posteriores también se pueden usar si el biodiesel se usa particularmente a temperaturas frías. Estos procesos incluyen el acondicionamiento para el invierno y el fraccionamiento. Ambos procesos se diseñan para mejorar el flujo en frío y el desempeño en el invierno del combustible mediante la reducción del punto de nube (la temperatura a la que el biodiesel empieza a cristalizar). Existen varios enfoques para preparar el biodiesel para el invierno. Un método consiste en mezclar el biodiesel con diesel de petróleo. Otro enfoque es usar aditivos que pueden bajar el punto de nube del biodiesel. Otro enfoque consiste en eliminar los ésteres de metilo saturados indiscriminadamente, mediante la mezcla en aditivos, permitiendo la cristalización de las grasas saturadas y luego filtrando los cristales. El fraccionamiento separa selectivamente los ésteres de metilo en componentes o fracciones individuales, lo que permite la eliminación o inclusión de los ésteres de metilo específicos. Los métodos de fraccionamiento incluyen el fraccionamiento con urea, fraccionamiento con solvente y destilación térmica.

Otro valioso combustible proporcionado por los métodos de la presente invención es el diesel renovable, el cual comprende alcanos, tales como C10:0, C12:0, C14:0, C16:0 y C18: 0, y por lo tanto se distingue del biodiesel. El diesel renovable de alta calidad se obtiene conforme a la norma ASTM D975. Los lípidos producidos por los métodos de la presente invención pueden servir como materia prima para producir el diesel renovable. Los lípidos producidos por los métodos de la presente invención pueden servir como materia prima para producir el diesel renovable. Por lo tanto, en otro aspecto de la presente invención, se proporciona un método para producir diesel renovable. Durante estos procesos, los triacilglicéridos producidos y aislados como se describe en la presente, se convierten a alcanos.

En una realización, el método para la producción del diesel renovable comprende (a) cultivar un microorganismo que contiene lípido usando los métodos descritos en la presente, (b) lisar el microorganismo para producir un lisado, (c) aislar el lípido a partir del microorganismo lisado, y (d) desoxigenar e hidrotratar el lípido para producir un alcano, produciéndose así el diesel renovable. Los lípidos adecuados para la fabricación del diesel renovable se pueden obtener a través de la extracción proveniente de la biomasa microbiana usando un solvente orgánico tal como el hexano, o a través de otros métodos, tales como los descritos en la patente de los Estados Unidos núm. 5,928,696. Algunos métodos apropiados pueden incluir prensado mecánico y centrifugación.

En algunos métodos, el lípido microbiano primero se craquea conjuntamente con el hidrotratamiento para reducir la longitud de la cadena de carbono y saturar los dobles enlaces, respectivamente. Después, el material se isomeriza conjuntamente también con el hidrotratamiento. La fracción de nafta se puede eliminar después mediante destilación, seguido de una destilación adicional para evaporar y destilar los componentes deseados en el combustible diesel con vistas a cumplir con una norma ASTM D975, mientras que los componentes más pesados de lo deseado se dejan para cumplir con la norma D975. Los métodos de hidrotratamiento, hidrocraqueo, desoxigenación e isomerización de los aceites modificados químicamente, incluyendo los aceites de triglicéridos, son bien conocidos en la materia. Ver por ejemplo las solicitudes de patente europea EP1741768 (A1); EP1741767 (A1); EP1682466 (A1); EP1640437 (A1); EP1681337 (A1); EP1795576 (A1); y las patentes de los Estados Unidos 7,238,277; 6,630,066; 6,596,155; 6,977,322; 7,041,866; 6,217,746; 5,885,440; 6,881,873.

En una realización del método de producción del diesel renovable, el tratamiento del lípido para producir un alcano se lleva a cabo por hidrotratamiento de la composición lipídica. En el procesamiento hidrotérmico, típicamente la biomasa se hace reaccionar en agua a una temperatura y presión elevadas para formar aceites y residuos sólidos. Las temperaturas de conversión son típicamente de 300º a 660ºF, con presión suficiente para mantener el agua primariamente como un líquido, de 100 a 170 atmósferas como estándar. Los tiempos de reacción son del orden de 15 a 30 minutos. Después que la reacción se completa, los orgánicos se separan del agua. De tal modo se produce un destilado apropiado para el diesel.

En algunos métodos para preparar el diesel renovable, la primera etapa del tratamiento de un triglicérido es el hidroprocesamiento para saturar los dobles enlaces, seguido por desoxigenación a temperatura elevada en presencia de hidrógeno y un catalizador. En algunos métodos, la hidrogenación y desoxigenación ocurren en la misma reacción. En otros métodos la desoxigenación ocurre antes de la hidrogenación. La isomerización se realiza entonces opcionalmente, también en presencia de hidrógeno y un catalizador. Los componentes de nafta se eliminan preferentemente por medio de destilación. Como ejemplos, ver las patentes de los Estados Unidos 5,475,160 (hidrogenación de los triglicéridos), 5,091,116 (desoxigenación, hidrogenación y extracción de gas),

6,391,815 (hidrogenación), y 5,888,947 (isomerización).

Un método apropiado para la hidrogenación de los triglicéridos incluye preparar una solución acuosa de sales de cobre, zinc, magnesio y lantano y otra solución de metales alcalinos o preferentemente, carbonato de amonio. Las dos soluciones se pueden calentar a una temperatura de aproximadamente 20°C a aproximadamente 85°C y se miden en conjunto en un recipiente de precipitación a velocidades tales que el pH en el recipiente de precipitación se mantenga entre 5.5 y 7.5 con el objetivo de formar un catalizador. Agua adicional se puede usar inicialmente en el recipiente de precipitación o añadirse al mismo tiempo que la solución de sal y la solución de precipitación. El precipitado resultante puede lavarse completamente, secarse, calcinarse a aproximadamente 300°C y activarse en hidrógeno a temperaturas que se encuentran en el intervalo de aproximadamente 100°C a aproximadamente 400°C. Uno o más triglicéridos se pueden contactar después y reaccionar con hidrógeno en presencia del catalizador descrito anteriormente en un reactor. El reactor puede ser un reactor de lecho de goteo, un reactor gas-sólido de lecho fijo, un reactor de columna de burbujas empacada, un reactor de tanque de agitación continua, un reactor de fase de suspensión, o cualquier otro tipo de reactor adecuado conocido en la materia. El proceso se puede llevar a cabo por lotes o en forma continua. Las temperaturas de reacción están típicamente en el intervalo de aproximadamente 170°C a aproximadamente 250°C, mientras que las presiones de reacción están típicamente en el intervalo de aproximadamente 300 psig a aproximadamente 2000 psig. Por otra parte, la relación molar del hidrógeno con respecto a los triglicéridos en el proceso de la presente invención, está típicamente en el intervalo de aproximadamente 20:1 a aproximadamente 700:1. El proceso se realiza, típicamente, a una velocidad espacial horaria por peso (WHSV) en el intervalo de aproximadamente 0.1 h^{-1} a aproximadamente 5 h^{-1}. Un experto en la materia reconocerá que el período de tiempo requerido para la reacción variará en función de la temperatura usada, la relación molar del hidrógeno con respecto a los triglicéridos y la presión parcial de hidrógeno. Los productos producidos por tales procesos de hidrogenación, incluyen alcoholes grasos, glicerol, rastros de parafinas y triglicéridos sin reaccionar. Estos productos se separan normalmente por medios convencionales tales como, por ejemplo, destilación, extracción, filtración, cristalización, y similares.

Las refinerías de petróleo usan el hidroprocesamiento para eliminar las impurezas mediante el tratamiento de las fuentes con hidrógeno. Las temperaturas de conversión del hidroprocesamiento son típicamente de 300º a 700º F. Las presiones son normalmente de 40 a 100 atmósferas. Los tiempos de reacción están típicamente en el orden de 10 a 60 minutos. Los catalizadores sólidos se emplean para aumentar ciertas velocidades de reacción, mejorar la selectividad de determinados productos, y optimizar el consumo de hidrógeno.

Los métodos adecuados para la desoxigenación de un aceite incluyen calentar un aceite a una temperatura en el intervalo de aproximadamente 350°F hasta aproximadamente 550°F y mantener el aceite caliente continuamente en contacto con el nitrógeno, bajo al menos una presión que se encuentra en el intervalo desde aproximadamente la presión atmosférica hasta las anteriores, durante al menos aproximadamente 5 minutos.

Los métodos más adecuados para la isomerización incluyen el uso de isomerización alcalina y otra isomerización de aceites conocida en la materia.

El hidrotratamiento y el hidroprocesamiento en última instancia conducen a una reducción en el peso molecular de la fuente de los triglicérido. La molécula de triglicéridos se reduce a cuatro moléculas de hidrocarburos bajo las condiciones de hidroprocesamiento: una molécula de propano y tres moléculas de hidrocarburos más pesados, típicamente en el intervalo de C8 a C18.

Así, en una realización, el producto de una o más reacción(es) química(s) que se llevan a cabo en las composiciones lipídicas de la invención, es una mezcla de alcanos que comprende el diesel renovable de acuerdo a la norma ASTM D975. La producción de hidrocarburos por microorganismos es revisada por Metzger y otros, Appl Microbiol Biotechnol (2005) 66: 486-496 y A Look Back en the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae, NREL/TP-580-24190, John Sheehan, Terri Dunahay, John Benemann y Paul Roessler (1998).

Las propiedades de la destilación de un combustible diesel se describen en términos de T10-T90 (temperatura a 10% y 90%, respectivamente, volumen destilado). El diesel renovable se produjo a partir del aceite de triglicéridos de *Prototheca moriformis* y se describe en el Ejemplo 14. La T10-T90 del material producido en el Ejemplo 14 fue de 57.9°C. Los métodos de hidrotratamiento, isomerización, y otra modificación covalente de los aceites que se divulgan en la presente, así como los métodos de destilación y fraccionamiento (tal como el filtrado en frío) descritos en la presente, se pueden emplear para generar composiciones de diesel renovable con otros intervalos de T90-T10, tales como 20, 25, 30, 35, 40, 45, 50, 60 y 65°C usando los aceites de triglicéridos producidos de acuerdo con los métodos divulgados en la presente.

La T10 del material producido en el Ejemplo 14 fue 242.1ºC. Los métodos de hidrotratamiento, isomerización, y otra modificación covalente de los aceites divulgados en la presente, así como los métodos de destilación y fraccionamiento (como el filtrado en frío) descritos en la presente, se pueden emplear para generar composiciones de diesel renovable con otros valores de T10, tales como T10 entre 180 y 295, entre 190 y 270, entre

210 y 250, entre 225 y 245, y al menos de 290.

La T90 del material producido en el Ejemplo 14 fue de 300°C. Los métodos de hidrotratamiento, isomerización, y otra modificación covalente de los aceites divulgados en la presente, así como los métodos de destilación y fraccionamiento (como el filtrado en frío) descritos en la presente, se pueden emplear para generar composiciones de diesel renovable con otros valores de T90, tales como T90 entre 280 y 380, entre 290 y 360, entre 300 y 350, entre 310 y 340, y al menos de 290.

La FBP del material producido en el Ejemplo 14 fue de 300°C. Los métodos de hidrotratamiento, isomerización, y otra modificación covalente de los aceites divulgados en la presente, así como los métodos de destilación y fraccionamiento (como el filtrado en frío) descritos en la presente, se pueden emplear para generar composiciones de diesel renovable con otros valores de FBP, tales como FBP entre 290 y 400, entre 300 y 385, entre 310 y 370, entre 315 y 360, y de al menos 300.

Otros aceites proporcionados por los métodos y las composiciones de la invención se pueden someter a combinaciones de hidrotratamiento, isomerización, y otra modificación covalente, incluidos los aceites con perfiles de lípidos que incluyen (a) al menos 4% de C8-C14, (b) al menos 0,3% de C8, (c) al menos 2% de C10; (d) al menos 2% de C12 y (e) al menos 30% de C8-C14.

Un diesel tradicional ultra bajo en azufre se puede producir a partir de cualquier forma de biomasa mediante un proceso de dos pasos. Primero, la biomasa se convierte en un gas sintético, una mezcla gaseosa rica en hidrógeno y monóxido de carbono. Entonces el gas sintético se convierte catalíticamente en líquidos. Típicamente, la producción de líquidos se logra usando la síntesis de Fischer-Tropsch (FT). Esta tecnología se aplica al carbón, gas natural y aceites pesados. Por lo tanto, en otra realización preferida del método para la producción de diesel renovable, el tratamiento de la composición lipídica para producir un alcano se realiza mediante la licuefacción indirecta de la composición lipídica.

La presente invención también proporciona métodos para producir combustible de reactor. El combustible de avión es de color claro a pajizo. El combustible más común es un combustible a base de aceite sin plomo/parafina clasificado como Aeroplane A-1, el cual se produce para un conjunto especificaciones normalizadas a nivel internacional. El combustible de avión es una mezcla de un gran número de diferentes hidrocarburos, posiblemente tanto como mil o más. El intervalo de su tamaño (peso molecular o números de carbono) está limitado por los requerimientos para el producto, por ejemplo, el punto de congelación o el punto de humo. El combustible de avión tipo Kerosone (incluyendo el Jet A y el Jet A-1) tiene una distribución del número de carbonos entre aproximadamente 8 y 16 números de carbono. El combustible de avión de tipo nafta o wide-cut (incluyendo el Jet B) típicamente tiene una distribución del número de carbonos entre aproximadamente 5 y 15 carbonos.

Ambos Aeroplanes (Jet A y Jet B) pueden contener un número de aditivos. Los aditivos útiles incluyen, pero sin limitarse a, antioxidantes, agentes antiestáticos, inhibidores de la corrosión, y agentes inhibidores de la congelación del sistema de combustible (FSII). Los antioxidantes previenen el engomado y usualmente se basan en fenoles alquilados, por ejemplo, AO-30, AO-31, o AO-37. Los agentes antiestáticos disipan la electricidad estática y evitan las chispas. El Stadis 450 con ácido dinonilnaftilsulfónico (DINNSA) como ingrediente activo, es un ejemplo. Los inhibidores de la corrosión, por ejemplo, el DNI-4A, se usa para combustibles civiles y militares y el ICD-6A se usa para combustibles militares. Los agentes FSII incluyen, por ejemplo, el Di-EGME.

En una realización de la invención, un combustible de reactor se produce mediante la mezcla de combustibles de algas con los combustibles existentes. Los lípidos producidos por los métodos de la presente invención pueden servir como materia prima para producir combustible de reactor. Por lo tanto, en otro aspecto de la presente invención, se proporciona un método para producir combustible de reactor. En la presente se proporcionan dos métodos para producir combustible de reactor a partir de los lípidos producidos por los métodos de la presente invención: craqueo catalítico fluido (FCC), e hidrodesoxigenación (HDO).

El craqueo catalítico fluido (FCC) es un método que se usa para producir olefinas, especialmente propileno a partir de fracciones pesadas del crudo. Los lípidos producidos por el método de la invención se pueden convertir en olefinas. El proceso involucra el flujo de los lípidos producidos a través de una zona de FCC y se recoge una corriente del producto compuesto de olefinas, el cual es útil como combustible de reactor. Los lípidos producidos se ponen contacto con un catalizador de craqueo en condiciones de craqueo para proporcionar una corriente del producto que comprende olefinas e hidrocarburos, útiles como combustible de reactor.

En una realización, el método para producir combustible de reactor comprende (a) cultivar un microorganismo que contiene lípido usando los métodos descritos en la presente, (b) lisar el microorganismo que contiene lípido para producir un lisado, (c) aislar el lípido a partir del lisado, y (d) tratar la composición lipídica, por medio de lo cual se produce el combustible de reactor. En una realización del método para producir un combustible de reactor, la composición lipídica puede fluir a través de una zona de craqueo catalítico fluido, la cual, en una realización, puede comprender el contacto de la composición lipídica con un catalizador de craqueo en condiciones

de craqueo, para proporcionar una corriente del producto que comprende olefinas de C_2-C_5.

En ciertas realizaciones de este método, puede ser deseable eliminar los contaminantes que pueden estar presentes en la composición lipídica. Por lo tanto, antes de hacer fluir la composición lipídica a través de una zona de craqueo catalítico fluido, la composición lipídica se somete a un tratamiento previo. El tratamiento previo puede incluir poner en contacto la composición lipídica con una resina de intercambio iónico. La resina de intercambio iónico es una resina de intercambio iónico ácida, tal como Amberlyst™-15 y se puede usar como lecho en un reactor a través del cual fluye la composición lipídica, ya sea hacia arriba o hacia abajo. Otros tratamientos previos pueden incluir lavados con ácidos leves por contacto de la composición lipídica con un ácido, tal como el ácido sulfúrico, acético, nítrico o clorhídrico. El contacto se realiza con una solución diluida del ácido, por lo general a temperatura ambiente y a presión atmosférica.

La composición lipídica, opcionalmente sometida a un tratamiento previo, se hace fluir a una zona de FCC, donde los componentes hidrocarburos se craquean a olefinas. El craqueo catalítico se lleva a cabo por el contacto de la composición lipídica en una zona de reacción con un catalizador compuesto de un material particulado finamente dividido. La reacción es de craqueo catalítico, a diferencia del hidrocraqueo, y se lleva a cabo en ausencia de hidrógeno añadido o consumo de hidrógeno. A medida que avanza la reacción de craqueo, cantidades sustanciales de coque se depositan en el catalizador. El catalizador se regenera a altas temperaturas por la quema del coque del catalizador en una zona de regeneración. El catalizador que contiene coque, que se refiere en la presente como "catalizador coquizado", se transporta continuamente desde la zona de reacción a la zona de la regeneración para regenerarse y reemplazarse por un catalizador regenerado esencialmente libre de coque a partir de la zona de regeneración. La fluidización de las partículas del catalizador mediante varias corrientes gaseosas permite el transporte del catalizador entre la zona de reacción y la zona de regeneración. Los métodos para el craqueo de los hidrocarburos, tal como los de la composición lipídica descritos en la presente, en una corriente fluidizada del catalizador, el transporte del catalizador entre las zonas de reacción y regeneración, y la combustión de coque en el regenerador, son bien conocidos por los expertos en la materia de los procesos de FCC. Las aplicaciones del FCC ejemplares y de catalizadores útiles para craquear la composición lipídica para producir olefinas C_2-C_5 se describen en las patentes de los Estados Unidos núms. 6,538,169 y 7,288,685.

Los catalizadores de FCC apropiados generalmente comprenden al menos dos componentes que pueden o no estar en la misma matriz. En algunas realizaciones, ambos componentes se pueden distribuir a través de todo el recipiente de reacción. El primer componente incluye generalmente cualquiera de los catalizadores conocidos que se utilizan en la materia de craqueo catalítico fluidizado, tal como un catalizador de tipo arcilla activo amorfo y/o un tamiz molecular cristalino de alta actividad. Los catalizadores de tamiz molecular pueden preferirse a los catalizadores amorfos, debido a su selectividad muy mejorada para los productos deseados. En algunas realizaciones preferidas, las zeolitas se pueden usar como tamiz molecular en los procesos de FCC. Preferentemente, el componente del primer catalizador comprende una zeolita de poro grande, tal como una zeolita tipo Y, un material de alúmina activa, un material aglutinante, que comprende ya sea sílice o alúmina y relleno inerte tal como caolín.

En una realización, el craqueo de la composición lipídica de la presente invención, se lleva a cabo en la sección del ascensor o, alternativamente, la sección del elevador, de la zona de FCC. La composición lipídica se introduce en el ascensor mediante una boquilla, lo que resulta en la evaporación rápida de la composición lipídica. Antes de ponerse en contacto con el catalizador, la composición lipídica normalmente tendrá una temperatura de aproximadamente 149ºC a aproximadamente 316ºC (300ºF a 600ºF). El catalizador se hace fluir desde un recipiente de mezclado hasta el ascensor, donde entra en contacto con la composición lipídica durante un tiempo de 2 segundos o menos.

El catalizador mezclado y los vapores de la composición lipídica reaccionada se descargan desde la parte superior del elevador a través de una conexión de salida, y se separan en una corriente de vapor del producto craqueado como olefinas y una colección de partículas del catalizador cubiertas con cantidades sustanciales de coque que se refiere generalmente como "catalizador coquizado". En un esfuerzo por minimizar el tiempo de contacto de la composición lipídica y el catalizador, que puede promover una conversión adicional de los productos deseados a otros productos no deseados, se puede usar cualquier disposición de separadores, tal como una disposición de brazo en espiral, para eliminar el catalizador coquizado de la corriente del producto rápidamente. El separador, por ejemplo, el separador de brazo en espiral, se sitúa en una porción superior de una cámara con una zona de extracción situada en la porción inferior de la cámara. El catalizador separado por el arreglo del brazo en espiral, gotea en la zona de extracción. La corriente de vapor del producto craqueado que comprende los hidrocarburos craqueados que incluyen olefinas ligeras y algún catalizador, salen de la cámara a través de un conducto que está en comunicación con los ciclones. Los ciclones eliminan los restos de partículas del catalizador de la corriente de vapor del producto para reducir las concentraciones de partículas a niveles muy bajos. La corriente de vapor del producto entonces sale de la parte superior del recipiente de separación. El catalizador, separado por los ciclones, se devuelve al recipiente de separación y luego a la zona de extracción. La zona de extracción elimina los hidrocarburos adsorbidos de la superficie del catalizador por el contacto en contracorriente con el vapor.

La baja presión parcial de los hidrocarburos opera a favor de la producción de olefinas ligeras. En consecuencia, la presión del ascensor se establece a aproximadamente de 172 a 241 kPa (25 a 35 psia) con una presión parcial de hidrocarburos de aproximadamente 35 a 172 kPa (5 a 25 psia), con una presión parcial de hidrocarburos preferida de aproximadamente 69 a 138 kPa (10 a 20 psia). Esta presión parcial, relativamente baja para los hidrocarburos, se logra mediante el uso de vapor como un diluyente en la medida en que el diluyente es del 10-55% en peso de la composición lipídica, y preferentemente de aproximadamente el 15% en peso de la composición lipídica. Otros diluyentes, tales como gas seco, se pueden usar para alcanzar presiones parciales de hidrocarburos equivalentes.

La temperatura de la corriente de craqueo en la salida del ascensor será de aproximadamente 510ºC a 621ºC (950ºF a 1150ºF). Sin embargo, las temperaturas de salida del ascensor por encima de 566ºC (1050ºF) producen más gas seco y más olefinas. Mientras que las temperaturas de salida del ascensor por debajo de 566ºC (1050ºF) producen menos etileno y propileno. Por consiguiente, se prefiere para llevar a cabo el proceso de FCC a una temperatura preferida de aproximadamente 566ºC a aproximadamente 630°C, presión preferida de aproximadamente de 138 kPa a aproximadamente 240 kPa (20 a 35 psia). Otra condición para el proceso es la relación del catalizador con respecto a la composición lipídica, la cual puede variar de aproximadamente 5 a aproximadamente 20 y preferentemente de aproximadamente 10 a aproximadamente 15.

En una realización del método para producir un combustible de reactor, la composición lipídica se introduce en la sección del elevador de un reactor de FCC. La temperatura en la sección del elevador estará muy caliente y se encuentra en el intervalo de aproximadamente 700°C (1292ºF) a aproximadamente 760°C (1400ºF), con una relación del catalizador con respecto a la composición lipídica de aproximadamente 100 a aproximadamente 150. Se prevé que la introducción de la composición lipídica en la sección del elevador producirá cantidades considerables de propileno y etileno.

En otra realización del método para producir un combustible de reactor con la composición lipídica o los lípidos producidos como se describe en la presente, la estructura de la composición lipídica o de los lípidos se rompe por un proceso conocido como hidrodesoxigenación (HDO). La HDO significa la eliminación de oxígeno por medio de hidrógeno, es decir, el oxígeno se elimina mientras que se rompe la estructura del material. Los dobles enlaces olefínicos se hidrogenan y todos los compuestos de azufre y nitrógeno se eliminan. La eliminación de azufre se llama hidrodesulfuración (HDS). El tratamiento previo y la pureza de las materias primas (composición lipídica o lípidos) contribuyen a la vida útil del catalizador.

Generalmente en la HDO/etapa de HDO, el hidrógeno se mezcla con el stock de la fuente (composición lipídica o lípidos) y luego la mezcla se pasa a través de un lecho del catalizador, como un flujo de co-corriente, ya sea como una sola fase o una materia prima de la fuente en dos fases. Después de la HDO/etapa de HDS, la fracción del producto se separa y se pasa a un reactor de isomerización separado. Un reactor de isomerización para material de partida biológico se describe en la literatura (FI 100 248) como un reactor de co-corriente.

El proceso para producir un combustible por hidrogenación de una fuente del hidrocarburo, por ejemplo, la composición lipídica o los lípidos de la presente, también se puede realizar pasando la composición lipídica o los lípidos como un flujo de co-corriente con gas de hidrógeno a través de una primera zona de hidrogenación, y posteriormente el efluente del hidrocarburo se hidrogena adicionalmente en una segunda zona de hidrogenación, pasando gas de hidrógeno a la segunda zona de hidrogenación como un flujo en contracorriente con respecto al efluente del hidrocarburo. Las aplicaciones ilustrativas de la HDO y de los catalizadores útiles para craquear la composición lipídica para producir olefinas C_2-C_5 se describen en la patente de los Estados Unidos núm. 7,232,935.

Típicamente, en la etapa de hidrodesoxigenación, la estructura del componente biológico, tal como la composición lipídica o los lípidos de la presente se descompone, el oxígeno, el nitrógeno, los compuestos de fósforo y de azufre y los hidrocarburos ligeros como gas se eliminan, y los enlaces olefínicos se hidrogenan. En la segunda etapa del proceso, es decir, en la etapa llamada de isomerización, se lleva a cabo la isomerización para la ramificación de la cadena del hidrocarburo y se mejora el rendimiento de la parafina a bajas temperaturas.

En la primera atapa, es decir, en la etapa de HDO del proceso de craqueo, el gas de hidrógeno y la composición lipídica o los lípidos de la presente que deben hidrogenarse, se pasan a un sistema de lecho del catalizador de HDO, ya sea como flujos de co-corriente o contracorriente, dicho sistema de lecho del catalizador que comprende uno o más lechos de catalizador(es), preferentemente de 1-3 lechos de catalizador. La etapa de HDO típicamente funciona de una manera en co-corriente. En el caso de un sistema de lecho del catalizador de HDO, el cual comprende dos o más lechos de catalizadores, uno o más de los lechos pueden funcionar usando el principio de flujo en contracorriente. En la etapa de HDO, la presión varía entre 20 y 150 bar, preferentemente entre 50 y 100 bar, y la temperatura varía entre 200 y 500°C, preferentemente en el intervalo de 300-400°C. En la etapa de HDO, se pueden usar los catalizadores de hidrogenación conocidos que contienen metales del grupo VII y/o VIB del sistema periódico. Preferentemente, los catalizadores de hidrogenación son catalizadores soportados de Pd, Pt, Ni, NiMo o un catalizador CoMo, siendo el soporte alúmina y/o sílice. Típicamente se usan los catalizadores $NiMo/Al_2O_3$ y $CoMo/Al_2O_3$.

Antes de la etapa de HDO, la composición lipídica o los lípidos en la presente, opcionalmente, se pueden tratar opcionalmente mediante prehidrogenación en condiciones más suaves, evitándose así reacciones secundarias de los dobles enlaces. Dicha prehidrogenación se lleva a cabo en presencia de un catalizador de prehidrogenación a temperaturas de 50 a 400°C y a presiones de hidrógeno de 1 a 200 bar, preferentemente a una temperatura entre 150 y 250°C y a una presión de hidrógeno entre 10 y 100 bar. El catalizador puede contener metales del grupo VIII y/o VIB del sistema periódico. Preferentemente, el catalizador de prehidrogenación es un catalizador soportado de Pd, Pt, Ni, NiMo o un catalizador CoMo, el soporte es alúmina y/o sílice.

Una corriente gaseosa de la etapa de HDO que contiene hidrógeno se enfría y después los compuestos de monóxido de carbono, dióxido de carbono, nitrógeno, fósforo y azufre, hidrocarburos ligeros gaseosos y otras impurezas se eliminan de la misma. Después de la compresión, el hidrógeno purificado o el hidrógeno reciclado se retornan de vuelta al primer lecho del catalizador y/o entre los lechos de catalizadores para compensar la corriente del gas eliminado. El agua se elimina a partir del líquido condensado. El líquido se pasa al primer lecho del catalizador o entre los lechos de los catalizadores.

Después de la etapa de HDO, el producto se somete a una etapa de isomerización. Es importante para el proceso que las impurezas se eliminen tan completamente como sea posible antes de que los hidrocarburos se pongan en contacto con el catalizador de isomerización. La etapa de isomerización comprende una etapa opcional de separación, en la cual el producto de reacción de la etapa de HDO se puede purificar por separación con vapor de agua o un gas adecuado, tal como hidrocarburos ligeros, nitrógeno o hidrógeno. La etapa opcional de separación se lleva a cabo en forma de contracorriente en una unidad ubicada corriente arriba del catalizador de isomerización, en la cual el gas y el líquido están en contacto uno con el otro, o antes del reactor de isomerización existente en una unidad de separación independiente que utiliza el principio de contracorriente.

Después de la etapa de separación, el gas de hidrógeno y la composición lipídica o los lípidos hidrogenados de la presente, y, opcionalmente, una mezcla de n-parafina, se pasan a una unidad de isomerización reactiva que comprende uno o varios lechos del(de los) catalizador(es). Los lechos de los catalizadores de la etapa de isomerización pueden operar en forma de co-corriente o contracorriente.

Es importante para el proceso que el principio del flujo en contracorriente se aplique en la etapa de isomerización. En la etapa de isomerización, esto se hace llevando a cabo cualquier etapa de separación opcional o la etapa de la reacción de isomerización o ambas en forma de contracorriente. En la etapa isomerización, la presión varía en el intervalo de 20-150 bar, de preferencia en el intervalo de 20-100 bar, y la temperatura está entre 200 y 500°C, preferentemente entre 300 y 400°C. En la etapa de isomerización, los catalizadores de isomerización conocidos en la materia se pueden usar. Los catalizadores de isomerización adecuados contienen el tamiz molecular y/o un metal del grupo VII y/o un vehículo. Preferentemente, el catalizador de isomerización contiene SAPO-11 o SAPO41 o ZSM-22 o ZSM-23 o ferrierita y Pt, Pd o Ni y Al_2O_3 y SiO_2. Los catalizadores de isomerización típicos son, por ejemplo, Pt/SAPO-11/Al_2O_3, Pt/ZSM-22/Al_2O_3, Pt/ZSM-23/Al_2O_3 y Pt/SAPO-11/SiO_2. La etapa de isomerización y la etapa HDO se pueden llevar a cabo en el mismo recipiente de presión o en recipientes de presión separados. La prehidrogenación opcional se puede llevar a cabo en un recipiente de presión separado o en el mismo recipiente de presión que las etapas de HDO y de isomerización.

De esta manera, en una realización, el producto de uno o más reacciones químicas es una mezcla de alcanos que comprende combustible de reactor según la norma ASTM D1655. En algunas realizaciones, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un contenido de azufre que es menor de 10 ppm. En otras realizaciones, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un valor T10 de la curva de destilación de menos de 205°C. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un punto de ebullición final (FBP) menor que 300°C. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un punto de inflamación de al menos 38°C. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene una densidad entre 775K/M^3 y 840K/M^3. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un punto de congelación que está por debajo de -47°C. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene una red de calor de combustión que es al menos de 42.8 MJ/K. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene un contenido de hidrógeno que es al menos 13.4% en masa. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene una estabilidad térmica, según pruebas realizadas por JFTOT gravimétrica cuantitativa a 260°C, que es inferior a 3 mm de Hg. En otra realización, la composición que se ajusta a las especificaciones del combustible de reactor ASTM 1655 tiene una goma existente que está por debajo de 7 mg/dl.

Por lo tanto, la presente invención describe una variedad de métodos en los cuales se lleva a cabo la modificación química de los lípidos de las microalgas para rendir productos útiles en una variedad de aplicaciones industriales y de otro tipo. Ejemplos de los procesos para la modificación de los aceites producidos por los métodos descritos en la presente incluyen, pero sin limitarse a, la hidrólisis del aceite, hidroprocesamiento del aceite, y

esterificación del aceite. La modificación del aceite de las microalgas produce productos oleoquímicos básicos que pueden modificarse adicionalmente en productos oleoquímicos derivados seleccionados para una función deseada. De una manera similar a la descrita anteriormente en relación con el proceso de producción del combustible, estas modificaciones químicas también se pueden realizar en los aceites generados por los cultivos microbianos descritos en la presente. Los ejemplos de productos oleoquímicos básicos incluyen, pero sin limitarse a, jabones, ácidos grasos, ésteres metílicos de ácidos grasos y glicerol. Los ejemplos de productos oleoquímicos derivados incluyen, pero no limitado a, nitrilos grasos, ésteres, ácidos diméricos, compuestos de amonio cuaternario, surfactantes, alcanolamidas grasas, sulfatos de alcoholes grasos, resinas, emulsificantes, alcoholes grasos, olefinas y alcanos superiores.

La hidrólisis de los constituyentes de los ácidos grasos provenientes de los glicerolípidos producidos por los métodos de la invención, produce ácidos grasos libres que pueden derivatizarse para producir otras sustancias químicas útiles. La hidrólisis se produce en presencia de agua y un catalizador que puede ser un ácido o una base. Los ácidos grasos libres liberados se pueden derivar para producir una variedad de productos, como se divulga en los siguientes documentos: patentes de los Estados Unidos núms. 5,304,664 (Ácidos grasos altamente sulfatados); 7,262,158 (Composiciones de limpieza); 7,115,173 (Composiciones suavizantes de tejidos); 6,342,208 (Emulsiones para el tratamiento de la piel); 7,264,886 (Composiciones repelentes de agua); 6,924,333 (Aditivos de pinturas); 6,596,768 (Alimentos para rumiantes enriquecidos con lípidos); y 6,380,410 (Surfactantes para detergentes y limpiadores).

Con respecto a la hidrólisis, en una realización de la invención, un aceite de triglicéridos primero se hidroliza opcionalmente en un medio líquido tal como agua o hidróxido de sodio, con el objetivo de obtener glicerina y jabones. Existen varios métodos de hidrólisis de los triglicéridos adecuados, que incluyen pero no limitan a, saponificación, hidrólisis ácida, hidrólisis alcalina, hidrólisis enzimática (denominado aquí como fraccionamiento), e hidrólisis usando agua comprimida caliente. Un experto en la materia reconocerá que un aceite de triglicéridos no necesita hidrolizarse con el objetivo de producir un producto oleoquímico; más bien el aceite puede convertirse directamente al producto oleoquímico deseado mediante otro proceso conocido. Por ejemplo, el aceite de triglicéridos puede convertirse directamente en un éster metílico de ácidos grasos por medio de esterificación.

En algunas realizaciones, la hidrólisis catalítica del aceite producido por los métodos descritos en la presente, mediante el fraccionamiento del aceite en glicerina y ácidos grasos. Como se mencionó anteriormente, los ácidos grasos pueden procesarse después a través de varias otras modificaciones para obtener los derivados oleoquímicos. Por ejemplo, en una realización, los ácidos grasos pueden sufrir una reacción de aminación para producir compuestos de nitrógeno graso. En otra realización, los ácidos grasos pueden sufrir ozonólisis para producir ácidos mono y dibásicos.

En otras realizaciones la hidrólisis puede ocurrir a través del fraccionamiento de los aceites producidos en la presente para crear productos oleoquímicos. En algunas realizaciones preferidas de la invención, un aceite de triglicéridos se puede fraccionar antes de que se lleven a cabo otros procesos. Un experto en la materia reconocerá que hay muchos métodos adecuados para el fraccionamiento de los triglicéridos, que incluyen pero no limitan al fraccionamiento enzimático y el fraccionamiento por presión.

Generalmente, los métodos de fraccionamiento enzimático del aceite usan enzimas, las lipasas, las cuales actúan como biocatalizadores en una mezcla de agua y aceite. El fraccionamiento enzimático fracciona el aceite o la grasa, respectivamente, en glicerol y ácidos grasos libres. El glicerol puede entonces migrar a la fase del agua, mientras que la fase orgánica se enriquece con los ácidos grasos libres.

Las reacciones de fraccionamiento enzimático generalmente tienen lugar en la delimitación entre de la fase orgánica y la fase la acuosa, donde la enzima está presente solamente en la delimitación de las fases. Los triglicéridos que se reúnen en la delimitación de las fases entonces contribuyen a o participan en la reacción de fraccionamiento. A medida que avanza la reacción, la densidad de ocupación o concentración de los ácidos grasos, todavía unidos químicamente como glicéridos, en comparación con los ácidos grasos libres, disminuye en la delimitación de las fases, de manera que la reacción es más lenta. En ciertas realizaciones, el fraccionamiento enzimático puede ocurrir a temperatura ambiente. Un experto en la materia conocerá las condiciones adecuadas para el fraccionamiento del aceite en los ácidos grasos deseados.

A modo de ejemplo, la velocidad de reacción puede acelerarse mediante el aumento de la superficie de delimitación de la interfase. Una vez que la reacción se completa, los ácidos grasos entonces se separan de la fase orgánica liberada de la enzima, y el residuo que aún contiene ácidos grasos unidos químicamente como glicéridos, se alimenta de nuevo o se recicla y se mezcla con el aceite o la grasa para someterse al fraccionamiento. De esta manera, los glicéridos reciclados se someten a un proceso de fraccionamiento enzimático adicional. En algunas realizaciones, los ácidos grasos libres se extraen a partir de un aceite o grasa parcialmente fraccionada de tal manera. De esa forma, si los ácidos grasos unidos químicamente (triglicéridos) se retornan o se retroalimentan en el proceso de fraccionamiento, el consumo de la enzima se puede reducir drásticamente.

El grado de fraccionamiento se determina como la relación entre el valor de la acidez medido dividido por el valor de la acidez teóricamente posible, lo cual se puede calcular para un aceite o grasa determinados. Preferentemente, el valor de la acidez se mide por medio de titulación, de acuerdo con los métodos estándares comunes. Alternativamente, la densidad de la fase acuosa del glicerol se puede tomar como una medida del grado de fraccionamiento.

En una realización, el proceso de fraccionamiento como se describe en la presente, también es adecuado para el fraccionamiento de los mono-, di- y triglicéridos que se encuentran en la llamada pasta de neutralización a partir de los procesos de refinación alcalina de los aceites producidos. De esta manera, la pasta de neutralización puede convertirse cuantitativamente, sin saponificación previa de los aceites neutros, en ácidos grasos. A tal efecto, los ácidos grasos que están químicamente unidos en los jabones se liberan, preferentemente antes del fraccionamiento, a través de una adición de ácido. En ciertas realizaciones, una solución tampón se usa además del agua y la enzima para el proceso de fraccionamiento.

En una realización, los aceites producidos de acuerdo con los métodos de la invención también pueden someterse a saponificación como un método de hidrólisis. Los aceites animales y vegetales se hacen típicamente de triacilgliceroles (TAG), los cuales son ésteres de ácidos grasos con el alcohol trihídrico, glicerol. En una reacción de hidrólisis alcalina, el glicerol en un TAG se elimina, dejando tres aniones ácidos carboxílicos, que se pueden asociar con los cationes de metales alcalinos tales como el sodio o el potasio, para producir sales de ácidos grasos. En este esquema, los constituyentes de los ácidos carboxílicos se disocian de la fracción del glicerol y se reemplazan con grupos hidroxilo. La cantidad de base (por ejemplo, KOH) que se usa en la reacción está determinada por el grado deseado de saponificación. Si el objetivo es, por ejemplo, producir un producto de jabón que comprenda algunos de los aceites originalmente presentes en la composición de TAG, se introduce en la mezcla de reacción una cantidad de base suficiente para convertir todos los TAG en sales de ácidos grasos. Normalmente, esta reacción se lleva a cabo en una solución acuosa y ocurre lentamente, pero se puede acelerar mediante la adición de calor. La precipitación de las sales de ácidos grasos se puede facilitar mediante la adición de sales, tales como los haluros de metales alcalinos solubles en agua (por ejemplo, NaCl o KCl), a la mezcla de reacción. Preferentemente, la base es un hidróxido de metal alcalino, tal como NaOH o KOH. Alternativamente, otras bases, tales como alcanolaminas, incluyendo por ejemplo la trietanolamina y el aminometilpropanol, se pueden usar en el esquema de reacción. En algunos casos, estas alternativas pueden ser preferibles para producir un producto jabonoso claro.

En algunos métodos, la primera etapa de modificación química puede ser el hidroprocesamiento para saturar los dobles enlaces, seguido por la desoxigenación a temperatura elevada en presencia de hidrógeno y un catalizador. En otros métodos, la hidrogenación y la desoxigenación pueden ocurrir en la misma reacción. En aún otros métodos, la desoxigenación ocurre antes de la hidrogenación. La isomerización se puede entonces llevar a cabo opcionalmente, también en presencia de hidrógeno y un catalizador. Finalmente, los gases y los componentes de nafta se pueden eliminar si se desea. Por ejemplo, ver las patentes de los Estados Unidos núms. 55,475,160 (Hidrogenación de triglicéridos); 5,091,116 (Desoxigenación, hidrogenación y eliminación de gas); 6,391,815 (Hidrogenación); y 5,888,947 (Isomerización).

En algunas realizaciones de la invención, los aceites triglicéridos se desoxigenan parcial o totalmente. Las reacciones de desoxigenación forman productos deseados, que incluyen pero sin limitarse a, ácidos grasos, alcoholes grasos, polioles, cetonas y aldehídos. En general, sin estar limitados por ninguna teoría en particular, las reacciones de desoxigenación involucran una combinación de varias vías de reacción diferentes, que incluyen pero sin limitarse a: hidrogenólisis, hidrogenación, hidrogenación-hidrogenólisis consecutiva, hidrogenólisis-hidrogenación consecutiva, y reacciones de hidrogenación-hidrogenólisis combinadas, dando como resultado, al menos, la eliminación parcial del oxígeno de los ácidos grasos o de los ésteres de ácidos grasos para producir productos de reacción, tales como alcoholes grasos, que se pueden convertir fácilmente a los productos químicos deseados mediante el procesamiento posterior. Por ejemplo, en una realización, un alcohol graso se puede convertir a olefinas mediante la reacción de FCC o a alcanos superiores a través de una reacción de condensación.

Una de tales modificaciones químicas es la hidrogenación, la cual es la adición de hidrógeno a los dobles enlaces en los ácidos grasos constituyentes de los glicerolípidos o de los ácidos grasos libres. El proceso de hidrogenación permite la transformación de los aceites líquidos en grasas semisólidas o sólidas, las cuales pueden ser más adecuadas para aplicaciones específicas.

La hidrogenación del aceite producido por los métodos descritos en la presente, se puede llevar a cabo en conjunto con uno o más de los métodos y/o materiales proporcionados en la presente, como se indica en los siguientes documentos: patentes de los Estados Unidos núms. 7,288,278 (Aditivos de alimentos o medicamentos); 5,346,724 (Productos para lubricación); 5,475,160 (Alcoholes grasos); 5,091,116 (Aceites comestibles); 6,808,737 (Grasas estructurales para margarina para untar); 5,298,637 (sustitutos de grasas reducidos en calorías); 6,391,815 (Catalizador de hidrogenación y adsorbente de azufre); 5,233,099 y 5,233,100 (Alcoholes grasos); 4,584,139 (Catalizadores de hidrogenación); 6,057,375 (Agentes supresores de espuma); y 7,118,773 (Emulsión comestible para untar).

Un experto en la materia reconocerá que se pueden usar varios procesos para hidrogenar los carbohidratos. Un método conveniente incluye poner en contacto los carbohidratos con hidrógeno o hidrógeno mezclado con un gas adecuado y un catalizador, bajo condiciones suficientes en un reactor de hidrogenación para formar un producto hidrogenado. El catalizador de la hidrogenación generalmente puede incluir Cu, Re, Ni, Fe, Co, Ru, Pd, Rh, Pt, Os, Ir, y aleaciones o cualquier combinación de los mismos, ya sean solos o con promotores, tales como W, Mo, Au, Ag, Cr, Zn, Mn, Sn, B, P, Bi, y aleaciones o cualquier combinación de los mismos. Otros materiales catalizadores de la hidrogenación efectivos incluyen tanto el níquel soportado o el rutenio modificado con renio. En una realización, el catalizador de la hidrogenación también incluye cualquiera de los soportes, dependiendo de la funcionalidad deseada del catalizador. Los catalizadores de la hidrogenación se pueden preparar por los métodos conocidos por los expertos en la materia.

En algunas realizaciones, el catalizador de la hidrogenación incluye un catalizador de metal del grupo VIII soportado y un material de esponja de metal (por ejemplo, un catalizador de esponja de níquel). El níquel Raney proporciona un ejemplo de un catalizador de esponja de níquel activado, adecuado para su uso en esta invención. En otra realización, la reacción de hidrogenación en la invención se lleva a cabo usando un catalizador que comprende un catalizador de níquel-renio o un catalizador de níquel-tungsteno modificado. Un ejemplo de un catalizador adecuado para la reacción de hidrogenación de la invención es un catalizador de níquel-renio soportado en carbono.

En una realización, un catalizador de níquel Raney adecuado se puede preparar mediante el tratamiento de una aleación de cantidades aproximadamente iguales en peso de níquel y aluminio con una solución alcalina acuosa, por ejemplo, que contiene aproximadamente 25% en peso de hidróxido de sodio. El aluminio se disuelve selectivamente por la solución alcalina acuosa, dando como resultado un material en forma de esponja que comprende mayormente níquel con pequeñas cantidades de aluminio. La aleación inicial incluye metales promotores (a saber, molibdeno y cromo) en cantidades tales que aproximadamente del 1 al 2% del peso se mantiene en el catalizador en la forma de esponja de níquel. En otra realización, el catalizador de la hidrogenación se prepara usando una solución de nitrosilnitrato de rutenio (III), cloruro de rutenio (III) en agua para impregnar un material de soporte adecuado. La solución entonces se seca para formar un sólido que tiene un contenido de agua de menos de aproximadamente el 1% en peso. El sólido puede entonces reducirse a la presión atmosférica en una corriente de hidrógeno a 300°C (sin calcinarse) o 400°C (calcinado) en un horno rotatorio de bolas durante 4 horas. Después de enfriarse y dar lugar al catalizador inerte con nitrógeno, 5% en volumen del oxígeno en el nitrógeno se pasa sobre el catalizador durante 2 horas.

En ciertas realizaciones, el catalizador descrito incluye un soporte de catalizador. El soporte del catalizador estabiliza y soporta el catalizador. El tipo de soporte del catalizador usado depende del catalizador escogido y de las condiciones de reacción. Los soportes adecuados para la invención incluyen, pero sin limitarse a, carbono, sílice, sílice-alúmina, zirconio, titania, ceria, vanadio, nitruro, nitruro de boro, heteropoliácidos, hidroxiapatito, óxido de zinc, cromía, zeolita, nanotubos de carbono, fullereno de carbono y cualquier combinación de éstos.

Los catalizadores utilizados en esta invención se pueden preparar por los métodos convencionales conocidos por los expertos en la materia. Los métodos adecuados pueden incluir, pero sin limitarse a, humedecimiento incipiente, impregnación evaporativa, deposición química de vapor, lavado de recubrimiento, técnicas de pulverización catódica, y similares.

Las condiciones para que llevar a cabo la reacción de hidrogenación pueden variar en función del tipo de materia prima y los productos deseados. Un experto en la materia, con el beneficio de esta divulgación, reconocerá las condiciones de reacción adecuadas. En general, la reacción de hidrogenación se lleva a cabo a temperaturas de 80°C a 250°C, y preferentemente de 90°C a 200°C, y con más preferencia de 100°C a 150°C. En algunas realizaciones, la reacción de hidrogenación se lleva a cabo a presiones de 500 kPa a 14 000 kPa.

El hidrógeno usado en la reacción de hidrogenólisis de la presente invención puede incluir hidrógeno externo, hidrógeno reciclado, hidrógeno generado in situ, y cualquier combinación de los mismos. Tal como se usa en la presente, el término "hidrógeno externo" se refiere al hidrógeno que no se origina de la reacción de la biomasa en sí, sino que se añade al sistema de otra fuente.

En algunas realizaciones de la invención, es deseable convertir los carbohidratos de partida en una molécula más pequeña que será más fácil de convertirse en los hidrocarburos superiores deseados. Un método conveniente para esta conversión es a través de una reacción de hidrogenólisis. Varios procesos se conocen para realizar la hidrogenólisis de los carbohidratos. Un método apropiado incluye poner en contacto un hidrato de carbono con hidrógeno o hidrógeno mezclado con un gas adecuado y un catalizador de hidrogenólisishidrogenólisis en un reactor de hidrogenólisis, bajo condiciones suficientes para formar un producto de reacción que comprende moléculas más pequeñas o polioles. Tal como se usa en la presente, el término "moléculas más pequeñas o polioles" incluye cualquier molécula que tiene un peso molecular más pequeño, que puede incluir un número menor de átomos de carbono o de átomos de oxígeno que los carbohidratos de partida. En una realización, los productos de reacción son moléculas más pequeñas que incluyen polioles y alcoholes. Una persona con conocimientos

comunes en la materia sería capaz de seleccionar el método apropiado para llevar a cabo la reacción de hidrogenólisis.

En algunas realizaciones, un alcohol de azúcar o azúcar de 5 ó 6 carbonos se puede convertir a propilenglicol, etilenglicol y glicerol con un catalizador de la hidrogenólisis. El catalizador de la hidrogenólisis puede incluir Cr, Mo, W, Re, Mn, Cu, Cd, Fe, Co, Ni, Pt, Pd, Rh, Ru, Ir, Os, y aleaciones o cualquier combinación de los mismos, ya sean solos o con promotores de tales como Au, Ag, Cr, Zn, Mn, Sn, Bi, B, O y aleaciones o cualquier combinación de los mismos. El catalizador de la hidrogenólisis puede incluir también un catalizador de piropolímero carbonoso que contiene metales de transición (por ejemplo, cromo, molibdeno, tungsteno, renio, manganeso, cobre, cadmio) o metales del grupo VIII (por ejemplo, hierro, cobalto, níquel, platino, paladio, rodio, rutenio, iridio y osmio). En ciertas realizaciones, el catalizador de la hidrogenólisis puede incluir cualquiera de los metales anteriores, combinado con un óxido de metal alcalinotérreo o adherido a un soporte catalíticamente activo. En ciertas realizaciones, el catalizador en la reacción de hidrogenólisis puede incluir un soporte de catalizador como el descrito anteriormente para la reacción de hidrogenación.

Las condiciones para que llevar a cabo la reacción de hidrogenólisis pueden variar en función del tipo de materia prima y los productos deseados. Un experto en la materia, con beneficio de esta divulgación, reconocerá las condiciones apropiadas a usar para llevar a cabo la reacción. En general, la reacción de hidrogenólisis se lleva a cabo a temperaturas de 110°C a 300°C, y preferentemente de 170°C a 220°C, y con más preferencia de 200°C a 225ºC. En algunas realizaciones, la reacción de hidrogenólisis se lleva a cabo en condiciones básicas, preferentemente a un pH de 8 a 13, y aún con más preferencia a pH de 10 a 12. En algunas realizaciones, la reacción de hidrogenólisis se lleva a cabo a presiones en un intervalo entre 60 kPa y 16 500 kPa, y preferentemente en un intervalo entre 1700 KPa y 14000 KPa, y aún con más preferencia entre 4800 KPa y 11000 KPa.

El hidrógeno usado en la reacción de hidrogenólisis de la presente invención puede incluir hidrógeno externo, hidrógeno reciclado, hidrógeno generado in situ, y cualquier combinación de los mismos.

En algunas realizaciones, los productos de reacción mencionados anteriormente pueden convertirse en hidrocarburos superiores a través de una reacción de condensación en un reactor de condensación (que se muestra esquemáticamente como el reactor de condensación 110 en la Figura 1). En tales realizaciones, la condensación de los productos de reacción se produce en presencia de un catalizador capaz de formar hidrocarburos superiores. Aunque sin pretender limitarse por la teoría, se cree que la producción de hidrocarburos superiores ocurre a través de una reacción de adición por pasos, la cual incluye la formación de enlaces carbono-carbono o enlaces carbono-oxígeno. Los productos de reacción resultantes incluyen cualquier número de compuestos que contienen estos residuos, tal como se describe con más detalle a continuación.

En ciertas realizaciones, los catalizadores de condensación adecuados incluyen un catalizador ácido, un catalizador básico, o un catalizador ácido/básico. Como se usa en la presente, el término “catalizador ácido/básico” se refiere a un catalizador que tiene una funcionalidad ácido y una base. En algunas realizaciones, el catalizador de condensación puede incluir, sin limitarse a, zeolitas, carburos, nitruros, zirconio, alúmina, sílice, aluminosilicatos, fosfatos, óxidos de titanio, óxidos de zinc, óxidos de vanadio, óxidos de lantano, óxidos de itrio, óxidos de escandio, óxidos de magnesio, óxidos de cerio, óxidos báricos, óxidos de calcio, hidróxidos, heteropoliácidos, ácidos inorgánicos, resinas ácidas modificadas, resinas básicas modificadas, y cualquier combinación de éstos. En algunas realizaciones, el catalizador de condensación puede incluir, además, un modificador. Los modificadores adecuados incluyen La, Y, Sc, P, B, Bi, Li, Na, K, Rb, Cs, Mg, Ca, Sr, Ba, y cualquier combinación de éstos. En algunas realizaciones, el catalizador de condensación puede incluir, además, un metal. Los metales adecuados incluyen Cu, Ag, Au, Pt, Ni, Fe, Co, Ru, Zn, Cd, Ga, In, Rh, Pd, Ir, Re, Mn, Cr, Mo, W, Sn, Os, aleaciones, y cualquier combinación de éstos.

En ciertas realizaciones, el catalizador descrito en la reacción de condensación puede incluir un soporte de catalizador como el descrito anteriormente para la reacción de hidrogenación. En ciertas realizaciones, el catalizador de la condensación es auto-soportado. Tal como se usa en la presente, el término "auto-soportado" significa que el catalizador no necesita de otro material que le sirva de soporte. En otras realizaciones, el catalizador de la condensación se usa conjuntamente con un soporte separado apropiado para suspender el catalizador. En una realización, el soporte del catalizador de la condensación es sílice.

Las condiciones bajo las cuales la reacción de condensación se produce pueden variar en función del tipo de materia prima y los productos deseados. Un experto en la materia, con beneficio de esta divulgación, reconocerá las condiciones apropiadas a usar para llevar a cabo la reacción. En algunas realizaciones, la reacción de condensación se lleva a cabo a una temperatura en la cual la termodinámica de la reacción propuesta sea favorable. La temperatura para la reacción de condensación puede variar dependiendo del poliol o alcohol de partida específico. En algunas realizaciones, la temperatura de la reacción de condensación se encuentra en un intervalo de 80°C a 500°C, y preferentemente de 125ºC a 450°C, y con la máxima preferencia de 125ºC a 250°C. En algunas realizaciones, la reacción de condensación se lleva a cabo a presiones en un intervalo de 0 kPa a 9000 kPa, y preferentemente en un intervalo entre 0 kPa y 7000 KPa, y aún con la máxima preferencia entre 0 kPa y 5000 KPa.

Los alcanos superiores formados por la invención incluyen, pero sin limitarse a, alcanos de cadena recta o ramificada que tienen de 4 a 30 átomos de carbono, alquenos de cadena recta o ramificada que tienen de 4 a 30 átomos de carbono, cicloalcanos que tienen de 5 a 30 átomos de carbono, cicloalquenos que tienen de 5 a 30 átomos de carbono, arilos, arilos fusionados, alcoholes, y cetonas. Los alcanos adecuados incluyen, pero sin limitarse a, butano, pentano, penteno, 2-metilbutano, hexano, hexeno, 2-metilpentano, 3-metilpentano, 2,2-dimetilbutano, 2,3-dimetilbutano, heptano, hepteno, octano, octeno, 2,2,4-trimetilpentano, 2,3-dimetilo hexano, 2,3,4-trimetilpentano, 2,3-dimetilpentano, nonano, noneno, decano, deceno, undecano, undeceno, dodecano, dodeceno, tridecano, trideceno, tetradecano, tetradeceno, pentadecano, pentadeceno, nonildecano, nonildeceno, eicosano, eicoseno, uneicosano, uneicoseno, doeicosano, doeicoseno, trieicosano, trieicoseno, tetraeicosano, tetraeicoseno, e isómeros de éstos. Algunos de estos productos pueden ser adecuados para usar como combustibles.

En algunas realizaciones, los cicloalcanos y los cicloalquenos son no sustituidos. En otras realizaciones, los cicloalcano y los cicloalquenos son mono-sustituidos. En aún otras realizaciones, los cicloalcano y los cicloalquenos son multi-sustituidos. En las realizaciones que comprenden los cicloalcanos y cicloalquenos sustituidos, el grupo sustituido incluye, sin limitarse a, un alquil de cadena recta o ramificada que tiene 1 a 12 átomos de carbono, un alquileno de cadena recta o ramificada que tiene 1 a 12 átomos de carbono, un fenilo, y cualquier combinación de éstos. Los cicloalcanos y cicloalquenos incluyen, pero sin limitarse a, ciclopentano, ciclopenteno, cicliclohexano, ciclohexeno, metil-ciclopentano, metil-ciclopenteno, etil-ciclopentano, etil-ciclopenteno, etil-cicliclohexano, etil-ciclohexeno, isómeros y cualquier combinación de éstos.

En algunas realizaciones, los arilos formados son no sustituidos. En otra realización, los arilos formados son mono-sustituidos. En las realizaciones que comprenden los arilos sustituidos, el grupo sustituido incluye, sin limitarse a, un alquil de cadena recta o ramificada que tiene 1 a 12 átomos de carbono, un alquileno de cadena recta o ramificada que tiene 1 a 12 átomos de carbono, un fenilo, y cualquier combinación de éstos. Los arilos adecuados para la invención incluyen, pero sin limitarse a, benceno, tolueno, xileno, etil benceno, para xileno, meta xileno, y cualquier combinación de éstos.

Los alcoholes producidos en la invención tienen de 4 a 30 átomos de carbono. En algunas realizaciones, los alcoholes son cíclicos. En otras realizaciones, los alcoholes son ramificados. En otra realización, los alcoholes son de cadena recta. Los alcoholes adecuados para la invención incluyen, pero sin limitarse a, butanol, pentanol, hexanol, heptanol, octanol, nonanol, decanol, undecanol, dodecanol, tridecanol, tetradecanol, pentadecanol, hexadecanol, heptildecanol, octildecanol, nonildecanol, eicosanol, uneicosanol, doeicosanol, trieicosanol, tetraeicosanol, e isómeros de éstos.

Las cetonas producidas en la invención tienen de 4 a 30 átomos de carbono. En una realización, las cetonas son cíclicas. En otra realización, las cetonas son ramificadas. En otra realización, las cetonas son de cadena recta. Las cetonas adecuadas para la invención incluyen, pero sin limitarse a, butanona, pentanona, hexanona, heptanona, octanona, nonanona, decanona, undecanona, dodecanona, tridecanona, tetradecanona, pentadecanona, hexadecanona, heptildecanona, octildecanona, nonildecanona, eicosanona, uneicosanona, doeicosanona, trieicosanona, tetraeicosanona, e isómero de éstos.

Otra de tales modificaciones químicas es la interesterificación. Los glicerolípidos producidos naturalmente no tienen una distribución uniforme de los constituyentes de los ácidos grasos. En el contexto de los aceites, la interesterificación se refiere al intercambio de radicales acilo entre dos ésteres de diferentes glicerolípidos. El proceso de interesterificación proporciona un mecanismo mediante el cual los constituyentes de los ácidos grasos de una mezcla de glicerolípidos pueden reordenarse para modificar el patrón de distribución. La interesterificación es un proceso químico conocido, y generalmente comprende calentar (a aproximadamente 200°C) una mezcla de aceites por un período de tiempo (por ejemplo, 30 minutos) en presencia de un catalizador, tal como un metal alcalino o metal alcalino alquilado (por ejemplo, metóxido de sodio). Este proceso se puede usar para seleccionar al azar el patrón de distribución de los constituyentes de los ácidos grasos de una mezcla de aceite, o puede dirigirse para producir un patrón de distribución deseado. Este método de modificación química de los lípidos se puede llevar a cabo en los materiales proporcionados en la presente, tal como la biomasa microbiana, con un porcentaje de peso seco de las células como lípidos de al menos el 20%.

La interesterificación dirigida, en la cual se busca un patrón de distribución específica de los ácidos grasos, se puede realizar mediante el mantenimiento de la mezcla de aceites a una temperatura por debajo del punto de fusión de algunos TAG que pudieran obtenerse. Esto resulta en la cristalización selectiva de estos TAG, lo cual los elimina eficientemente de la mezcla de reacción donde cristalizan. El proceso puede continuar hasta que la mayoría de los ácidos grasos en el aceite hayan precipitado, por ejemplo. Un proceso de interesterificación dirigida puede usarse, por ejemplo, para producir un producto con un contenido calórico inferior a través de la sustitución de los ácidos grasos de cadena larga con sus homólogos de cadena más corta. La interesterificación dirigida también se puede usar para producir un producto con una mezcla de grasas que pueden proporcionar las características de fusión deseadas y los rasgos estructurales buscados en los aditivos o productos alimentarios (por ejemplo, la margarina) sin recurrir a la hidrogenación, la cual puede producir isómeros trans no deseados.

La interesterificación de aceites producidos por los métodos descritos en la presente se puede realizar conjuntamente con uno o varios de los métodos y/o materiales, o para producir productos, como se divulga en los siguientes documentos: patente de los Estados Unidos núms. 6,080,853 (Sustitutos de grasas no digeribles); 4,288,378 (Estabilizador de mantequilla de maní); 5,391,383 (Atomizador de aceite comestible); 6,022,577 (Grasas comestibles para productos alimenticios); 5,434,278 (Grasas comestibles para productos alimenticios); 5,268,192 (Productos de nueces bajos en calorías); 5,258,197 (Composiciones comestibles reducidas en calorías); 4,335,156 (Producto de grasa comestible); 7,288,278 (Aditivos de alimentos o medicamentos); 7,115,760 (Proceso de fraccionamiento); 6,808,737 (Grasas estructurales); 5,888,947 (Lubricantes para motor); 5,686,131 (Mezclas de aceites comestibles); y 4,603,188 (Composiciones de uretanos curables).

En una realización de conformidad con la invención, la transesterificación del aceite, según lo descrito anteriormente, se sigue por la reacción del producto transesterificado con poliol, como se divulgó en la patente de los Estados Unidos núm. 6,465,642, para producir el poliol poliéster de ácido graso. Este proceso de esterificación y separación puede comprender los siguientes pasos d: reaccionar un éster de alquilo inferior con el poliol en presencia de jabón; eliminar los restos de jabón de la mezcla de productos; lavar con agua y secar la mezcla de productos para eliminar las impurezas; blanquear la mezcla de productos para su refinación; separar al menos una porción del éster de alquilo inferior sin reaccionar del poliol poliéster de ácido graso en la mezcla de productos; y reciclar el éster de alquilo inferior sin reaccionar separado.

La transesterificación también se puede llevar a cabo en la biomasa microbiana con ésteres de ácidos grasos de cadena corta, como se describe en la patente de los Estados Unidos núm. 6,278,006. En general, la transesterificación se puede realizar mediante la adición de un éster de ácido graso de cadena corta a un aceite en presencia de un catalizador adecuado, y calentando la mezcla. En algunas realizaciones, el aceite comprende de aproximadamente 5% a aproximadamente 90% de la mezcla de reacción en peso. En algunas realizaciones, los ésteres de ácidos grasos de cadena corta pueden constituir aproximadamente 10% a 50% de la mezcla de reacción en peso. Los ejemplos no limitantes de catalizadores incluyen catalizadores básicos, metóxido de sodio, catalizadores ácidos, incluyendo ácidos inorgánicos tales como ácido sulfúrico y arcillas acidificadas, ácidos orgánicos tales como ácido metano sulfónico, ácido bencenosulfónico y ácido toluenosulfónico, y resinas ácidas, tales como Amberlyst 15. Los metales tales como sodio y magnesio, y los hidruros de metales son también catalizadores útiles.

Otra modificación química es la hidroxilación, la cual involucra la adición de agua a un doble enlace lo que resulta en saturación e incorporación de un radical hidroxilo. El proceso de hidroxilación proporciona un mecanismo para la conversión de uno o más constituyentes de los ácidos grasos de un glicerolípido a un ácido graso hidroxilo. La hidroxilación se puede llevar a cabo, por ejemplo, a través del método descrito en la patente de los Estados Unidos núm. 5,576,027. Los ácidos grasos hidroxilados, incluyendo el aceite de ricino y sus derivados, son útiles como componentes en diversas aplicaciones industriales, incluyendo los aditivos alimentarios, agentes surfactantes, agentes humectantes, pigmentos, agentes antiespumantes, aditivos impermeabilizantes, agentes plastificantes, emulsificantes cosméticos y/o agentes desodorantes, así como en la electrónica, productos farmacéuticos, pinturas, tintas, adhesivos, y lubricantes. Un ejemplo de cómo se puede realizar la hidroxilación de un glicérido es como sigue: la grasa se puede calentar, preferentemente a aproximadamente 30-50°C combinada con heptano y se mantiene a la temperatura durante treinta minutos o más; después se puede agregar ácido acético a la mezcla seguido de una solución acuosa de ácido sulfúrico seguido por una solución acuosa de peróxido de hidrógeno la cual se añade en pequeños incrementos a la mezcla durante una hora; después del peróxido de hidrógeno acuoso, la temperatura puede incrementarse hasta al menos aproximadamente 60°C y se agita al menos durante seis horas; después de la agitación la mezcla se deja reposar y una capa acuosa inferior formada por la reacción puede eliminarse mientras que la capa superior de heptano formada por la reacción se puede lavar con agua caliente a una temperatura de aproximadamente 60°C, la capa de heptano lavada se puede neutralizar con una solución acuosa de hidróxido de potasio a un pH de aproximadamente 5 a 7 y después se elimina por destilación al vacío; el producto de rreacción puede secarse al vacío a 100°C y el producto seco se desodoriza con vapor en condiciones de vacío y se filtra a aproximadamente 50° a 60°C usando tierra de diatomáceas.

La hidroxilación de los aceites microbianos producidos por los métodos descritos en la presente se puede realizar conjuntamente con uno o varios de los métodos y/o materiales, o para producir los productos como se divulga en los siguientes documentos: patente de los Estados Unidos núm. 6,590,113 (Recubrimientos y tinta basados en aceite); 4,049,724 (Proceso de hidroxilación); 6,113,971 (Mantequilla de aceite de oliva); 4,992,189 (Lubricantes y aditivos de lubricantes); 5,576,027 (Leche hidroxilada); y 6,869,597 (Cosméticos).

Los glicerolípidos hidroxilados se pueden convertir en estolides. Los estolides consisten de un glicerolípido en el cual un constituyente de ácido graso hidroxilado se ha esterificado a otra molécula de ácido graso. La conversión de glicerolípidos hidroxilados a estolides se puede llevar a cabo mediante el calentamiento de una mezcla de glicerolípidos y ácidos grasos, y poniendo en contacto la mezcla con un ácido mineral, según lo descrito por Isbell y otros, *JAOCS* 71(2): 169-174 (1994). Los estólidos son útiles en una variedad de aplicaciones, incluyendo sin limitación a las descritas en los siguientes documentos: patentes de los Estados Unidos núms. 7,196,124 (Materiales elastoméricos y recubrimientos de pisos); 5,458,795 (Aceites engrosados para aplicaciones de

altas temperaturas); 5,451,332 (Fluidos para aplicaciones industriales); 5,427,704 (Aditivos de combustibles); y 5,380,894 (Lubricantes, grasas, plastificantes, y tintas para impresión).

Otras reacciones químicas que se pueden realizar en los aceites microbianos incluyen reaccionar los triacilgliceroles con un agente de ciclopropanación para aumentar la fluidez y/o estabilidad oxidativa, como se reporta en la patente de los Estados Unidos 6,051,539; fabricar ceras a partir de triacilgliceroles, como se reporta en la patente de los Estados Unidos 6,770,104; y epoxidación de triacilgliceroles, como se reporta en "The effect of fatty acid composition on the acrylation kinetics of epoxidized triacylglycerols", Journal of the American Oil Chemists' Society, 79:1, 59-63, (2001) y en Free Radical Biology and Medicine, 37:1, 104-114 (2004).

La generación de biomasa microbiana oleaginosas para combustibles y productos químicos como se describe anteriormente, resulta en la producción de harina de biomasa deslipidada. La harina deslipidada es un subproducto de la preparación de aceite de algas y es útil para la alimentación animal para animales de granja, por ejemplo, rumiantes, aves de corral, cerdos y para la acuicultura. La harina resultante, aunque tiene un contenido reducido de aceites, todavía contiene proteínas de alta calidad, carbohidratos, fibra, cenizas, restos de aceite y otros nutrientes adecuados para una alimentación animal. Debido a que las células están predominantemente lisadas por el proceso de separación de los aceites, la harina de deslipidada es fácilmente digerible por estos animales. La harina deslipidada opcionalmente se puede combinar con otros ingredientes, tales como granos, en un alimento animal. Debido a que la harina deslipidada tiene una consistencia de polvo, se puede comprimir en gránulos usando un extrusor o extensor u otro tipo de máquina, las cuales se encunetran comercialmente disponibles.

La invención, que se describió en detalle anteriormente, se ejemplifica en los siguientes ejemplos, que se ofrecen para ilustrar pero no para limitar, la invención reivindicada.

VII. EJEMPLOS

Ejemplo 1: Métodos para el cultivo de *Prototheca*

Las cepas de Prototheca se cultivaron para lograr un alto porcentaje de aceite por peso seco de células. Las células criopreservadas se descongelaron a temperatura ambiente y 500 µl de células se añadieron a 4.5 ml de medio (4.2 g/l de K_2HPO_4, 3.1 g/l de NaH_2PO_4, 0.24 g/l de $MgSO_4 \cdot 7H_2O$, 0.25 g/l de ácido cítrico monohidrato, 0.025 g/l de $CaCl_2$ $2H_2O$, 2 g/l de extracto de levadura) más 2% de glucosa y crecieron durante 7 días a 28°C con agitación (200 rpm) en una placa de 6 pocillos. El peso seco de las células se determinó por centrifugación de 1 ml de cultivo a 14000 rpm durante 5 min en un tubo Eppendorf pre-pesado. El sobrenadante de cultivo se desechó y el sedimento celular resultante se lavó con 1 ml de agua desionizada. El cultivo se centrifugó de nuevo, se desechó el sobrenadante, y el sedimento de células se colocó a -80°C hasta que se congeló. Las muestras se liofilizaron después por 24 horas y se calculó el peso seco de las células. Para la determinación de lípidos totales en los cultivos, se retiraron 3 ml de cultivo y se sometieron a análisis mediante el uso de un sistema de Ankom (Ankom Inc., Macedon, NY), según el protocolo del fabricante. Las muestras se sometieron a extracción por solventes con un extractor Amkom XT10 según el protocolo del fabricante. Los lípidos totales se determinaron como la diferencia de masa entre las muestras secas hidrolizadas con ácido y las muestras secas extraídas con disolvente. Las mediciones del porcentaje de aceite del peso seco de las células se muestran en la Tabla 8.

Tabla 8. Porcentaje de aceite por peso seco de las células

Especie	Cepa	% Aceite
Prototheca stagnora	UTEX 327	13.14
Prototheca moriformis	UTEX 1441	18.02
Prototheca moriformis	UTEX 1435	27.17

Se determinó el genotipo de las muestras de microalgas de las cepas que se enumeran en la Tabla 22. El ADN genómico se aisló a partir de la biomasa de algas de la siguiente manera. Las células de cultivos líquidos (aproximadamente 200 mg) se centrifugaron 5 minutos a 14,000 x g. Las células se resuspendieron después en agua destilada estéril, se centrifugaron 5 minutos a 14,000 xg y se desechó el sobrenadante. Una sola perla de vidrio de ~ 2 mm de diámetro se añadió a la biomasa y los tubos se colocaron a -80 °C durante al menos 15 minutos. Las muestras se retiraron y se añadieron 150 µl de tampón de molturación (Sarkosil 1%, sacarosa 0.25 M, NaCl 50 mM, EDTA 20 mM, Tris-HCl 100 mM, pH 8.0, ARNasa A 0.5 µg/µl). Los sedimentos se resuspendieron con una breve agitación, seguido por la adición de 40 µl de NaCl 5M. Las muestras se agitaron brevemente, seguido por la adición de 66 µl de CTAB 5% (bromuro cetil trimetilamonio) y una agitación breve final. Las muestras se incubaron a continuación a 65°C durante 10 minutos después de lo cual se centrifugaron a 14,000 x g durante 10 minutos. El sobrenadante se transfirió a un tubo nuevo y se extrajo una vez con 300 µl de fenol:cloroformo:alcohol isoamílico a 12:12:1, seguido por la centrifugación durante 5 minutos a 14,000 x g. La fase acuosa resultante se transfirió a un tubo nuevo que contenía 0.7 vol de isopropanol (~ 190 µl), se mezcló por inversión y se incubó a temperatura ambiente durante 30 minutos o toda la noche a 4 °C. El ADN se recuperó a través de la centrifugación a 14,000 x g durante 10 minutos. El precipitado resultante se lavó dos veces con 70% etanol, seguido de un lavado final con

100% etanol. Los sedimentos se secaron al aire durante 20-30 minutos a temperatura ambiente, seguido de la resuspensión en 50 µl de TrisCl 10 mM, EDTA 1 mM (pH 8.0).

Cinco µl del ADN total de algas, que se preparó como se describió anteriormente, se diluyeron 1:50 en Tris 10 mM, pH 8.0. Las reacciones de PCR, de volumen final de 20 µl, se establecieron de la siguiente manera. Diez µl de la mezcla maestra iProof HF 2 x (BIO-RAD) se añadieron a 0.4 µl del cebador SZ02613 (5'-TGTTGAAGAATGAGCCGGCGAC-3' '(SEQ ID NO: 9) a una concentración patrón de 10 mM). Esta secuencia del cebador corre de la posición 567-588 en el núm. de acceso L43357 del Gen Bank y está altamente conservada en los genomas de plantas superiores y plastidios de algas. Esto se siguió con la adición de 0.4 µl del cebador SZ02615 (5'-CAGTGAGCTATTACGCACTC-3' (SEQ ID NO: 10) a la concentración patrón de 10 mM). Esta secuencia del cebador es complementaria a la posición de 1112-1093 en el núm. de acceso L43357 del Gen Bank y es muy conservada en los genomas de plantas superiores y plastidios de algas. A continuación, se añadieron 5 µl del ADN total diluido y 3.2 µl dH_2O. Las reacciones de PCR se llevaron a cabo de la siguiente manera: 98°C, 45", 98°C, 8", 53 °C , 12", 72°C, 20" por 35 ciclos seguidos de 72°C durante 1 min y se mantuvo a 25°C. Para la purificación de los productos de la PCR, se añadieron a cada reacción 20 µl de Tris 10 mM, pH 8.0, seguido por la extracción con 40 µl de fenol:cloroformo:alcohol isoamílico 12:12: 1, la agitación y centrifugación a 14,000 x g durante 5 minutos. Las reacciones de PCR se aplicaron a columnas S-400 (GE Healthcare) y se centrifugaron durante 2 minutos a 3,000 x g. Los productos de la PCR purificados posteriormente se clonaron en TOPO en el PCR8/GW/TOPO y los clones positivos se seleccionaron en placas LB /Spec. El ADN del plásmido purificado se secuenció en ambas direcciones mediante el uso de los cebadores M13 directo e inverso. En total, doce cepas de *Prototheca* se seleccionaron para tener las secuenciar de ADN de sus ARNr 23S y las secuencias se enumeran en la lista de secuencias. Un resumen de las cepas y los números del listado de secuencia se incluyen a continuación. Las secuencias se analizaron para la divergencia general a partir de la secuencia de la UTEX 1435 (SEQ ID NO: 15). Surgieron dos pares (UTEX 329/UTEX 1533 y UTEX 329/UTEX 1440) como los de máxima divergencia. En ambos casos, el alineamiento de pares resultó en 75.0% de identidad de secuencia por pares. El porcentaje de identidad de secuencia con UTEX 1435 también se incluye a continuación.

Especie	Cepa	% identidad nt	SEQ ID NO
Prototheca kruegani	UTEX 329	75.2	SEQ ID NO: 11
Prototheca wickerhamii	UTEX 1440	99	SEQ ID NO: 12
Prototheca stagnora	UTEX 1442	75.7	SEQ ID NO: 13
Prototheca moriformis	UTEX 288	75.4	SEQ ID NO: 14
Prototheca moriformis	UTEX 1439; 1441; 1435; 1437	100	SEQ ID NO: 15
Prototheca wikerhamii	UTEX 1533	99.8	SEQ ID NO: 16
Prototheca moriformis	UTEX 1434	75.9	SEQ ID NO: 17
Prototheca zopfii	UTEX 1438	75.7	SEQ ID NO: 18
Prototheca moriformis	UTEX 1436	88.9	SEQ ID NO: 19

Las muestras de lípidos de un subconjunto de las cepas enumeradas anteriormente se analizaron para el perfil lipídico usando HPLC. Los resultados se muestran a continuación en la Tabla 9.

Tabla 9. Diversidad de cadenas de lípidos en especies de microalgas

Cepa	**C14:0**	**C16:0**	**C16:1**	**C18:0**	**C18:1**	**C18:2**	**C18:3**	**C20:0**	**C20:1**
UTEX 327	0	12.01	0	0	50.33	17.14	0	0	0
UTEX 1441	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0
UTEX 1435	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

Los péptidos de tránsito al plastidio de algas se identificaron a través del análisis de UTEX 1435 (*Prototheca moriformis*) o de UTEX 250 (*Chlorella prototheco*ides) como se describen en el Ejemplo 12 y en el Ejemplo 11. Los ADNc que codifican las proteínas potencialmente dirigidas a los plastidios, en base a los aciertos de homología de BLAST con otras proteínas conocidas dirigidas a los plastidios, se sometieron a análisis adicional mediante los programas de software PSORT (http://psort.ims.u-tokyo.ac.jp/form.html), ChloroP (http://www.cbs.dtu.dk/services/ChloroP/) y TargetP (http://www.cbs.dtu.dk/services/TargetP/). Los péptidos de tránsito a plastidios candidatos que se identificaron a través de al menos uno de estos tres programas se amplificaron entonces por la PCR del ADN genómico apropiado. A continuación se muestra un resumen de las secuencias de aminoácidos de las secuencias dirigidas a los plastidio de algas (PTS), que se identificaron a partir de esta selección. También se incluyen las secuencias de aminoácidos de las acil grasos ACP tioesterasas de plantas que se usaron en los ejemplos de expresión heteróloga a continuación.

ADNc	SEQ ID NO.
P.moriformis isopentenil difosfato sintasa PTS	SEQ ID NO: 127
P.moriformis delta 12-desaturasa de ácido graso PTS	SEQ ID NO: 128
P.moriformis estearoil ACP desaturasa PTS	SEQ ID NO: 129
C.protothecoides estearoil ACP desaturasa PTS	SEQ ID NO: 130
Cuphea hookeriana acilo graso-ACP tioesterasa (C8-10)	SEQ ID NO: 131
Umbellularia californica acil graso-ACP tioesterasa (C12)	SEQ ID NO: 132
Cinnamomum camphora acil graso-ACP tioesterasa (C14)	SEQ ID NO: 133

Ejemplo 2: Cultivo de *Prototheca* en diversas materias primas

A. Sorgo

Las siguientes cepas demostraron ser capaces de usar el sorgo como única fuente de carbono: las cepas de *Prototheca moriformis* UTEX 1435, UTEX 1437, UTEX 288, UTEX 1439, UTEX 1441 y UTEX 1434, y la cepa de *Prototheca stagnora* UTEX 1442. La designación "UTEX" indica el número de cepa de la colección de cultivos de algas de la Universidad de Texas, Universidad del Estado A6700, Austin, Texas 78712-0183.

El sorgo puro se adquirió de Maasdam Sorghum Mills (Lynnville, Iowa) con un perfil de azúcar 21.0% p/p de fructosa, 28.0% p/p de dextrosa, 16.0% p/p de sacarosa y <0.5% p/p de maltosa. Los cultivos se cultivaron en medio líquido que contenía 2%, 5% o 7% (v/v) de sorgo puro (diluido del patrón puro) como la única fuente de carbono y los cultivos se cultivaron heterótrofamente en la oscuridad, con agitación a ~ 350 rpm. Se sacaron muestras de los cultivos a las 24, 40, 48, 67 y 89 horas y el crecimiento se midió mediante el uso de lecturas en un espectrofotómetro a A750. Se observó crecimiento en cada una de las cepas que se probaron como se muestra en las Figuras 1-2

B. Celulosa

El rastrojo de maíz explotado, el *Miscanthus*, el sorgo forrajero, la pulpa de remolacha y el bagazo de caña de azúcar húmedos, se prepararon por el National Renewable Energy Laboratory (Golden, CO) mediante la cocción en una solución 1.4% de ácido sulfúrico y desecación de la mezcla resultante. El porcentaje de sólidos se determinó gravimétricamente por secado y fue el siguiente: los sólidos del rastrojo de maíz 25%; los sólidos de *Miscanthus* 28.7%, los sólidos del sorgo forrajero 26.7%, y los sólidos del bagazo de la caña de azúcar 26%.

Muestras húmedas de 100 gramos los materiales celulósicos explotados (rastrojo de maíz o pasto forrajero) se resuspendieron en agua desionizada hasta un volumen final de 420 ml y el pH se ajustó a 4.8 con NaOH 10N. Para la pulpa de remolacha, 9.8 gramos de los sólidos secos se llevaron a 350 ml con agua desionizada y el pH se ajustó a 4.8 con NaOH 10 N. Para todas las materias primas anteriores, se usó la Accellerase 1000 (Genencor, Nueva York) en una proporción de 0.25 ml de la enzima por gramo de biomasa seca para la sacarificación de los materiales celulósicos. Las muestras se incubaron con agitación (110 rpm) a 50 0 C durante 72 horas. El pH de cada una de las muestras se ajustó a 7.0 con NaOH (con cambio de volumen despreciable), se esterilizó por filtración a través de un filtro de 0.22 µm y se usó en los procesos que se detallan a continuación. Para los procesos de mayor escala, se siguió el mismo procedimiento para la sacarificación, excepto un paso adicional de filtración de flujo tangencial (TFF) o un paso de microfiltración que se llevó a cabo para ayudar en la esterilización por filtración de las materias primas. Una muestra de cada una de las materias primas que se prepararon se reservó para la determinación de la concentración de glucosa y xilosa mediante un sistema que se basa en HPLC/ELSD o un kit que se basa en la hexoquinasa (Sigma). Además, para la pulpa de remolacha, el material se llevó inicialmente a un volumen como con las otras materias primas, el pH se ajustó a 4.0 y se realizó un tratamiento con pectinasa a 50 0 C durante 24 horas. El pH se ajustó entonces a 4.8 si no se llevaron a cabo los procesos de lavado o 5.3 si se llevaron a cabo los procesos de lavado. La sacarificación enzimática se realizó con el mismo procedimiento que usó para las otras materias primas como se describió anteriormente.

La cepa UTEX 1435 de microalgas *Prototheca moriformis* se evaluó por su capacidad de crecer en una serie de materias primas celulósicas que se prepararon como se describió anteriormente (el rastrojo de maíz, la pulpa de remolacha, la caña de sorgo, el *Miscanthus* y el control de glucosa). El cultivo de microalgas creció en las condiciones descritas en el Ejemplo 1 anterior, con excepción de la fuente de carbono. La fuente de carbono fue 4% de glucosa (para las condiciones de control) o 4% de glucosa medida por la glucosa disponible en los materiales celulósicos. El crecimiento se evaluó por las lecturas a A750 y el tiempo de cultivo fue de 168 horas, con lecturas a A750 a las 48, 72, 96, 120, 144 y 168 horas después del inicio del cultivo. Como se puede observar en la Figura 7a, el cultivo de *Prototheca moriformis* creció mejor en el rastrojo de maíz. Las otras materias primas celulósicas que se usaron, el *Miscanthus*, la caña de sorgo y la pulpa de remolacha, todas mostraron inhibición del crecimiento.

En base a los resultados anteriores con los azúcares celulósicos derivados del rastrojo de maíz, también se evaluó la acumulación de lípidos en *Prototheca moriformis* mediante el uso de diferentes niveles de azúcares celulósicos derivados del rastrojo de maíz y glucosa reactivo como control. Los cultivos crecieron en 18 g/l de glucosa que era completamente de azúcares celulósicos derivados del rastrojo de maíz (condición 100% del rastrojo

de maíz en la Figura 7b), 9 g/l de glucosa de azúcares celulósicos derivados del rastrojo de maíz que se suplementó con 9 g/l de glucosa reactivo (condición de 50% del rastrojo de maíz que se suplementó con glucosa a 18 g/l en la Figura 7b), 9 g/l de glucosa de azúcares celulósicos derivados del rastrojo de maíz (50% del rastrojo de maíz, que no se suplementó; glucosa a condición de 9 g/l en la Figura 7b) y un cultivo de control de 42g/l de la glucosa reactivo y 13 g/l de la xilosa reactivo para el control de la osmolaridad. Todos los cultivos se alimentaron con azúcares celulósicos para mantener la concentración de glucosa de 20 g/l, a excepción del cultivo de control, que se alimentó con la glucosa reactiva para mantener la concentración de glucosa a 20 g/l. El crecimiento se midió en base al peso seco de la célula del cultivo y la productividad de los lípidos se determinó como un porcentaje del peso seco de las células. Los lípidos totales se determinaron gravimétricamente mediante el uso de una sistema de Ankom de hidrólisis ácida/extracción por solvente como se describió en el Ejemplo 1.

Como se puede observar en la Figura 7b, en base a la acumulación de biomasa (medida por el PSC), todas las concentraciones de los celulósicos derivados del rastrojo de maíz superaron (mayor PSC) el medio de control que se alimentó con glucosa sola. La producción de lípidos como porcentaje del PSC se calculó también para todas las condiciones. Además de la acumulación más alta de biomasa que se observó para el crecimiento en el rastrojo de maíz, la acumulación de lípidos también fue mayor en las condiciones de los celulósicos derivados del rastrojo de maíz en comparación con la condición de control de glucosa. Estos datos demostraron que, además de proporcionar azúcares derivados de celulósicos, el rastrojo de maíz proporciona nutrientes adicionales o componentes que contribuyen a una mayor acumulación de biomasa (crecimiento) y un mayor rendimiento del producto.

Debido a que las materias primas celulósicas contienen componentes, además de la glucosa, algunos de estos componentes adicionales se pueden acumular a niveles indeseables en el cultivo, cuando otros azúcares derivados celulósicos se introducen en el cultivo como la principal fuente de carbono (por lo general, pero sin limitarse a, la glucosa) se consumen. Por ejemplo, la xilosa, el azúcar presente en la materia prima derivada de celulosa se puede acumular durante el cultivo de alta densidad de microalgas a niveles inhibitorios del crecimiento y la producción del producto final. Para probar los efectos de la acumulación de xilosa durante el cultivo de *Prototheca*, los cultivos se cultivaron con 4% de glucosa en el medio y se suplementó con 0.10 g/l, 25 g/l, 50 g/l y 100 g/l de xilosa. Después de 6 días de cultivo, se evaluó el crecimiento y la acumulación de lípidos mediante los métodos que se describieron anteriormente. Como se observa en la Figura 7c, sorprendentemente, las mayores concentraciones de xilosa que se probaron no inhibieron la capacidad de la *Prototheca moriformis* para crecer y acumular lípidos y el cultivo en realidad creció más y acumuló más lípidos en las mayores concentraciones de xilosa. Para explorar este fenómeno, se llevó a cabo un experimento similar con sacarosa, una fuente de carbono, que el tipo salvaje de *Prototheca moriformis* es incapaz de metabolizar. No se observó impacto positivo con sacarosa, lo que sugiere que el aumento del crecimiento y la acumulación de lípidos que se vio con la xilosa son atribuibles a un mecanismo distinto del estrés osmótico a partir de altas concentraciones de componentes sin metabolizar en el medio y es específico de la xilosa.

Además de los azúcares sin metabolizar, las sales se pueden acumular a niveles inhibitorios como consecuencia de la concentración de azúcares derivados de lignocelulósicos. Debido a la etapa de hidrólisis ácida con H_2SO_4 durante la preparación típica de materiales celulósicos, seguido de la neutralización del ácido con NaOH, se forma Na_2SO_4 durante la generación de azúcares lignocelulósicos. Para evaluar el impacto de la concentración de sal en el crecimiento y la producción de lípidos, los cultivos de *Prototheca moriformis* se cultivaron en concentraciones de Na_2SO_4 en el intervalo desde 0 hasta 700 mM en medio suplementado con 4% de glucosa. Como se muestra en la Figura 7d, se observó una inhibición significativa del crecimiento, medida por la acumulación de PSC, donde las concentraciones de Na_2SO_4 estuvieron por encima de 25 mM, específicamente a concentraciones de 80 mM, 240 mM y 700 mM. Además, se evaluó en la misma prueba el impacto del antiespumante P2000. El compuesto antiespumante tuvo un impacto positivo significativo sobre la productividad de la biomasa. La productividad de los lípidos también se evaluó para cada condición, y las concentraciones de Na2SO4 superiores a 80 mM, específicamente 240 mM y 700 mM, inhibieron, mientras que la adición del antiespumante P2000 aumentó significativamente la productividad de los lípidos. Así, en una realización, las etapas de los métodos de cultivo de la presente invención incluyen el cultivo en medio que contiene un agente antiespumante.

En base a los resultados que se discutieron anteriormente y que se resumen en la Figura 7a, los inhibidores que probablemente estaban presentes en las materias primas de base celulósica exhibieron un crecimiento deficiente. La presente invención proporciona los medios de eliminación de tales compuestos mediante el lavado del material con agua caliente (tratamiento hidrotermal). La Figura 8 resume los resultados del crecimiento, medido por A750, mediante el uso del azúcar derivado de materias primas celulósicas con un solo lavado con agua caliente. Las condiciones de cultivo fueron idénticas a las que se usaron en los procesos que se resumen en la Figura 7a. En comparación con los resultados que se muestran en la Figura 7a, después de sólo un lavado con agua caliente, los cultivos de *Prototheca moriformis* crecieron mejor en todas las materias primas celulósicas que se probaron, específicamente el bagazo de la caña de azúcar, la caña de sorgo, el *Miscanthus* y la pulpa de remolacha, en comparación con el control de glucosa. La productividad de los lípidos también se evaluó en cada una de las condiciones. A excepción de la condición de la pulpa de remolacha, que fue comparable al control de glucosa, los cultivos que crecieron en los azúcares derivados a partir de materiales celulósicos sometidos a un lavado de agua

caliente exhibieron una mejor productividad de lípidos que el control de glucosa.

Un impacto potencial del tratamiento hidrotérmico (lavado con agua caliente) de la biomasa celulósica es la eliminación de furfurales y hidroximetilfurfurales , liberados por la explosión ácida del material. La presencia de furfurales e hidroximetilfurfurales puede que contribuyera al crecimiento limitado que se observó en algunos de los procesos que se resumen en la Figura 7a. Para evaluar cómo el tratamiento hidrotérmico afectó los niveles de furfurales (FA) y hidroximetilfurfurales (HMF), los sobrenadantes resultantes de uno hasta tres lavados de la biomasa celulósica derivada del bagazo de la caña de azúcar (B), la caña de sorgo (S), el *Miscanthus* (M) o de la pulpa de remolacha (BP) se evaluaron por HPLC para los FA y los HMF. Como se muestra en la Figura 8, los niveles de AF y HMF disminuyeron de forma significativa con cada etapa de lavado. Este resultado es consistente con la observación de que los FA y los HMF pueden ser inhibidores del crecimiento de microalgas (como se ve en la Figura 7a) y que el tratamiento hidrotérmico elimina estos compuestos y resulta en el aumento del crecimiento de microalgas, incluso mejor que el crecimiento en las condiciones de control de glucosa (como ve en la Figura 8).

Se evaluó el impacto sobre el perfil lipídico de los cultivos de *Prototheca moriformis* que crecieron en los diversos azúcares derivados de lignocelulósicos con tratamiento hidrotérmico. Los cultivos de *Prototheca moriformis* se cultivaron en las siguientes materias primas celulósicas que se lavaron 4 veces: el *Miscanthus*, el bagazo de la caña de azúcar y la caña de sorgo, con los niveles de glucosa que se mantuvieron a 20 g/l a través de la alimentación de los azúcares celulósicos. Al final del cultivo, en la biomasa de microalgas a partir de cada condición se analizó el perfil lipídico mediante el uso de los métodos que se describieron en el Ejemplo 1. Los resultados del análisis del perfil lipídico (que se expresó en % del área) se resumen en la Tabla 10. Cada condición se probó por duplicado, y se incluyeron los resultados de cada uno de los duplicados de las condiciones de prueba. El crecimiento en materias primas celulósicas resultó en una significativa redistribución en el perfil lipídico en comparación con el control de glucosa. Por ejemplo, hubo un aumento significativo de C18:0 % del área en todas las condiciones de materias primas celulósicas en comparación con la condición de control de glucosa.

Tabla 10. Perfil de lípidos de *Prototheca moriformis* cultivada en azúcares derivadas de glucosa y materias primas celulósicas.

	glucosa 1 (ctrl)	glucosa 2 (ctrl)	bagazo 1	bagazo 2	sorgo 1	sorgo 2	*Miscan* 1	*Miscan* 2
C10:0	n.d.	n.d.	0.03	0.02	n.d.	n.d.	n.d.	n.d.
C12:0	0.04	0.05	0.04	0.04	0.05	0.04	0.04	0.04
C14:0	1.64	1.64	1.07	1.10	1.17	1.14	1.08	1.12
C14:1	0.03	0.04	0.04	0.04	0.06	0.06	0.03	0.03
C15:0	0.04	0.05	0.07	0.05	0.08	0.08	0.06	0.06
C16:0	26.80	26.81	22.32	22.81	22.09	22.19	23.45	23.62
C16:1	0.75	0.82	1.68	1.70	1.92	2.12	1.38	1.23
C17:0	0.14	0.16	0.28	0.17	0.29	0.27	0.21	0.19
C17:1	0.07	0.06	0.10	0.10	0.13	0.12	0.10	0.09
C18:0	3.56	3.64	15.88	10.40	15.30	12.37	10.15	8.69
C18:1	54.22	54.01	49.87	53.87	49.35	50.80	54.05	55.26
C18:2	11.23	11.11	6.54	7.91	7.47	8.80	7.71	7.88
C18:3 alfa	0.84	0.85	0.39	0.56	0.47	0.53	0.56	0.60
C20:0	0.31	0.30	0.85	0.63	0.76	0.69	0.63	0.56
C20:1	0.15	0.15	0.33	0.28	0.32	0.32	0.27	0.25
C20:3	0.06	0.06	0.13	0.12	0.14	0.12	0.11	0.11
C24:0	0.12	0.12	0.22	0.19	0.22	0.20	0.18	0.15

n.d. denota ninguno detectado

La secuencia de azúcar celulósica se generó a partir del rastrojo de maíz explotado, sacarificado mediante el uso de la enzima Accellerase y se concentró por evaporación al vacío. Esta secuencia de azúcar se probó en ensayos de crecimiento de *Prototheca moriformis* a una concentración de 4% de glucosa. Los resultados de los ensayos de crecimiento mostraron un crecimiento muy pobre y se probó la conductividad de la secuencia del azúcar celulósica (contenido de sal). La conductividad fue muy alta, mucho mayor de 700 mM equivalentes de sodio, un nivel que se mostró que tenía efectos inhibidores del crecimiento como se describió anteriormente y que se muestra en la Figura 7d. Los métodos de la invención incluyen métodos en los que se reduce o se elimina la sal de azúcares derivados de lignocelulósicos antes de usar estas materias primas en la producción de aceite de microalgas derivados de lignocelulósicos. Sorprendentemente, sin embargo, no se pueden usar resinas para la desalinización de las secuencias de azúcar concentrado, primero hay que diluir la secuencia de azúcar concentrada. Para demostrar esta realización de la invención, los azúcares celulósicos derivados del material de rastrojo de maíz se diluyeron ocho veces antes de eliminar con la resina las sales contaminantes. La conductividad inicial del material de partida concentrado fue de 87 mS/cm mientras que, la de la secuencia diluida ocho veces, fue 10990 µS/cm a un pH

de 5.61. Estudios anteriores indicaron que la falla al diluir la secuencia de azúcar concentrada antes de la desionización resultó en una incapacidad para eliminar cuantitativamente las sales, así como una pérdida significativa de glucosa a partir de la secuencia de azúcar. Se usaron tres volúmenes diferentes de lecho de la resina IEX (DOWEX Marathon MR3) (1:2, 1:4 y 1:10). La Tabla 11 resume los resultados que demuestran la capacidad de un lecho mixto de resina de intercambio iónico (IEX) para reducir las sales (medida por la conductividad) de manera significativa, en una previamente concentrada secuencia de azúcar celulósico derivado del rastrojo de maíz, en las materias primas diluidas.

Tabla 11. Capacidad de la resina para reducir las sales.

Volumen del lecho de resina : celulósicos	pH post-desionización	Conductividad post-desionización (µS/cm)	Conductividad post-desionización calculada y re-concentración 8x (µS/cm)	equivalentes Na^+ (basado en la curva std) en mM
1:2	3.1	74	592	7.42
1:4	3.1	97	776	9.7
1:10	5.25	6320	50560	634

Un proceso que emplea un volumen de lecho: materias primas celulósicas de 1:4 y la re-concentración del material de ocho veces resultaría en una concentración de sodio que está bien dentro del intervalo para la biomasa normal y la acumulación de lípidos. Por otra parte, la desionización o eliminación de sales se puede realizar antes de la sacarificación o después de la sacarificación, pero antes de la concentración de la secuencia de azúcar. Si la eliminación de sal se realiza antes de la concentración de la cadena de azúcar, un paso de dilución de la cadena de azúcar antes de la eliminación de sal, probablemente no sería necesario.

Este ejemplo demuestra la eficacia del lavado del material celulósico explotado para el uso en la producción de aceite celulósico. Como se describió anteriormente, la concentración de azúcares derivados de celulosa sin la eliminación de sales (inherente a la producción de material celulósico explotado y el posterior tratamiento) resultaron en fermentaciones sub-óptimas. Los materiales tratados en el proceso descrito a continuación estaban en el pH apropiado para la posterior sacarificación. Además, la conductividad de este material se redujo significativamente (más de 100 veces) a partir de la materia prima de partida. Por lo tanto, los azúcares concentrados posteriormente que se usaron en fermentaciones, no inhibieron, debido a la presencia excesiva de sales. Una ventaja adicional es vista por la eliminación de furfurales del material celulósico. Cualquier xilosa o glucosa eliminada en la fracción hemicelulósica se puede desechar o preferentemente re-concentrar para su uso en fermentaciones.

El bagazo de la caña de azúcar explotado, húmedo (NREL, Colorado), con una masa inicial de partida de 65 kg de peso húmedo y la conductividad de 15000 µS/cm, pH 2.4 se llevó a 128 kg con agua desionizada y se ajustó el pH a 4.6 con NaOH 10 N, lo que hizo la conductividad resultante de 6800 µS/cm). El porcentaje de sólidos se evaluó mediante el retiro de una alícuota de los materiales en suspensión a una cacerola de aluminio tarada (peso = t), se registró el peso húmedo (peso = w), seguido por el secado durante tres horas a 110°C. Después de secar las muestras se trasladaron a un desecador y se dejaron a temperatura ambiente (25°C) y en ese momento, se volvieron a pesar (peso = d). El porcentaje de sólidos se calculó como: % de sólidos = [(d-t/w-t)] x 100. La conductividad se midió en un medidor de conductividad Thermo Electron Orion 3 Star.

El bagazo de la caña de azúcar se lavó de manera semi-continua mediante la mezcla continua de la pasta de celulosa (porcentaje de sólidos iniciales del 8.2%) a una temperatura de 50°C en un reactor de acero inoxidable (150 l de capacidad). Los celulósicos se descargaron de la vasija del reactor a través de una bomba de carga rotatoria a una velocidad de flujo de 1.9-3.8 kg/min a una centrífuga decantadora Sharples Modelo 660. El permeado líquido se mantuvo discontinuo (aprox. 35-175 kg alícuotas, véase la Tabla 12 a continuación) y se retiraron alícuotas homogéneas para la evaluación de los azúcares totales (glucosa y xilosa) y del porcentaje de sólidos como se describió en la Tabla 12. La conductividad y el pH del material celulósico se controlaron a través de la adición de agua desionizada y NaOH 10 N, respectivamente. Las muestras de 1-10 en la Tabla 12 representan el permeado que se decantó en la centrífuga, y, como tal, los sólidos y los azúcares presentes en estas fracciones se eliminaron del lavado final de los materiales celulósicos. Un cálculo de balance de masa de sólidos totales comparó los sólidos eliminados menos los sólidos que se perdieron más los sólidos finales para la sacarificación, lo que resultó en una recuperación del 99 % en el proceso anterior. La Figura 8 resume la concentración de los furfurales y los hidroximetilfurfurales (mg/l) en cada una de los 11 permeados de centrifugas que se reunieron y se describieron en la Tabla 12. Estos datos demuestran una eliminación clara de los furfurales y los hidroximetilfurfurales del bagazo de la caña de azúcar.

Tabla 12. Balance de masa para el tratamiento hidrotérmico semi-continuo de bagazo de caña de azúcar.

Muestra	kg (húmedo)	kg (seco)	pH	Conductividad µS/cm	total de xilosa eliminada (g)	total de glucosa eliminada (g)
1 (material inical)	128	10.50	4.60	6,880	0	0
2	81.8	2.03		3,280	1030.68	286.3
3	76.5	0.49		2,500	298.35	76.50
4	106	0.41			254.40	63.60
5	173.9	0.30	3.74	1,260	226.07	69.56
6	101.8	0.08	4.40	791	71.26	20.36
7	110.6	0.04	4.86	327	44.24	0
8	77.2	0			0	0
9	108.6	0.02	4.7	221	0	0
10	101.5	0			0	0
11	34.8	0	4.7	146	0	0
Sólidos eliminados (muestras 1-10) que se pierden en el proceso		3.37				
Total de xilosa eliminada					1925.00	
Total de glucosa eliminada						516.32
Sólidos finales para sacarificación	7.03					

En otra demostración de la capacidad de la Prototheca para usar materia prima derivada de celulosa, la *Prototheca moriformis* (UTEX 1435) se cultivó en biorreactores de tres litros con azúcar derivada de celulosa como materia prima de carbono fijo. El inóculo se preparó a partir de células criopreservadas, las cuales se descongelaron a temperatura ambiente y 1 ml de células se añadió 300 ml del medio de inóculo que se basa en el medio basal de microalgas que se describió en el Ejemplo 1 con 1 g/l $(NH_4)_2SO_4$, 4 g/l de extracto de levadura y una solución de elementos traza, más 4% de glucosa y se cultivó durante 1 día a 28°C con agitación (200 rpm). Este cultivo se usó para inocular un biorreactor de tres litros que contenía 1l de medio más 0.26 ml del Antiespumante 204 (Sigma, Estados Unidos.). El fermentador se controló a 28°C y el pH se mantuvo en 6.8 por adición de KOH. El oxígeno disuelto se mantuvo a 30% de saturación mediante agitación en cascada y flujo de aire. El cultivo se alimentó con la materia prima de azúcar celulósica a partir del rastrojo de maíz para mantener 0-10 g/l de glucosa. La desalinización de materias primas de azúcar celulósico a menos de 300 mM de sal fue esencial para asegurar un peso de células secas similar y el rendimiento de la acumulación de lípidos en comparación con controles de materia prima del azúcar purificado. La desalinización de la materia prima del azúcar celulósico se realizó mediante los métodos que se describieron anteriormente. Se tomaron muestras del fermentador para supervisar el rendimiento de la fermentación. La acumulación de la masa celular se supervisó por la densidad óptica y el peso seco de las células. Las concentraciones de glucosa, xilosa, amoniaco, potasio, sodio y furfural también se determinaron y se supervisaron a lo largo del curso temporal de la fermentación. La concentración de lípidos se determinó por los métodos gravimétricos que se discutieron anteriormente.

Ejemplo 3: Métodos para la transformación de *Prototheca*

A. <u>Método general para la transformación biolística de *Prototheca*</u>

Los transportadores de oro S550d de Seashell Technology se prepararon de acuerdo con el protocolo del fabricante. El plásmido linealizado (20 µg) se mezcló con 50 µl de tampón de unión y 60 µl (30 mg) de los transportadores de oro S550d y se incubaron en hielo durante 1 min. El tampón de precipitación (100 µl) se añadió, y la mezcla se incubó en hielo durante otro minuto. Después de agitación, las partículas de ADN recubierto se sedimentaron por rotación a 10,000 rpm en una microcentrífuga Eppendorf 5415C durante 10 segundos. El precipitado de oro se lavó una vez con 500 µl de etanol frío al 100%, se sedimentó por rotación breve en la microcentrífuga, y se resuspendió con 50 µl de etanol helado. Después de una breve (1-2 seg) sonicación, 10 µl de partículas de ADN recubiertas se trasladaron de inmediato a la membrana transportadora.

Las cepas de Prototheca se cultivaron en medio proteosa (2 g/l de extracto de levadura, NaNO3 2.94 mM, CaCl2 • 2H2O 0.17 mM, MgSO4 • 7H2O 0.3 mM, K2HPO4 0.4 mM, KH2PO4 1.28 mM, NaCl 0.43 mM) en un agitador giratorio hasta alcanzar una densidad celular de $2x10^6$ células/ml. Las células se cosecharon, se lavaron una vez con agua destilada estéril, y se resuspendieron en 50 µl de medio. Se extendieron 1 x 10^7 células en el

tercer centro de una placa con medio no selectivo de proteosa. Las células se bombardearon con el sistema biolístico de entrega de partículas PDS-1000/He (Bio-Rad). Se usaron los discos de ruptura (1100 y 1350 psi), y las placas se colocaron 9 y 12 cm por debajo del ensamblaje de la pantalla/macro-transportador. A las células se les permitió recuperarse a 25°C durante 12-24 h. Tras la recuperación, las células se rasparon de las placas con una espátula de goma, se mezclaron con 100 µl de medio y se dispersaron en las placas que contenían la selección de antibióticos adecuada. Después de 7-10 días de incubación a 25°C, las colonias que representan las células transformadas eran visibles en las placas a partir de los discos de ruptura de 1100 y 1350 psi y a partir de 9 y 12 cm de distancia. Las colonias se recogieron y se aplicaron en forma de puntos en placas de agar selectivo para una segunda ronda de selección.

B. Transformación de *Prototheca* con el gen de resistencia al G418

La Prototheca moriformis y otras cepas de *Prototheca* sensibles al G418 se pueden transformar mediante el uso de los métodos que se describen a continuación. El G418 es un antibiótico aminoglucósido que inhibe la función de los ribosomas 80S y por ello, inhibe la síntesis de proteínas. El correspondiente gen de resistencia funciona a través de la fosforilación, lo que resulta en la inactivación del G418. Las cepas de *Prototheca* UTEX 1435, UTEX 1439 y UTEX 1437 se seleccionaron para la transformación. A las tres cepas de *Prototheca* se les determinó el genotipo mediante el uso de los métodos que se describieron anteriormente. Las tres cepas de *Prototheca* tuvieron secuencias genómicas idénticas del ARNr 23S (SEQ ID NO: 15).

Todos los casetes de transformación se clonaron como fragmentos EcoRI-SacI en pUC19. Se usaron técnicas estándar de biología molecular en la construcción de todos los vectores de acuerdo a Sambrook y Russell, 2001. El promotor de la beta-tubulina/5'UTR de *C. reinhardtii* se obtuvo del plásmido pHyg3 (Berthold *y otros*, (2002) Protist: 153(4), pp 401-412) por la PCR como un fragmento EcoRI-AscI. La nitrato reductasa 3'UTR de *Chlorella vulgaris* se obtuvo a partir de ADN genómico aislado de la cepa UTEX 1803 a través de la PCR mediante el uso de los pares de cebadores siguientes:

Directo: 5'TGACCTAGGTGATTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCG 3' (SEQ ID NO:35)

Inverso: 5' CTACGAGCTCAAGCTTTCCATTTGTGTTC CCATCCCACTACTTCC 3' (SEQ ID NO:36)

El promotor/UTR de la glutamato deshidrogenasa de *Chlorella sorokiniana* se obtuvo a través de la PCR del ADN genómico aislado de la cepa UTEX 1230 a través de la PCR mediante el uso de los pares de cebadores siguientes:

Directo:5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (SEQ ID NO:37)

Inverso:5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (SEQ ID NO:38)

La optimización de codón se basó en los codones en la Tabla 1 para la *Prototheca moriformis*. La secuencia del casete de la neomicina fosfotransferasa en la que no se optimizaron los codones (nptII) se sintetizó como un fragmento AscI-XhoI y se basó en la secuencia con núm. de acceso YP_788126 en el GenBank. El casete nptII con los codones optimizados se basó también en este número de acceso en el Genbank.

Las tres cepas de *Prototheca* se transformaron mediante el uso de los métodos biolísticos que se describieron anteriormente. Brevemente, las cepas de *Prototheca* se cultivaron heterofílicamente en medio líquido que contenía 2% de glucosa hasta alcanzar la densidad celular deseada ($1x10^7$ células/ml a $5x10^7$ células/ml). Las células se cosecharon, se lavaron una vez con agua destilada estéril y se resuspendieron a $1x10^8$ células/ml. Después, 0.5 ml de células se extendieron en una placa con medio sólido no selectivo y se dejó secar en una campana estéril. Las células se bombardearon con el sistema biolístico de entrega de partículas PDS-1000/He (BioRad). A las células se les permitió recuperarse a 25°C durante 24 horas. Tras la recuperación, las células se retiraron mediante el lavado de las placas con 1 ml de medio estéril y se transfirieron a nuevas placas que contenían 100 mg/ml de G418. Las células se dejaron secar en una campana estéril y se dejó que se formaran colonias en la placa a temperatura ambiente durante un máximo de tres semanas. Las colonias de UTEX 1435, UTEX 1439 y UTEX 1437 se recogieron y se aplicaron en forma de puntos en placas de agar selectivo para una segunda ronda de selección.

Un subconjunto de las colonias que sobrevivieron a una segunda ronda de selección descrita anteriormente, se cultivaron en pequeño volumen y el ADN genómico y el ARN se extrajeron por medio de técnicas de biología molecular. Se hizo la transferencia de tipo Southern sobre el ADN genómico extraído de las no transformadas (WT), los transformantes y el ADN del plásmido. El ADN de cada muestra se corrió en geles de agarosa 0.8% después de los siguientes tratamientos: sin digerir (U), digerido con AvrII (A), digerido con NcoI (N), digerido con SacI (S). Los ADN de estos geles se transfirieron a membranas de Nylon + (Amersham). Estas membranas se probaron con un fragmento que corresponde a toda la región de codificación del gen nptII (sonda NeoR). La Figura 4 muestra los mapas de los casetes que se usaron en las transformaciones. La Figura 5 muestra

los resultados del análisis de la transferencia de tipo Southern en tres transformantes (todos generados en la cepa UTEX 1435) (1, 2, y 3) que se transformaron ya sea con la beta-tubulina::neo::nit (SEQ ID NO: 39) (los transformantes 1 y 2) o con la glutamato deshidrogenasa:neo:nit (SEQ ID NO: 40) (transformante 3). El plásmido de transformación glutamato deshidrogenasa:neo: nit se corrió como un control y se cortó tanto con NcoI y con SacI. La AvrII no cortó en este plásmido. El ADN genómico aislado de la cepa sin transformar UTEX 1435 no mostró hibridación de la sonda NeoR.

Las construcciones de transformantes adicionales que contienen la glutamato deshidrogenasa:neo:nit con codón optimizado (SEQ ID NO: 41) y la β-tubulina::neo::nit con codón optimizado (SEQ ID NO: 42) se recogieron y se analizaron por análisis de transferencia de tipo Southern. Como era de esperar, sólo la digestión con la SacI mostró linealización del ADN de transformación. Estos eventos de transformación son consistentes con los eventos de integración que se producen en forma de oligómeros del plásmido de transformación. Sólo la digestión con las enzimas de restricción que cortan el ADN del plásmido de transformación hace en estas moléculas la disminución del tamaño del plásmido de transformación.

El análisis de transferencia de tipo Southern se realizó también en transformantes generados en la transformación de las cepas de *Prototheca* UTEX 1437 y UTEX 1439 con el casete de la glutamato deshidrogenasa::neo::nit. La transferencia se probó con la sonda NeoR y los resultados son similares a los transformantes UTEX 1435. Los resultados son indicativos de eventos de integración que se caracterizaron por la oligomerización y la integración del plásmido de transformación. Este tipo de evento de integación se sabe que ocurre con bastante frecuencia en *Dictyostelium discoideum* (véase, por ejemplo, Kuspa, A. y Loomis, W. (1992) *PNAS*, 89:8803-8807 y Morio *y otros*, (1995) *J. Plant Res.* 108:111-114).

Para confirmar aún más la expresión del plásmido de transformación, se realizaron en los transformantes que se seleccionaron el análisis por transferencia de tipo Northern y el análisis por RT-PCR. La extracción de ARN se realizó mediante el uso del reactivo Trizol según las instrucciones del fabricante. El análisis de transferencia de tipo Northern se llevó a cabo de acuerdo con los métodos publicados en Sambrook y Russell, 2001. El ARN total (15 µg) aislado de cinco transformantes UTEX 1435 y UTEX 1435 sin transformar (carriles de control) se separó en gel de agarosa-formaldehído 1% y se transfirió a membrana de nylon. La transferencia se hibridó con la sonda neo sin optimizar, específica para las secuencias transgénicas en los transformantes 1 y 3. Los otros dos ARN transformantes, expresaron la versión de los neo-transgénicos con codón optimizado y, como era de esperar, en base a la homología de secuencia entre los neo genes que se optimizaron y los que no se optimizaron, mostraron la señal de hibridación significativamente menor. Se extrajo ARN (1 µg) de la cepa de *Prototheca* UTEX 1435 sin transformar y dos representantes transformantes de UTEX 1435 y se realizó la transcripción inversa mediante el uso de un cebador oligo dT o un cebador específico del gen. Posteriormente estos ADNc (por duplicado) se sometieron a análisis de qPCR en el termociclador ABI Veriti que usa la química SYBR-Green de qPCR mediante el uso de los siguientes cebadores (nptII):

Directo: 5' GCCGCGACTGGCTGCTGCTGG 3' (SEQ ID NO: 43)

Inverso: 5' AGGTCCTCGCCGTCGGGCATG 3' (SEQ ID NO: 44)

La posible contaminación de ADN genómico se descartó por una muestra de control negativo sin transcriptasa inversa. Los resultados indican que los genes NeoR que se usaron para transformar estas cepas se transcribieron activamente en los transformantes.

C. <u>Transformación de *Prototheca* con sacarosa invertasa secretada heteróloga</u>

Los siguientes experimentos se realizaron mediante el uso de medio líquido/placas de agar en base al medio basal descrito en Ueno y otros, (2002) *J Bioscience and Bioengineering* 94(2):160-65, con la adición de trazas de minerales descrita en la patente de Estados Unidos núm. 5,900,370, y el cóctel de vitamina DAS 1x (solución 1000x): tricina: 9g, tiamina HCL: 0.67g, biotina: 0.01g, cianocobalamina (vitamina B12): 0.008g, pantotenato de calcio: 0.02 g y ácido p-aminobenzoico: 0.04 g).

Dos construcciones de plásmido se ensamblaron mediante el uso de las técnicas estándar de ADN recombinante. Los genes de la sacarosa invertasa de levadura (uno con codones optimizados y uno con sin codones optimizados), suc2, estaban bajo el control del promotor de la beta-tubulina de *Chlorella reinhardtii*/5'UTR y tenía la nitrato reductasa de *Chlorella vulgaris* 3'UTR. Las secuencias (se incluyeron las secuencias 5'UTR y 3'UTR) para la construcción con codón no optimizado (Crβ-tub::NCO-suc2::CvNitRed), SEQ ID NO: 57, y la construcción con codón optimizado (Crβ-tub::CO-suc2::CvNitRed), SEQ ID NO: 58, se enumeran en el listado de secuencias. La optimización de codones se basó en la Tabla 1 para la *Prototheca sp.* La Figura 6 muestra un esquema de las dos construcciones con los correspondientes sitios de restricción de clonación y las flechas indican la dirección de la transcripción. La selección se proporcionó por el Neo R (con codón optimizado, Tabla 1).

Preparación del microtransportador ADN/oro: los microtransportadores ADN/oro se prepararon

inmediatamente antes de su uso y se almacenaron en hielo hasta que se aplicaron a los macrotransportadores. El ADN del plásmido (en tampón TE) se añadió a 50 µl de tampón de unión. La saturación de las perlas de oro se alcanzó a los 15 µg de ADN del plásmido para 3 mg del transportador de oro. El tampón de unión y el ADN se mezclaron bien a través de la agitación. El ADN y el tampón de unión se deben pre-mezclar antes de la adición de oro para asegurar la unión uniforme del plásmido a las partículas del transportador de oro. Se añadieron 60 µl del transportador de oro S550d (Seashell Technologies, San Diego, CA) a la mezcla ADN/tampón de unión. Para un patrón de oro a 50 mg/ml, la adición de 60 µl resulta en una proporción óptima de 15 µg ADN/3 mg del transportador de oro. La mezcla del transportador de oro/ADN se dejó incubar en hielo durante 1 minuto y entonces se añadieron 100 µl de tampón de precipitación. La mezcla se dejó incubar nuevamente en hielo durante 1 minuto y entonces brevemente se agitó y se centrifugó a 10,000 rpm a temperatura ambiente durante 10 segundos para sedimentar el transportador de oro. Se retiró el sobrenadante cuidadosamente con una pipeta y el sedimento se lavó con 500 µl de etanol 100% helado. Las partículas de oro se re-sedimentaron por centrifugación de nuevo a 10,000 rpm durante 10 segundos. El etanol se eliminó y se añadieron 50 µl de etanol helado a la mezcla de oro. Inmediatamente antes de aplicar el oro a los macrotransportadores, el oro/etanol se resuspendió con un breve pulso de 1-2 segundos en el nivel 2 en un baño de ultrasonido Misonix mediante el uso de la micro punta. Inmediatamente después de la resuspensión, 10 µl de la dispersión de partículas de oro se trasladaron al macrotransportador y se dejaron secar en una campana estéril.

Las dos cepas de *Prototheca moriformis* (UTEX 1435 y 1441) se cultivaron heterofílicamente en medio líquido que contenía 2% de glucosa de los frascos que se criopreservaron. Cada cepa se cultivó a una densidad de 10^7 células/ml. Este cultivo de semillas entonces se diluyó con medio fresco a una densidad de 10^5 células/ml y se dejó crecer durante 12-15 horas para alcanzar una densidad celular final de aproximadamente 10^6 células/ml. Las microalgas se formaron en alícuotas en tubos cónicos de 50 ml y se centrifugaron durante 10 minutos a 3500 rpm. Las células se lavaron con medio fresco y se centrifugaron de nuevo durante 10 minutos a 3500 rpm. Las células se resuspendieron a continuación a una densidad de 1.25 x 10^8 células/ml en medio fresco.

En una campana estéril se retiraron 0.4 ml de las células que se prepararon anteriomente y se colocaron directamente en el centro de una placa de agar (sin agente de selección). La placa se agitó suavemente con un nivel de movimiento circular para distribuir uniformemente las células a un diámetro de no más de 3 cm. Las células se dejaron secar en las placas en la campana estéril durante aproximadamente 30-40 minutos y entonces se bombardearon a una presión del disco de ruptura de 1350 psi y una placa con el macrotransportador a 6 cm de distancia. Las placas se cubrieron y se envolvieron con parafilm y se dejaron incubar con poca luz durante 24 horas.

Después de la recuperación de 24 horas, se introdujo 1 ml de medio estéril (sin glucosa) en el césped de células. Las células se resuspendieron mediante el uso de un asa estéril, que se aplicó en un movimiento circular al césped de las células y las células resuspendidas se recogieron con una pipeta estéril. Las células se sembraron a continuación, sobre una placa de agar fresco con 2% de glucosa y 100 mg/ml de G418. La aparición de colonias se produjo 7-12 días después de que se sembraron. Las colonias individuales se recogieron y se cultivaron en medio selectivo con 2% de glucosa y 100 µg/ml de G418. El tipo salvaje (sin transformar) y las células transgénicas se analizaron para la introducción, la integración y la expresión exitosa del transgén.

El ADN genómico de *Prototheca moriformis* UTEX 1435 y 1441 transformadas y sus contrapartes de tipo salvaje (sin transformar) se aislaron mediante el uso de métodos estándar. Brevemente, las células se centrifugaron durante 5 minutos a 14,000 rpm en una centrífuga de mesa estándar Eppendorf (modelo 5418) y se congelaron rápidamente antes de la extracción de ADN. Los sedimentos de células se lisaron mediante la adición de 200 µl de tampón de lisis (Tris HCl 100 mM, pH 8.0, lauril sarcosina 1%, NaCl 50 mM, EDTA 20 mM, sacarosa 0.25 M, ARNasa A 0.5 mg/ml) por cada 100 - 200 mg de células (peso húmedo) y agitación durante 30-60 segundos. El bromuro de cetil trimetil amonio (CTAB) y el NaCl se llevaron a 1% y 1 M, respectivamente, y los extractos de células se incubaron a 60-65°C durante 10 minutos. Posteriormente, los extractos se aclararon por centrifugación a 14,000 rpm durante 10 minutos y el sobrenadante resultante se extrajo con un volumen igual de fenol/cloroformo/alcohol isoamílico (25:24:1). Las muestras se centrifugaron durante 5 minutos a 14,000 rpm y la fase acuosa se eliminó. El ADN se precipitó con 0.7 volúmenes de isopropanol. El ADN se granuló a través de centrifugación a 14,000 rpm durante 10 minutos y se lavó dos veces con etanol al 80%, y una vez con etanol. Después del secado, el ADN se resuspendió en Tris HCl 10 mM, pH 8.0 y se determinaron las concentraciones de ADN mediante el uso del ensayo de cuantificación de la fluorescencia PicoGreen (Molecular Probes).

El ARN de *Prototheca moriformis* UTEX 1435 y 1441 y sus contrapartes de tipo salvaje (sin transformar) se aislaron mediante el uso de métodos estándar. Brevemente, las células se centrifugaron durante 5 minutos a 14,000 rpm en una centrífuga de mesa estándar Eppendorf (modelo 5418) y se congelaron rápidamente antes de la extracción del ARN. Los sedimentos de células se lisaron mediante la adición de 1 ml del reactivo Trizol (Sigma) por cada 100 mg de células (peso húmedo) y mediante agitación durante 1-2 minutos. Las muestras se incubaron a temperatura ambiente durante 5 minutos y posteriormente ajustaron con 200 µl de cloroformo por 1 ml del reactivo Trizol. Después de amplia agitación, las células se incubaron a temperatura ambiente durante 15 minutos y entonces se sometieron a centrifugación a 14,000 rpm durante 15 minutos en una microcentrífuga de mesa refrigerada. La partición del ARN en la fase acuosa superior se retiró y se precipitó por la adición de isopropanol (500 µl por 1 ml del

reactivo Trizol). El ARN se recogió por centrifugación durante 10 minutos y el sedimento resultante se lavó dos veces con 1 ml de etanol al 80%, se secó, y se resuspendió en agua libre de ARNasa. La concentración de ARN se estimó mediante el ensayo de cuantificación de fluorescencia RiboGreen (Molecular Probes).

La expresión del gen de la neomicina fosfotransferasa que confiere resistencia al antibiótico G418 y la levadura invertasa se probaron en las *Prototheca moriformis* no transformadas UTEX 1435 y 1441 y los transformantes T98 (UTEX transformante 1435) y T97 (UTEX 1441 transformante) mediante el uso del análisis por PCR cuantitativa en transcripción inversa (RT-PCR cuantitativa). 20 ng de ARN total (aislado como se describió anteriormente) se sometieron a una etapa de análisis de RT-PCR cuantitativa mediante el uso del kit de RT-PCR iScript SYBR Green (BioRad Laboratories) y pares de cebadores dirigidos al gen de resistencia a la neomicina (cebador directo 5'CCGCCGTGCTGGACGTGGTG 3' y el cebador inverso 5' GGTGGCGGGGTCCAGGGTGT 3', SEQ ID NO: 65 y 66, respectivamente) y los transcritos de *suc2* de invertasa (cebador directo 5' CGGCCGGCGGCTCCTTCAAC 3' y cebador inverso 5' GGCGCTCCCGTAGGTCGGGT 3', SEQ ID NO: 67 y 68, respectivamente). Los transcritos de la beta tubulina endógena sirvieron como control interno positivo para la amplificación por PCR y como referencia de normalización para estimar los niveles relativos de transcripción.

Ambas construcciones con codón optimizado y con sin codón optimizado se transformaron en las células de *Prototheca moriformis* UTEX 1435 y 1441 como se describió anteriormente. Inicialmente, los transformantes se obtuvieron con las dos construcciones y la presencia del transgén se comprobó mediante el análisis de transferencia de tipo Southern seguido por RT-PCR para confirmar la presencia del ADN y el ARNm del transgén. Para el análisis de transferencia de tipo Southern, el ADN genómico aislado como se describió anteriormente se sometió a electroforesis en geles de agarosa al 0.7% en tampón TAE 1x. Los geles se procesaron como se describe en Sambrook y otros (Molecular Cloning; A Laboratory Manual, 2da Edición. Cold Spring Harbor Laboratory Press, 1989). Las sondas se prepararon por cebado aleatorio e hibridaciones que se realizaron como se describe en Sambrook y otros. Los transformantes de las construcciones ya sea con codón optimizado y sin codón optimizado, mostraron la presencia del casete de la invertasa, mientras que el control no transformado fue negativo. El ARNm de la invertasa se detectó también en transformantes con las construcciones ya sea con codón optimizado y sin codón optimizado.

Para confirmar que los transformantes expresaban una proteína invertasa activa, los transformantes se sembraron en placas de sacarosa. Los transformantes que contenían el casete de codón no optimizado no crecieron en las placas que contenían sacarosa, lo que indicaba que, si bien los genes y el ARNm que codifica la proteína suc2 estaban presentes, la proteína o bien (1) no se tradujo, o (2) se tradujo, pero no se acumuló a niveles suficientes para permitir el crecimiento en la sacarosa como única fuente de carbono. Los transformantes con el casete de codón optimizado crecieron en las placas que contenían sacarosa. Para evaluar los niveles de la invertasa que expresaban estos transformantes, dos clones (T98 y T97) se sometieron a un ensayo de la invertasa de todas las células raspadas del medio sólido y el muestreo directo y la cuantificación de azúcares en los sobrenadantes de cultivo después de 48 horas de crecimiento en el medio líquido que contenía 2% de sacarosa como única fuente de carbono.

Para el ensayo de la invertasa, las células (T98 y T97) se cultivaron en placas que contenían 2% de sacarosa, se rasparon y se probó la actividad de la invertasa. Se mezclaron 10 µl de las células tomadas por raspado con 40 µl de NaOAc 50mM pH 5.1. Se añadieron a la mezcla de la células 12.5 µl de sacarosa 0.5 M y se incubaron a 37°C durante 10-30 minutos. Para detener la reacción se añadieron 75 µl de K_2HPO_4 0.2 M. Para el ensayo de la glucosa liberada, se añadieron 500 µl del reactivo reconstituido (glucosa oxidasa/peroxidasa + o-dianisidina) de Sigma (kit de prueba GAGO-20) a cada tubo y se incubaron a 37°C durante 30 minutos. Una curva de glucosa estándar se creó también en este momento (intervalo: 25 µl a 0.3 µl de glucosa). Después de la incubación, se añadieron 500 µl de HCl 6N a detener la reacción y desarrollar el color. Las muestras se leyeron a 540 nm. La cantidad de glucosa liberada se calculó a partir de la curva estándar de glucosa mediante la fórmula y = mx+c, donde y es la lectura de 540 nm, y x son los µl de glucosa. El peso de glucosa se convirtió en moles de glucosa, y dada la relación equimolar entre moles de sacarosa hidrolizada a los moles de glucosa generada, los datos se expresaron como nmoles de sacarosa hidrolizada por unidad de tiempo. El ensayo mostró que tanto los clones T98 y T97 hidrolizaron la sacarosa, lo que indicó que las células produjeron y secretaron una invertasa de sacarosa funcional.

Para el análisis de azúcar en los medios de cultivo líquido después de 48 horas de crecimiento de las algas, las células T97 y T98 se cultivaron en medio que contenía 2% de sacarosa durante 48 horas y los medios de cultivo se procesaron para el análisis de azúcar. Los caldos de cultivo de cada transformante (y el control negativo de células sin transformar) se centrifugaron a 14,000 rpm durante 5 minutos. El sobrenadante resultante se retiró y se sometió a HPLC/ELSD (detección de dispersión de la luz por evaporación). La cantidad de azúcar en cada muestra se determinó mediante el uso de estándares externos y el análisis de regresión de línea. Los niveles de sacarosa en el medio de cultivo de los transformantes fueron muy bajos (menos de 1.2 g/l, y en la mayoría de los casos 0 g/l). En los controles negativos, los niveles de sacarosa se mantuvieron altos, aproximadamente a 19 g/l después de 48 horas de crecimiento.

Estos resultados fueron consistentes con los resultados de actividad de la invertasa, y en conjunto, indicaron que los transformantes con codones optimizados, T97 y T98, secretaron una sacarosa invertasa activa que les permitió a las microalgas usar la sacarosa como única fuente de carbono, en contraste con (1) los transformantes sin codones optimizados y (2) las microalgas no transformadas de tipo salvaje, las cuales no pudieron usar sacarosa como única fuente de carbono en el medio de cultivo.

Las cepas de Prototheca moriformis, T98 y T97, que expresaban un transgén de sacarosa invertasa segregada (suc2), funcional se probaron para el crecimiento y la producción de lípidos mediante el uso de la sacarosa como única fuente de carbono.

Las cepas T98 y T97, de tipo salvaje (sin transformar), se cultivaron en medios de cultivo (como se describe anteriormente) que contenían 4% de glucosa o 4% de sacarosa como única fuente de carbono en condiciones heterótrofas durante unos 6 días. El crecimiento, según lo determinado por las lecturas de densidad óptica A750 que se tomaron de las cuatro muestras cada 24 horas y el peso seco de las células de los cultivos y los perfiles de lípidos se determinaron después de 6 días de crecimiento. Las lecturas de densidad óptica de las cepas transgénicas que crecieron tanto en las condiciones de glucosa y de sacarosa fueron comparables al crecimiento de las cepas de tipo salvaje en las condiciones de glucosa. Estos resultados indican que las cepas transgénicas crecieron en glucosa o sacarosa como única fuente de carbono a una velocidad igual a las cepas del tipo salvaje en condiciones de glucosa. Las cepas de tipo salvaje, sin transformar, no crecieron en la condición de sacarosa sola.

Se analizó el perfil lipídico en la biomasa de la cepa de tipo salvaje que creció en glucosa y de la cepa T98 que creció en sacarosa. Las muestras de lípidos se prepararon a partir de la biomasa seca (liofilizada) mediante el uso de un sistema de hidrólisis ácida (Ankom Technology, NY) de acuerdo con las instrucciones del fabricante. Las determinaciones del perfil lipídico se realizaron como se describió en el Ejemplo 4. El perfil lipídico de la cepa de *Prototheca moriformis* sin transformar UTEX cepa 1435, que creció en glucosa como única fuente de carbono y dos cepas colonal T98 (UTEX 1435 transformada con un transgén de sacarosa invertasa), que crecieron en sacarosa como única fuente de carbono, se describen en la Tabla 13 (tipo salvaje UTEX 1435 y clon 8 y el clon 11 de T98 a continuación. El lípido C:19:0 se usó como un control de calibración interna.

Tabla 13. Perfil lipídico de UTEX 1435 tipo salvaje y los clones de UTEX 1435 con el transgén *suc2* .

Nombre	tipo salvaje (% área - ISTD)	T98 clon 11 (% área - ISTD)	T98 clon 8 (% área - ISTD)
C 12:0	0.05	0.05	0.05
C 14:0	1.66	1.51	1.48
C 14:1	0.04	ND	ND
C 15:0	0.05	0.05	0.04
C 16:0	27.27	26.39	26.50
C 16:1	0.86	0.80	0.84
C 17:0	0.15	0.18	0.14
C 17:1	0.05	0.07	0.05
C 18:0	3.35	4.37	4.50
C 18:1	53.05	54.48	54.50
C 18:2	11.79	10.33	10.24
C 19:0 (ISTD)	--	--	--
C 18:3 alfa	0.90	0.84	0.81
C 20:0	0.32	0.40	0.38
C 20:1	0.10	0.13	0.12
C 20:1	0.04	0.05	0.04
C 22:0	0.12	0.16	0.12
C 20:3	0.07	0.08	0.07
C 24:0	0.12	0.11	0.10
nd - denota ninguno detectado			

El aceite extraído de *Prototheca moriformis* UTEX 1435 de tipo salvaje (a través de extracción con disolventes o mediante el uso de una prensa expulsora (ver métodos en el Ejemplo 44 anterior) se analizó para los carotenoides, la clorofila, los tocoferoles, otros esteroles y tocotrienoles. Los resultados se resumen a continuación en la Tabla 14.

Tabla 14. Análisis de carotinoide, clorofila, tocoferol/esteroles y tocotrienol en aceite extraido de *Prototheca moriformis* (UTEX 1435).

	Aceite prensado (mcg/ml)	Aceite extraído del solvente (mcg/ml)
cis-Luteína	0.041	0.042
trans-Luteína	0.140	0.112
trans-Zeaxantina	0.045	0.039
cis-Zeaxantina	0.007	0.013
t-alfa-Critoxantina	0.007	0.010
t-beta-Critoxantina	0.009	0.010
t-alfa-Caroteno	0.003	0.001
c-alfa-Caroteno	ninguno detectado	ninguno detectado
t-beta-Caroteno	0.010	0.009
9-cis-beta-Caroteno	0.004	0.002
Licopeno	ninguno detectado	ninguno detectado
Total de Carotenoides	**0.267**	**0.238**
Clorofila	<0.01 mg/kg	<0.01 mg/kg
Tocoferoles y Esteroles		
	Aceite prensado (mg/100g)	Aceite extraido del solvente (mg/100g)
gamma Tocoferol	0.49	0.49
Campesterol	6.09	6.05
Estigmasterol	47.6	47.8
Beta-sitosterol	11.6	11.5
Otros esteroless	445	446
Tocotrienoles		
	Aceite prensado (mg/g)	Aceite extraido del solvente (mg/g)
alfa Tocotrienol	0.26	0.26
beta Tocotrienol	< 0.01	< 0.01
gamma Tocotrienol	0.10	0.10
detal Tocotrienol	< 0.01	< 0.01
Total de Tocotrienoles	**0.36**	**0.36**

La capacidad de usar sacarosa como única fuente de carbono como el factor de selección de los clones que contenían la construcción del transgén *suc2* en lugar del G418 (u otro antibiótico) se evaluó mediante el uso de transformantes positivo para el gen *suc2*. Un subconjunto de los transformantes positivos se cultivó durante 24 doblajes en placas que contenían sacarosa como única fuente de carbono y sin selección de antibiótico. Los clones entonces se retaron, con placas que contenían glucosa como única fuente de carbono y G418. Hubo un subconjunto de los clones que no creció con la condición de glucosa + G418, lo que indicó una pérdida de la expresión del transgén. Un experimento adicional se realizó mediante el uso de una placa que contenía sacarosa como única fuente de carbono y no G418 y se estrió un clon que expresaba el transgén *suc2* en la mitad de la placa y la *Prototheca moriformis* de tipo salvaje en la otra mitad de la placa. El crecimiento se vio tanto con el tipo salvaje como con las células de *Prototheca moriformis* que contenían el transgen. La *Prototheca moriformis* de tipo salvaje no demostró capacidad de crecer en sacarosa, por lo tanto, este resultado muestra que a diferencia de la resistencia a los antibióticos, el uso de la selección con sacarosa/invertasa no es por células autónomas. Es muy probable que los transformantes secretaran suficiente sacarosa invertasa en la placa/medio para apoyar el crecimiento del tipo salvaje mientras que la sacarosa se hidrolizara en fructosa y glucosa.

Ejemplo 4: *Prototheca* recombinantes con los genes TE exógenos

Como se describió anteriormente, las cepas de *Prototheca* se pueden transformar con genes exógenos. La *Prototheca moriformis* (UTEX 1435) se transformó, mediante el uso de los métodos que se describieron anteriormente, ya sea con el gene de la tioesterasa C12 de *Umbellularia californica* o con el gen de la tioesterasa C14 de *Cinnamomum camphora* (ambas con codones optimizados de acuerdo a la Tabla 1). Cada una de las construcciones de la transformación contenían un promotor de glutamato deshidrogenasa de *Chlorella sorokiniana*/región 5'UTR (SEQ ID NO: 69) para dirigir la expresión del transgén de la tioesterasa. Las regiones de codificación de los transgenes de las tioesterasas, la tioesterasa C12 de *Umbellularia californica* (SEQ ID NO: 70) o la tioesterasa C14 de *Cinnamomum camphora* (SEQ ID NO: 71), cada uno con la secuencia dirigida a plastidio putativa nativa. Inmediatamente después de la secuencia codificante de la tioesterasa está la secuencia codificante

para una etiqueta 3x FLAG en c-terminal (SEQ ID NO: 72), seguida de la 3'UTR de la nitrato reductasa de *Chlorella vulgaris* (SEQ ID NO: 73). Un diagrama de las construcciones de las tioesterasas que se usaron en las transformaciones de la *Prototheca moriformis* se muestra en la Figura 9.

La preparación del ADN, el microtransportador de oro, y las células *Prototheca moriformis* (UTEX 1435) se realizó mediante el uso de los métodos que se describieron anteriormente en el Ejemplo 3. Las microalgas se bombardearon con la mezcla de ADN con el microtransportador de oro y se sembraron en las placas de selección que contenían 2% de glucosa y 100µg/ml de G418. A las colonias se les permitió el desarrollo de 7 a 12 días y las colonias se recogieron de cada placa de transformación para la detección de la incorporación de la construcción de ADN mediante el uso de ensayos de transferencia del tipo Southern y la expresión de las construcciones de tioesterasa se detectó mediante el uso de la RT-PCR.

Los clones positivos se recogieron tanto de las placas de transformación de la tioesterasa C12 y C14 y se detectó la incorporación de la construcción mediante ensayos de transferencia de tipo Southern. Los ensayos de transferencia de tipo Southern se llevaron a cabo mediante el uso de métodos estándar (y se describieron anteriormente en el Ejemplo 3), mediante el uso de una sonda c optimizada, que se basó en la secuencia SEQ ID NO: 70 y SEQ ID NO: 71. El ADN del plásmido de transformación se corrió como un control positivo. Fuera de los clones que resultaron positivos para la incorporación de la construcción, un subgrupo se seleccionó para el análisis por PCR cuantitativa con transcripción inversa (RT-PCR cuantitativa) para la expresión de la tioesterasa C12 y la tioesterasa C14.

El aislamiento del ARN se realizó mediante el uso de los métodos que se describieron en el Ejemplo 3 anterior y se realizaron las RT-qPCR de los clones positivos mediante el uso de 20 ng de ARN total de cada clon, se usó el par de cebadores que se describen a continuación y el kit de RT-PCR iScript SYBR Green (Bio-Rad Laboratories) según el protocolo del fabricante. El ARN total de *Prototheca moriformis* de tipo salvaje (no transformado) se incluyó como control negativo. La expresión del ARNm se expresó como las veces de expresión relativa (RFE) en comparación con el control negativo. Los cebadores que se usaron en las detecciones RT-qPCR de a C12 de transformación de la tioesterasa fueron:

U. californica cebadores PCR de tioesterasa C12:

Directo: 5' CTGGGCGACGGCTTCGGCAC 3' (SEQ ID NO: 74)
Inverso: 5' AAGTCGCGGCGCATGCCGTT 3' (SEQ ID NO: 75)

Los cebadores que se usaron en la detección RT-qPCR de la transformación de tioesterasa C14 fueron: *Cinnamomum camphora* cebadores de tioesterasa C14:

Directo: 5' TACCCCGCCTGGGGCGACAC 3' (SEQ ID NO: 76)
Inverso: 5' CTTGCTCAGGCGGCGGGTGC 3' (SEQ ID NO: 77)

La RT-PCR cuantitativa para la expresión de la tioesterasa C12 resultó en los clones positivos que mostraron un aumento de RFE de aproximadamente 40 veces a más de 2000 veces mayor expresión en comparación con el control negativo. Se observaron resultados similares con la expresión de la tioesterasa C14 en los clones positivos con un aumento de RFE de aproximadamente 60 veces a más de 1200 veces mayor expresión en comparación con el control negativo.

Se seleccionó un subconjunto de los clones positivos de cada transformación (como los que se detectaron en los ensayos de transferencia de tipo Southern y RT-PCR cuantitativa) y se cultivaron en condiciones abundantes de nitrógeno y se analizó la producción y el perfil lipídico totales. Se prepararon muestras de los lípidos a partir de la biomasa seca de cada clon. Se resuspendieron 20-40 mg de biomasa seca de cada clon transgénico en 2 ml de H_2SO_4 3% en MeOH, y se añadieron 200 µl de tolueno que contenía una cantidad apropiada de un patrón interno adecuado (C19:0). La mezcla se sometió a sonicación brevemente para dispersar la biomasa, entonces, se calentó a 65-70°C durante dos horas. Se añadieron 2 ml de heptano para extraer los ésteres metílicos de ácidos grasos, seguido por la adición de 2 ml de K_2CO_3 (ac) 6% para neutralizar el ácido. La mezcla se agitó vigorosamente, y una parte de la capa superior se transfirió a un frasco que contenía Na_2SO_4 (anhidro) para el análisis por cromatografía gaseosa mediante el uso del método estándar FAME GC/FID (cromatografía gaseosa de éster metílico de ácidos grasos con detección de ionización de llama). El perfil lipídico (se expresó como % del área) de los clones positivos en comparación con el control negativo de tipo salvaje se resume en las Tablas 15 y 16 a continuación. Como se muestra en la Tabla 15, las veces de incremento de la producción de C12 en los transformantes C12 estuvo en el intervalo de aproximadamente un aumento de 5 veces (clon C12-5) a un aumento de más de 11 veces (clon C12-1). Las veces de incremento de la producción de C14 en los transformantes C14 estuvo en el intervalo de aproximadamente un aumento de 1.5 veces a aproximadamente un aumento de 2.5 veces.

Tabla 15. Sumario del perfil lipídico total de transformantes de tioesterasa C12*Prototheca moriformis*.

	Tipo salvaje	C12-1	C12-2	C12-3	C12-4	C12-5	C12-6	C12-7	C12-8
C6:0	0.03	ND	ND	ND	ND	ND	ND	ND	ND
C8:0	0.11	0.09	ND	0.11	ND	ND	ND	ND	ND
C10:0	ND	ND	ND	0.01	0.01	ND	ND	0.01	ND
C12:0	0.09	1.04	0.27	0.72	0.71	0.50	0.67	0.61	0.92
C14:0	2.77	2.68	2.84	2.68	2.65	2.79	2.73	2.56	2.69
C14:1	0.01	ND	ND	0.02	ND	ND	ND	0.01	ND
C15:0	0.30	0.09	0.10	0.54	0.19	0.09	0.13	0.97	0.09
C15:1	0.05	ND	ND	0.02	ND	ND	ND	ND	ND
C16:0	24.13	23.12	24.06	22.91	22.85	23.61	23.14	21.90	23.18
C16:1	0.57	0.62	0.10	0.52	0.69	0.63	0.69	0.49	0.63
C17:0	0.47	0.24	0.27	1.02	0.36	0.17	0.26	2.21	0.19
C17:1	0.08	ND	0.09	0.27	0.10	0.05	0.09	0.80	0.05
C18:0	ND	ND	2.14	1.75	2.23	2.16	2.38	1.62	2.47
C18:1	22.10	23.15	24.61	21.90	23.52	19.30	22.95	20.22	22.85
C18:1	ND	0.33	0.24	ND	ND	0.09	0.09	ND	0.11
C18:2	37.16	34.71	35.29	35.44	35.24	36.29	35.54	36.01	35.31
C18:3 alfa	11.68	11.29	9.26	11.62	10.76	13.61	10.64	11.97	10.81
C20:0	0.15	0.16	0.19	0.16	0.16	0.14	0.18	0.14	0.18
C20:1	0.22	0.17	0.19	0.20	0.21	0.19	0.21	0.20	0.21
C20:2	0.05	ND	0.04	0.05	0.05	0.05	0.04	0.05	0.04
C22:0	ND	ND	ND	0.01	ND	ND	ND	0.02	ND
C22:1	ND	ND	ND	ND	ND	0.01	ND	0.01	ND
C20:3	0.05	ND	0.07	0.06	0.06	0.10	0.07	0.05	0.06
C20:4	ND	ND	ND	ND	ND	0.02	ND	ND	ND
C24:0	ND	ND	0.24	0.01	0.20	0.19	0.19	0.14	0.20

Tabla 16. Sumario del perfil lipídico total de transformantes de tioesterasa *Prototheca moriformis*.

	Tipo salvaje	C14-1	C14-2	C14-3	C14-4	C14-5	C14-6	C14-7
C6:0	0.03	ND	ND	ND	ND	ND	ND	ND
C8:0	0.11	ND	ND	ND	ND	ND	ND	ND
C10:0	ND	0.01	ND	0.01	ND	0.01	ND	ND
C12:0	0.09	0.20	0.16	0.25	0.21	0.19	0.40	0.17
C14:0	2.77	4.31	4.76	4.94	4.66	4.30	6.75	4.02
C14:1	0.01	ND	0.01	ND	ND	0.01	ND	ND
C15:0	0.30	0.43	0.45	0.12	0.09	0.67	0.10	0.33
C15:1	0.05	ND	ND	ND	ND	ND	ND	ND
C16:0	24.13	22.85	23.20	23.83	23.84	23.48	24.04	23.34
C16:1	0.57	0.65	0.61	0.60	0.60	0.47	0.56	0.67
C17:0	0.47	0.77	0.76	0.21	0.19	1.11	0.18	0.54
C17:1	0.08	0.23	0.15	0.06	0.05	0.24	0.05	0.12
C18:0	ND	1.96	1.46	2.48	2.34	1.84	2.50	2.06
C18:1	22.10	22.25	19.92	22.36	20.57	19.50	20.63	22.03
C18:1	ND	ND	ND	ND	ND	ND	0.10	ND
C18:2	37.16	34.97	36.11	34.35	35.70	35.49	34.03	35.60
C18:3 alfa	11.68	10.71	12.00	10.15	11.03	12.08	9.98	10.47
C20:0	0.15	0.16	0.19	0.17	0.17	0.14	0.18	0.16
C20:1	0.22	0.20	0.12	.019	0.19	0.19	0.17	0.20
C20:2	0.05	0.04	0.02	0.03	0.04	0.05	0.03	0.04
C22:0	ND	ND	ND	ND	0.02	0.01	ND	ND
C22:1	ND	0.01	ND	ND	ND	ND	ND	0.01
C20:3	0.05	0.08	0.03	0.06	0.09	0.05	0.05	0.07
C20:4	ND	0.01	ND	ND	ND	ND	0.02	ND
C24:0	ND	0.17	0.14	0.19	0.20	0.16	0.22	0.17

Los experimentos que se describieron indican la transformación exitosa de *Prototheca moriformis* (UTEX 1435) con las construcciones de los transgenes de dos tioesterasas diferentes (C12 y C14), que implicaron no sólo la expresión exitosa del transgén, sino también la correcta dirección de la proteína expresada a plastidios y un efecto funcional (el cambio esperado en el perfil lipídico) como resultado de la transformación. El mismo experimento de transformación se realizó mediante el uso de una construcción de expresión que contenía (según la Tabla 1) la región de codificación con codones optimizados de la tioesterasa C8-10 de *Cuphea hookeriana* con la secuencia dirigida a los plastidio, nativa, (SEQ ID NO: 78) que no produjo ningún cambio en el perfil de lípidos. Si bien la introducción del transgén C8-10 de *Cuphea hookeriana* en *Prototheca moriformis* (UTEX 1435) fue un éxito y se confirmó mediante análisis de transferencia de tipo Southern, ningún cambio en la producción de ácidos grasos C10 o C8 se detectó en los transformantes en comparación con la cepa de tipo salvaje.

Ejemplo 5: **Generación de la cepa de *Prototheca moriformis* con TE de plantas exógenas con secuencias dirigidas a plastidios de algas**

Con el fin de investigar si el uso de las secuencias dirigidas a los cloroplastos/plastidios de algas mejorarían la expresión de la tioesterasa de cadena mediana (C8-C14) y la posterior producción de lípidos de cadena mediana de *Prototheca moriformis* (UTEX 1435), varias secuencias dirigidas a plastidios de algas putativas, se clonaron a partir de *Chlorella protothecoides* y *Prototheca moriformis*. Las construcciones de tioesterasa que se basaban en la tioesterasa C8-10 de *Cuphea hookeriana*, la tioesterasa C12 de *Umbellularia californica* y la tioesterasa C14 de *Cinnamomum camphora* se hicieron mediante el uso del promotor de la glutamato deshidrogenasa de *Chlorella sorokiniana*/5'UTR y una nitrato reductasa de *Chlorella vulgaris* 3'UTR. Las secuencias codificantes de la tioesterasa se modificaron mediante la eliminación de las secuencias dirigidas a plastidios, nativas y su reemplazo con secuencias dirigidas a plastidios de los genomas de *Chlorella protothecoides* y de *Prototheca moriformis*. Las construcciones de expresión de tioesterasa y sus correspondientes números de identificación de secuencia se enumeran a continuación. Cada plásmido de transformación contenía también una construcción de resistencia a Neo que era idéntico a los que se describieron en el Ejemplo 3 anterior. Además, otro promotor derivado de algas, el promotor de la β-tubulina de *Chlamydomonas reinhardtii*, se probó también conjuntamente con las construcciones de tioesterasa. La secuencia dirigida al plastido, "nativa", se refiere a la secuencia dirigida a plastidio de tioesterasa de plantas superiores. Un resumen de las construcciones usadas en estos experimentos se proporciona a continuación:

Construcción Nombre	Promotor/ 5'UTR	sec. dirigida a plastidio	Gen	3'UTR	SEQ ID NO.
Construcción 1	*C. sorokiniana* glutamato deshidrogenasa	*C.prototheco ides* estearoil ACP desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 79
Construcción 2	*C. sorokiniana* glutamato deshidrogenasa	*P. moriformis* delta 12 graso ácido desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO:80
Construcción 3	*C. sorokiniana* glutamato deshidrogenasa	*P. moriformis* isopentenil difosfato sintasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 81
Construcción 4	*C. sorokiniana* glutamato deshidrogenasa	*P. moriformis* isopentenil difosfato sintasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 82
Construcción 5	*C. sorokiniana* glutamato deshidrogenasa	*P. moriformis* estearoil ACP desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 83
Construcción 6	*C. sorokiniana* glutamato deshidrogenasa	*C.prototheco ides* estearoil ACP desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 84
Construcción 7	*C. sorokiniana* glutamato deshidrogenasa	*P. moriformis* delta 12 graso ácido desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 85
Construcción 8	*C. sorokiniana* glutamato deshidrogenasa	*C.prototheco ides* estearoil ACP desaturasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 86

Construcción 9	*Chlamydomonas reinhardtii* β-tubulina	Native	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO:113
Construcción 10	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* isopentenil difosfato sintasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 114
Construcción 11	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* delta 12 graso ácido desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 115
Construcción 12	*Chlamydomonas reinhardtii* β-tubulina	*C.prototheco ides* estearoil ACP desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 116
Construcción 13	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* estearoil ACP desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 117
Construcción 14	*Chlamydomonas reinhardtii* β-tubulina	Nativa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 118
Construcción 15	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* isopentenil difosfato	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 119
Construcción 16	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* delta 12 graso ácido desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 120
Construcción 17	*Chlamydomonas reinhardtii* β-tubulina	*C.prototheco ides* estearoil ACP desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 121
Construcción 18	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* estearoil ACP desaturasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 122
Construcción 19	*Chlamydomonas reinhardtii* β-tubulina	Nativa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 123
Construcción 20	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* isopentenil difosfato sintasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 124
Construcción 21	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* delta 12 graso ácido desaturasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO:
Construcción 22	*Chlamydomonas reinhardtii* β-tubulina	*C.prototheco ides* estearoil ACP desaturasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 87
Construcción 23	*Chlamydomonas reinhardtii* β-tubulina	*P. moriformis* estearoil ACP desaturasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 88

Cada construcción se transformó en *Prototheca moriformis* (UTEX 1435) y la selección se realizó con G418 mediante el uso de los métodos que se describieron en el Ejemplo 4 anterior. Varios clones positivos de cada

transformación se recogieron y se examinaron en busca de la presencia de transgenes de la tioesterasa mediante el uso del análisis de transferencia de tipo Southern. La expresión del transgén de la tioesterasa se confirmó mediante la RT-PCR. Un subconjunto de los clones positivos (según se confirmó por el análisis de transferencia de tipo Southern y la RT-PCR) de cada transformación se seleccionó y se cultivó para el análisis del perfil lipídico. Las muestras de los lípidos se prepararon a partir de muestras de la biomasa seca de cada clon y el análisis del perfil lipídico se realizó mediante el uso de los métodos de hidrólisis ácida que se describieron en el Ejemplo 4. Los cambios en el porcentaje del área de los ácidos grasos correspondientes al transgén de la tioesterasa se compararon con los niveles del tipo salvaje, y con los clones transformados con una tioesterasa con la secuencia dirigida a plastidio, nativa.

Como se mencionó en el Ejemplo 4, los clones transformados con la construcción de la tioesterasa C8-10 de *Cuphea hookeriana* con la secuencia dirigida a plastidio, nativa, tenían el mismo nivel de ácidos grasos C8 y C10 que el tipo salvaje. Los clones transformados con la construcción de la tioesterasa C8-10 de *Cuphea hookeriana* (Construcción 1-3) con secuencias dirigidas a plastidios de algas tenían un aumento de más de 10 veces en los ácidos grasos C10 para la Construcción 3 y un aumento de mas de 40 veces en los ácidos grasos C10 de las Construcciones 1 y 2 (en comparación con el tipo salvaje). Los clones transformados con las construcciones de la tioesterasa C12 de *Umbellularia californica* con la secuencia dirigida a los plastidio nativa, tuvo un aumento moderado de 6-8 veces en los niveles de ácidos grasos C12 en comparación con el tipo salvaje. Los clones transformados con las construcciones de la tioesterasa C12 de *Umbellularia californica* con las construcciones dirigidas a plastidios de algas (Construcciones 4-7) tuvieron un aumento de más de 80 veces en el nivel del ácido graso C12 para la Construcción 4, aproximadamente un aumento de 20 veces en el nivel de ácido graso C12 para la Construcción 6, aproximadamente un aumento de 10 veces en el nivel de ácido graso C12 para la Construcción 7 y aproximadamente un aumento de 3 veces en el nivel de ácido graso C12 para la Construcción 5 (todos en comparación con tipo salvaje). Los clones transformados con la tioesterasa C14 de *Cinnamomum camphora*, ya sea con la secuencia dirigida a plastidio nativa o la Construcción 8 (con la secuencia dirigida a plastidio de la estearoil ACP desaturasa de *Chlorella protothecoides*) tuvieron un aumento de 2-3 veces en los niveles de ácidos grasos C14 en comparación con el tipo salvaje. En general, los clones transformados con las construcciones de la tioesterasa con una secuencia dirigida a los plastidio de algas tuvieron un aumento de veces más alto en los correspondientes niveles de longitud de cadena de los ácidos grasos, que cuando se usó la secuencia dirigida nativa de plantas superiores.

A. *El promotor de la β-tubulina de* Clamydomonas reinhartii

La expresión adicional de las construcciones de la tioesterasa heteróloga se elaboraron mediante el uso del promotor de la β-tubulina de *Chlamydomonas reinhardtii* en lugar del promotor de la glutamato deshidrogenasa de *C. sorokinana*. Los elementos de la construcción y la secuencia de las construcciones de expresión se enumeraron anteriormente. Cada construcción se transformó en las células hospederas *Prototheca moriformis* UTEX 1435 mediante el uso de los métodos que se describieron anteriormente. Los perfiles de lípidos se generaron a partir de un subconjunto de los clones positivos para cada construcción con el fin de evaluar el éxito y la productividad de cada construcción. Los perfiles de lípidos comparan los niveles de ácidos grasos (expresados en % del área) con las células hospederas de tipo salvaje. La columna "media" representa el promedio numérico del subconjunto de los clones positivos. La columna "muestra" representa el mejor clon positivo que se examinó (el mejor se definió como la muestra que produce el mayor cambio en % del área de la producción del ácido graso de la correspondiente longitud de cadena). La columna "bajo-alto" representa el mas bajo % de área y el más alto % área de los ácidos grasos de los clones que se examinaron. Los perfiles de lípidos resultantes de las Construcciones 9-23 se resumen a continuación.

Construcción 9. *Cuphea hookeriana* C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.05	0.30	0-0.29
C 10:0	0.01	0.63	2.19	0-2.19
C 12:0	0.03	0.06	0.10	0-0.10
C 14:0	1.40	1.50	1.41	1.36-3.59
C 16:0	24.01	24.96	24.20	
C 16:1	0.67	0.80	0.85	
C 17:0	0	0.16	0.16	
C 17:1	0	0.91	0	
C 18:0	4.15	17.52	3.19	
C 18:1	55.83	44.81	57.54	
C 18:2	10.14	7.58	8.83	
C 18:3α	0.93	0.68	0.76	
C 20:0	0.33	0.21	0.29	
C 24:0	0	0.05	0.11	

Construcción 10. *Cuphea hookeriana* C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.01	0.02	0-0.03
C 10:0	0	0.16	0.35	0-0.35
C 12:0	0.04	0.05	0.07	0-0.07
C 14:0	1.13	1.62	1.81	0-0.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.05	0.05	
C 16:0	19.94	26.42	28.08	
C 16:1	0.84	0.96	0.96	
C 17:0	0.19	0.14	0.13	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.62	3.43	
C 18:1	63.96	54.90	53.91	
C 18:2	9.62	9.83	9.11	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.79	0.73	
C 20:0	0.26	0.35	0.33	
C 20:1	0.06	0.08	0.09	
C 20:1	0.08	0.06	0.07	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.13	0.11	

Construcción 11. *Cuphea hookeriana* C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.82	1.57	0-1.87
C 10:0	0	3.86	6.76	0-6.76
C 12:0	0.04	0.13	0.20	0.03-0.20
C 14:0	1.13	1.80	1.98	1.64-2.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	25.60	25.44	
C 16:1	0.84	1.01	1.02	
C 17:0	0.19	0.13	0.11	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	2.98	2.38	
C 18:1	63.96	51.59	48.85	
C 18:2	9.62	9.85	9.62	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.91	0.92	
C 20:0	0.26	0.29	0.26	
C 20:1	0.06	0.06	0	
C 20:1	0.08	0.06	0.03	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.06	0	

Construcción 12. *Cuphea hookeriana* C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.31	0.85	0-0.85
C 10:0	0	2.16	4.35	0.20-4.35
C 12:0	0.04	0.10	0.15	0-0.18
C 14:0	1.13	1.96	1.82	1.66-2.97
C 14:1	0	0.03	0.04	
C 15:0	0.06	0.07	0.07	
C 16:0	19.94	26.08	25.00	
C 16:1	0.84	1.04	0.88	
C 17:0	0.19	0.16	0.16	
C 17:1	0.10	0.05	0.07	
C 18:0	2.68	3.02	3.19	
C 18:1	63.96	51.08	52.15	
C 18:2	9.62	11.44	9.47	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.98	0.90	
C 20:0	0.26	0.30	0.28	
C 20:1	0.06	0.06	0.05	
C 20:1	0.08	0.04	0	
C 22:0	0	0.07	0	
C 24:0	0.13	0.05	0	

Construción14. *Umbellularia californica* C12 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.02	0.03	0.02-0.03
C 12:0	0.03	2.62	3.91	0.04-3.91
C 14:0	1.40	1.99	2.11	1.83-2.19
C 16:0	24.01	27.64	27.01	
C 16:1	0.67	0.92	0.92	
C 18:0	4.15	2.99	2.87	
C 18:1	55.83	53.22	52.89	
C 18:2	10.14	8.68	8.41	
C 18:3α	0.93	0.78	0.74	
C 20:0	0.33	0.29	0.27	

Construcción 15. *Umbellularia californica* C12 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.05	0.08	0-0.08
C 12:0	0.04	8.12	12.80	4.35-12.80
C 13:0	0	0.02	0.03	0-0.03
C 14:0	1.13	2.67	3.02	2.18-3.37
C 14:1	0	0.04	0.03	0.03-0.10
C 15:0	0.06	0.07	0.06	
C 16:0	19.94	25.26	23.15	
C 16:1	0.84	0.99	0.86	
C 17:0	0.19	0.14	0.14	
C 17:1	0.10	0.05	0.05	
C 18:0	2.68	2.59	2.84	
C 18:1	63.96	46.91	44.93	
C 18:2	9.62	10.59	10.01	
C 18:3α	0.63	0.92	0.83	
C 20:0	0.26	0.27	0.24	
C 20:1	0.06	0.06	0.06	
C 20:1	0.08	0.05	0.04	
C 22:0	0	0.07	0.09	
C 24:0	0.13	0.13	0.12	

Construcción 16. *Umbellularia californica* C12 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.03	0.04	0.02-0.04
C 12:0	0.04	2.43	5.32	0.98-5.32
C 13:0	0	0.01	0.02	0-0.02
C 14:0	1.13	1.77	1.93	1.62-1.93
C 14:1	0	0.03	0.02	0.02-0.04
C 15:0	0.06	0.06	0.05	
C 16:0	19.94	24.89	22.29	
C 16:1	0.84	0.91	0.82	
C 17:0	0.19	0.16	0.15	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.81	3.67	
C 18:1	63.96	53.19	52.82	
C 18:2	9.62	10.38	10.57	
C 18:3α	0.63	0.80	0.77	
C 20:0	0.26	0.35	0.32	
C 20:1	0.06	0.06	0.07	
C 20:1	0.08	0.07	0.08	
C 22:0	0	0.08	0.07	
C 24:0	0.13	0.15	0.14	

Construcción 17. *Umbellularia californica* C12 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.04	0.07	0.03-0.08
C 12:0	0.04	7.02	14.11	4.32-14.11
C 13:0	0	0.03	0.04	0.01-0.04
C 14:0	1.13	2.25	3.01	1.95-3.01
C 14:1	0	0.03	0.03	0.02-0.03
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	23.20	21.46	
C 16:1	0.84	0.82	0.77	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.47	2.93	
C 18:1	63.96	50.30	45.17	
C 18:2	9.62	10.33	9.98	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.84	0.86	
C 20:0	0.26	0.32	0.27	
C 20:1	0.06	0.07	0.06	
C 20:1	0.08	0.06	0.06	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.14	0.13	

Construcción 18. *Umbellularia californica* C12 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.03	0.05	0.01-0.05
C 12:0	0.04	5.06	7.77	0.37-7.77
C 13:0	0	0.02	0	0-0.03
C 14:0	1.13	2.11	2.39	1.82-2.39
C 14:1	0	0.03	0.03	0.02-0.05
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	24.60	23.95	
C 16:1	0.84	0.86	0.83	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.31	2.96	
C 18:1	63.96	51.26	49.70	
C 18:2	9.62	10.18	10.02	
C18:3 γ	0	0.01	0.02	
C 18:3α	0.63	0.86	0.86	
C 20:0	0.26	0.32	0.29	
C 20:1	0.06	0.05	0.05	
C 20:1	0.08	0.07	0.04	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.13	0.13	

Construcción 19. *Cinnamomum camphora* C14 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.02	0.01	0.01	0.01-0.02
C 12:0	0.05	0.27	0.40	0.08-0.41
C 14:0	1.52	4.47	5.81	2.10-5.81
C 16:0	25.16	28.14	28.55	
C 16:1	0.72	0.84	0.82	
C 18:0	3.70	3.17	2.87	
C 18:1	54.28	51.89	51.01	
C 18:2	12.24	9.36	8.62	
C 18:3α	0.87	0.74	0.75	
C 20:0	0.33	0.33	0.31	

Construcción 20. *Cinnamomum camphora* C14 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.01	0.02	0.01-0.02
C 12:0	0.03	0.39	0.65	0.08-0.65
C 13:0	0	0.01	0.01	0.01-0.02
C 14:0	1.40	5.61	8.4	2.1-8.4
C 14:1	0	0.03	0.03	0.02-0.03
C 15:0	0	0.06	0.07	
C 16:0	24.01	25.93	25.57	
C 16:1	0.67	0.75	0.71	
C 17:0	0	0.13	0.12	
C 17:1	0	0.05	0.05	
C 18:0	4.15	3.30	3.23	
C 18:1	55.83	51.00	48.48	
C 18:2	10.14	10.38	10.35	
C 18:3α	0.93	0.91	0.88	
C 20:0	0.33	0.35	0.32	
C 20:1	0	0.08	0.08	
C 20:1	0	0.07	0.07	
C 22:0	0	0.08	0.08	
C 24:0	0	0.14	0.13	

Construcción 21. *Cinnamomum camphora* C14 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.01	0.01	0-0.01
C 12:0	0.03	0.10	0.27	0.04-0.27
C 14:0	1.40	2.28	4.40	1.47-4.40
C 16:0	24.01	26.10	26.38	
C 16:1	0.67	0.79	0.73	
C 17:0	0	0.15	0.16	
C 17:1	0	0.06	0.06	
C 18:0	4.15	3.59	3.51	
C 18:1	55.83	53.53	50.86	
C 18:2	10.14	10.83	11.11	
C 18:3α	0.93	0.97	0.87	
C 20:0	0.33	0.36	0.37	
C 20:1	0	0.09	0.08	
C 20:1	0	0.07	0.07	
C 22:0	0	0.09	0.09	

Construcción 22. *Cinnamomum camphora* C14 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.02	0.02	0.02-0.02
C 12:0	0.03	1.22	1.83	0.59-1.83
C 13:0	0	0.02	0.03	0.01-0.03
C 14:0	1.40	12.77	17.33	7.97-17.33
C 14:1	0	0.02	0.02	0.02-0.04
C 15:0	0	0.07	0.08	
C 16:0	24.01	24.79	24.22	
C 16:1	0.67	0.64	0.58	
C 17:0	0	0.11	0.10	
C 17:1	0	0.04	0.04	
C 18:0	4.15	2.85	2.75	
C 18:1	55.83	45.16	41.23	
C 18:2	10.14	9.96	9.65	
C 18:3α	0.93	0.91	0.85	
C 20:0	0.33	0.30	0.30	
C 20:1	0	0.07	0.06	
C 20:1	0	0.06	0.05	
C 22:0	0	0.08	0.08	

Construcción 23. *Cinnamomum camphora* C14 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.01	0.02	0-0.02
C 12:0	0.05	0.57	1.08	0.16-1.08
C 13:0	0	0.02	0.02	0-0.02
C 14:0	1.45	7.18	11.24	2.96-11.24
C 14:1	0.02	0.03	0.03	0.02-0.03
C 15:0	0.06	0.07	0.07	
C 16:0	24.13	25.78	25.21	
C 16:1	0.77	0.72	0.66	
C 17:0	0.19	0.13	0.11	
C 17:1	0.08	0.05	0.04	
C 18:0	3.53	3.35	3.12	
C 18:1	56.15	49.65	46.35	
C 18:2	11.26	10.17	9.72	
C 18:3α	0.84	0.95	0.83	
C 20:0	0.32	0.34	0.32	
C 20:1	0.09	0.08	0.09	
C 20:1	0.07	0.05	0.06	
C 22:0	0.07	0.08	0.08	
C 24:0	0.13	0.13	0.12	

Las Construcciones 9-13 eran vectores de expresión que contenían la construcción de la tioesterasa C8-10 de *Cuphea hookeriana*. Como se puede observar en los resúmenes de los datos anteriores, los mejores resultados se vieron con la Construcción 11, con la muestra de ácido graso C8 que fue 1.57 % del área (en comparación con el 0 en el tipo salvaje) y el ácido graso C10 que fue 6.76 % del área (en comparación con el 0 en el tipo salvaje). También hubo un aumento moderado de los ácidos grasos C12 (un aumento de aproximadamente 2-5 veces). Si bien la secuencia dirigida a los plastidio nativa, no produjo cambio bajo el control del promotor de la glutamato deshidrogenasa de *C.sorokinana*, la misma construcción de expresión conducida por el promotor de la β-tubulina de *C.reinhardtii* produjo cambios significativos en los ácidos grasos C8-10 en la célula hospedera. Esta es una prueba más de la idiosincrasia de la expresión heteróloga de tioesterasas en especies de *Prototheca*. Todos los clones que contenían la construcción de la tioesterasa C8-10 con el promotor de la β-tubulina de *C.reinhardtii* tuvieron mayores aumentos en los ácidos grasos C8-10 que los clones que contenían la construcción de la tioesterasa C8-10 con el promotor de la glutamato deshidrogenasa de *C.sorokinana*. Los datos del perfil lipídico para la Construcción 13 no se obtuvieron y, por tanto, no se incluyeron anteriormente.

Las Construcciones 14-18 eran vectores de expresión que contenían la construcción de la tioesterasa C12 de *Umbellularia californica*. Como se puede observar en los resúmenes de los datos anteriores, los mejores resultados se observarn con la Construccion 15 (la secuencia dirigida a plastidio de la isopentenil difosfato sintasa de *P.moriformis*) y la 17 (la secuencia dirigida a plastidio de la estearoil ACP desaturasa de *C.prototheco*ides). El mayor cambio en la producción del ácido graso C12 se vio con la Construcción 17, con niveles de ácidos grasos C12 de hasta 14.11% del área, en comparación con el 0.04% del área en el tipo salvaje. También se observaron cambios moderados (aproximadamente de 2 veces) con los niveles de ácidos grasos C14. En comparación con la misma construcción con el promotor de la glutamato deshidrogenasa de la *C.sorokinana*, las mismas tendencias fueron ciertas con el promotor de la β-tubulina de *C.reinhardtii* – las secuencias dirigidas a plastidios de, la estearoil ACP desaturasa de *C.prototheco*ides y la isopentenil difosfato sintasa de *P.moriformis* produjeron el mayor cambio en los niveles de ácidos grasos C12 con ambos promotores.

Las Construcciones 19-23 eran vectores de expresión que contenían la construcción de la tioesterasa C14 de *Cinnamomum camphora*. Como se puede observar en los resúmenes de los datos anteriores, los mejores resultados se vieron con la Construcción 22 y la Construcción 23. El mayor cambio en la producción del ácido graso C14 se vio con la Construcción 22 y con niveles del ácido graso C14 de hasta el 17.35 % del área (cuando los valores de C140 y C141 se combinan), en comparación con el 1.40 % en el tipo salvaje. Los cambios en los ácidos grasos C12 también se observaron (5-60 veces). En comparación con la misma construcción con el promotor de la glutamato deshidrogenasa de *C.sorokinana*, las mismas tendencias fueron ciertas con el promotor de la β-tubulina – las secuencias dirigidas a plastidios de la estearoil ACP desaturasa de *C.prototheco*ides y la estearoil ACP desaturasa de *P.moriformis* produjeron el mayor cambio en los niveles de ácidos grasos C14 con ambos promotores. En consonancia con todas las construcciones de expresión de tioesterasa, la construcción con el promotor de la β-tubulina de *C.reinhardtii* produjo los mayores cambios en los niveles de ácidos grasos C8-14 que el de la glutamato

deshidrogenasa de *C.sorokiniana.*

Dos clones positivos de la Construcción de 22 se seleccionaron y se cultivaron bajo alta presión selectiva (50 mg/l de G418). Después de 6 días de cultivo, los clones se cosecharon y se determinó el perfil lipídico mediante el uso de los métodos que se describieron anteriormente. Los datos del perfil lipídico se resumen a continuación y se expresaron en % del área.

Construcción 22 clones + 50mg/l G418		
Ácido graso	Construcción 22 A	Construcción 22 B
C 12:0	3.21	3.37
C 14:0	27.55	26.99
C 16:0	25.68	24.37
C 16:1	0.99	0.92
C 18:0	1.37	1.23
C 18:1	28.35	31.07
C 18:2	11.73	11.05
C 18:3α	0.92	0.81
C 20:0	0.16	0.17

Ambos clones, cuando se cultivaron a presión selectiva alta, constante, produjeron un aumento en la cantidad de los ácidos grasos C14 y C12, aproximadamente el doble de los niveles que se observaron con la Construcción 22 anterior. Estos clones produjeron mas de 30% del área de los ácidos grasos C12-14, en comparación con el 1.5 % del área de los ácidos grasos C12-14 que se vio en las células de tipo salvaje.

Ejemplo 6: Expresión heteróloga de la tioesterasa de *Cuphea palustris* y de *Ulmus americanca* en *Prototheca*

Dado el éxito de la expresión heteróloga de las tioesterasas en las especies de *Prototheca*, que se describió anteriormente, se construyeron y se describen a continuación casetes de expresión que contenían las secuencias que codifican las acil graso ACP tioesterasas con codones optimizados (según la Tabla 1) de *Cuphea palustris* y *Ulmus americana*.

Construcción Nombre	Promotor/ 5'UTR	Sec. dirigida a Plastidio	Gen	3'UTR	SEQ ID NO.
Construcción 27	*C.reinhardtii* β-tubulina	*C.prototheco*ides estearoil ACP desaturasa	*Cuphea palustris* tioesterasa	*C. vulgaris* nitrato reductasa	SEQ ID NO: 107

La *Ulmus americana* (secuencia codificante con codón optimizado) se puede insertar en el casete de expresión. La secuencia codificante con codones optimizados sin la secuencia dirigida a los plastidio nativa, para la tioesterasa de *Ulmus americana* se muestra como SEQ ID NO: 108 y se puede fundir cualquier secuencia dirigida a plastidio que se desee y el elemento de expresión (es decir, el promotor/5'UTR y 3'UTR)

Estos casetes de expresión se pueden transformar en las especies de *Prototheca* mediante el uso de los métodos que se describieron anteriormente. Los clones positivos se pueden detectar con la inclusión de un gen de resistencia a los antibióticos (por ejemplo, el neoR) en la construcción de expresión y detectarlo en las placas/medio que contienen el G418. Los clones positivos se pueden confirmar mediante uso de los ensayos de transferencia de tipo Southern con sondas específicas para la región que codifica la tioesterasa heteróloga y la construcción de expresión también se puede confirmar mediante el uso de la RT-PCR y cebadores específicos para la región que codifica la tioesterasa heteróloga. La confirmación secundaria de los clones positivos se puede lograr mediante la búsqueda de cambios en los niveles de ácidos grasos en el perfil lipídico de la célula hospedera. Como se vio en los ejemplos anteriores, la expresión de la tioesterasa heteróloga en las especies de *Prototheca* puede ser idiosincrásica a la tioesterasa particular. Los elementos promotor y secuencias dirigidas a plastidios (y otros elementos de regulación de la expresión) se pueden intercambiar hasta que la expresión de la tioesterasa (y el consiguiente aumento en el ácido graso correspondiente) alcance el nivel deseado.

Ejemplo 7: Transformantes dobles - expresión simultánea de dos proteínas heterólogas

La cepa de microalgas *Prototheca moriformis* (UTEX 1435) se transformó mediante el uso de los métodos que se describieron anteriormente con una construcción de expresión que contenía el gen *suc2* de la sacarosa invertasa de levadura que codifica la forma secretada de la invertasa de *S. cerevisiae*. La expresión exitosa de este gen y la orientación al periplasma resultó en la capacidad de la célula hospedera para crecer (y utilizar) la sacarosa como única fuente de carbono en condiciones heterótrofas (como se demostró en el Ejemplo 3). El segundo conjunto de genes que se expresaron son las tioesterasas que son responsables de la escisión del grupo acil de la proteína

transportadora de acil. Específicamente, las tioesterasas de *Cuphea hookeriana* (preferentemente una tioesterasa C8-10), de *Umbellularia californica* (preferentemente una tioesterasa C12) y de *Cinnamomum camphora* (preferentemente una tioesterasa C14). Estos casetes de expresión de tioesterasas se clonaron como fusiones con las secuencias dirigidas a plastidios de microalgas N-terminales, ya sea de *Prototheca moriformis* o de *Chlorella protothecoides*, que se demostró (en los ejemplos anteriores) ser más óptima que las secuencias dirigidas a plastidios, nativas de plantas superiores. La expresión exitosa de los genes de la tioesterasa y la orientación a plastidios resultó en cambios medibles en el perfil de ácidos grasos en la célula hospedera. Estos cambios en los perfiles son consistentes con la especificidad enzimática o preferencia de cada tioesterasa. A continuación se muestra un resumen de construcciones de doble expresión que se ensamblaron y se transformaron en *Prototheca moriformis* (UTEX 1435). Cada construcción contenía el gen *suc2* de levadura bajo el control del promotor de la β-tubulina de *C .reinhardtii* /5'UTR y contenía la nitrato reductasa 3'UTR de *C. vulgaris* y una tioesterasa de planta superior con una secuencia dirigida a plastidio de microalgas en sustitución de la secuencia nativa bajo el control de la glutamato deshidrogenasa 5'UTR de *C. sorokinana* y contenía la nitrato reductasa 3'UTR de *C. vulgaris*. A continuación se muestra un resumen de la parte de la tioesterasa de las construcciones que se ensamblaron y se transformaron en *Prototheca moriformis* (UTEX 1435). El casete entero de doble expresión con el gen *suc2* y el gen de la tioesterasa se enumera en la lista de identificación de secuencias.

Construcción Nombre	Promotor/ 5'UTR	sec. dirigida a plastidio	Gen	3'UTR	SEQ ID NO.
Construcción 24	*C. sorokiniana* glutamato deshidrogenasa	*C.prototheco ides* estearoil ACP desaturasa	*Cuphea hookeriana* C8-10 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 109
Construcción 25	*C. sorokinana* glutamato deshidrogenasa	*P. moriformis* isopentenil difosfato sintasa	*Umbellularia californica* C12 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 110
Construcción 26	*C. sorokinana* glutamato deshidrogenasa	*C.prototheco ides* estearoil ACP desaturasa	*Cinnamomum camphora* C14 TE	*C. vulgaris* nitrato reductasa	SEQ ID NO: 111

Las construcciones similares de doble expresión con los casetes de la tioesterasa que se describieron en el Ejemplo 5 (por ejemplo, bajo el control de un promotor diferente, tal como el promotor de la β-tubulina/5'UTR de *C. reinhardtii*) también se pueden generar mediante el uso de métodos estándar de biología molecular y los métodos que describen en la presente.

Los clones positivos que contienen cada una de las construcciones de expresión se examinaron mediante el uso de su capacidad de crecer en placas que contenían sacarosa, donde la sacarosa es la única fuente de carbono, como factor de selección. Un subconjunto de estos clones positivos de cada construcción de transformación se seleccionó y la presencia de la construcción de expresión se confirmó mediante ensayos de transferencia de tipo Southern. La función de la sacarosa invertasa de levadura también se confirmó mediante un ensayo de hidrólisis de la sacarosa. Los clones positivos se seleccionaron y se cultivaron en medios que contenían sacarosa como única fuente de carbono a una concentración inicial de 40 g/l. También se incluyó un control negativo de *Prototheca moriformis* (UTEX 1435) de tipo salvaje que se cultivó en medio que contenía glucosa como única fuente de carbono a la misma concentración de partida de 40 g/l. La utilización de la sacarosa se midió a través del curso del experimento por mediciones del nivel de sacarosa en el medio, mediante el uso de un analizador bioquímico YSI 2700 con una membrana específica de sacarosa. Después de seis días de cultivo, los cultivos se cosecharon y se procesaron para el perfil lipídico, mediante el uso de los mismos métodos que se describieron anteriormente. Los resultados del perfil lipídico se resumen a continuación en la Tabla 17 y se muestran en % del área.

Tabla 17. Perfiles lipídicos de transformantes dobles con *suc2* sacarosa invertasa y tioesterasa.

Ácido graso	PP	C24 A	C24 B	C24 c	C25 A	C25 B	C25 c	C26 A	C26 B	C26 c
C 10:0	0.01	**0.03**	**0.04**	**0.08**	0.01	0.01	0.01	0.01	0.01	0.0
C 12:0	0.04	0.04	0.04	0.04	**0.28**	**0.40**	**0.10**	0.04	0.04	0.13
C 14:0	1.6	1.55	1.53	1.56	1.59	1.59	1.60	**1.65**	**1.56**	**2.69**
C 14:1	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.03
C 15:0	0.04	0.03	0.03	0.04	0.04	0.03	0.03	0.03	0.03	0.04
C 16:0	29.2	29.1	29.0	28.6	28.9	28.6	29.0	28.8	29.5	27.5
C 16:1	0.86	0.80	0.79	0.82	0.77	0.81	0.82	0.79	0.79	0.86
C 17:0	0.1	0.08	0.08	0.09	0.09	0.08	0.09	0.08	0.08	0.09
C 17:1	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.03	0.04
C 18:0	3.26	3.33	3.37	3.27	3.36	3.28	3.18	3.33	3.36	3.03
C 18:1	54.5	53.9	54.1	53.9	53.5	53.7	53.5	54.2	53.9	52.7
C 18:2	8.72	9.35	9.22	9.45	9.68	9.65	9.87	9.31	9.06	10.8
C 18:3 alfa	0.63	0.71	0.69	0.73	0.74	0.73	0.75	0.71	0.66	0.83
C 20:0	0.29	0.31	0.31	0.31	0.32	0.32	0.31	0.32	0.31	0.29

Todos los clones positivos seleccionados para el ensayo de la utilización de sacarosa hidrolizaron la sacarosa en el medio y al final del período de 6 días de cultivo, no había niveles medibles de sacarosa en el medio. Estos datos, además del uso exitoso de la sacarosa como una herramienta de selección de los clones positivos, indican que el gen *suc2* de la sacarosa invertasa exógena de levaduras se orientó correctamente y se expresó en los transformantes. Como se mostró en la Tabla 17 anterior, los clones que expresaban la Construcción 24 (la tioesterasa C8-10) tuvieron un aumento medible en ácidos grasos C10 (tan alto como un aumento de ocho veces). Asimismo hubo aumentos medibles en los clones que expresaban la Construcción 25 (la tioesterasa C12) y la Construcción 26 (la tioesterasa C14) en los correspondientes ácidos grasos de cadena mediana. En conjunto, los datos muestran la expresión simultánea exitosa en *Prototheca moriformis* de dos proteínas recombinantes (por ejemplo, la sacarosa invertasa y una acil ACP tioesterasa de ácidos grasos), los cuales confieren cambios fenotípicos útiles y cuantificables en el organismo hospedero.

Ejemplo 8: Efectos del glicerol en la producción de ácidos grasos C10-C14 en transformantes de la tioesterasa C14

Los clones de todas las transformaciones de tioesterasas se seleccionaron y se evaluaron adicionalmente. Un clon que expresaba la Construcción 8 (C14 TE de *Cinnamomum camphora*) se cultivó heterofílicamente mediante el uso de distintas fuentes de carbono: sólo glucosa, sólo fructosa y sólo glicerol. La condición de sólo glucosa resultó en mayor crecimiento celular y producción de lípidos totales, en comparación con las condiciones de sólo fructosa y sólo glicerol. Sin embargo, la proporción de ácidos grasos C12-14 que se produjeron en la condición de sólo glicerol fue dos veces mayor que la que se alcanzó en la condición de sólo glucosa.

Ejemplo 9: Expresión de la invertasa de *Arabidopsis thaliana* en *Protoheca moriformis*

La cepa de microalga *Prototheca moriformis* (UTEX 1435) se transformó mediante el uso de los métodos que se describieron anteriormente, con una construcción de expresión que contenía una invertasa de *Arabidopsis thaliana* asociada a la pared celular con codón optimizado (según la Tabla 1). La secuencia de la invertasa de *Arabidoposis* se modificó para incluir el N-terminal de 39 aminoácidos de la levadura invertasa (proteína suc2) para garantizar la eficiente dirección al RE y en última instancia al periplasma. Para ayudar a la detección, se añadió un epítopo FLAG al C-terminal de la proteína recombinante. El transgén se clonó en un vector de expresión con una región promotor/5'UTR de la glutamato deshidrogenasa de *Chlorella sorokinianna* y una región nitrato reductasa 3'UTR de *Chlorella vulgaris*. La secuencia de ADN de este casete del transgen se muestra como SEQ ID NO: 89 y la secuencia de aminoácidos traducida se muestra como SEQ ID NO: 90. Los clones positivos se examinaron y seleccionaron mediante el uso de placas/medio que contenían sacarosa. Un subconjunto de los clones positivos se confirmó por la presencia del transgén y la expresión de la invertasa mediante el análisis de transferencia de tipo Southern y Western para la invertasa con la etiqueta FLAG. A partir de este examen, se seleccionaron 10 clones positivos por la productividad de lípidos y las pruebas de la utilización de sacarosa. Los 10 clones se cultivaron en un medio que contenía sacarosa como única fuente de carbono y se incluyó también un control positivo del transformante de invertasa suc2. El control negativo, *Prototheca moriformis* de tipo salvaje también se cultivó pero en un medio que contenía glucosa. Después de seis días, las células se cosecharon y se secaron y se determinó el porcentaje total de lípidos por peso seco de células. Los medios también se analizaron para el consumo total de sacarosa.

Los diez clones positivos hidrolizaron la sacarosa, sin embargo, la mayoría de los clones crecieron aproximadamente la mitad, ya sea el tipo salvaje o el control positivo del transformante de levadura invertasa suc2 según se determinó por el peso seco de las células al final del experimento. Del mismo modo, los diez clones positivos produjeron aproximadamente la mitad del total de lípidos tanto cuando se comparó con el tipo salvaje o con el transformante del control positivo. Estos datos demuestran el éxito de la expresión heteróloga de diversas sacarosas invertasas en *Prototheca.*

Ejemplo 10: Expresión heteróloga de la levadura invertasa (suc2) en *Prototheca krugani, Prototheca stagnora* y *Prototheca zopfii*

Para probar la aplicabilidad general de los métodos de transformación para su uso en las especies del género *Prototheca*, otras tres especies de *Prototheca* se seleccionaron: *Prototheca krugani* (UTEX 329), *Prototheca stagnora* (UTEX 1442) y *Prototheca zopfii* (UTEX 1438). Estas tres cepas se cultivaron en los medios y las condiciones que se describieron en el Ejemplo 1 y sus perfiles de lípidos se determinaron mediante el uso de los métodos que se describieron anteriormente. Un resumen de los perfiles de lípidos de las tres cepas de *Prototheca* se resume a continuación en % del área.

Ácido graso	*P. krugani* (UTEX 329)	*P. stagnora* (UTEX 1442)	*P. zopfii* (UTEX 1438)
C 10:0	0.0	0.0	0.0
C 10:1	0.0	0.0	0.0
C 12:0	1.5	0.8	2.1
C 14:0	1.2	0.9	1.7
C 16	15.1	17.1	19.7
C 18:0	3.3	4.1	5.4
C 18:1	66.0	61.5	53.8
C 18:2	12.9	15.6	17.3

Estas tres cepas se transformaron con un casete de expresión de una levadura invertasa (suc2) (SEQ ID NO: 58) mediante el uso de los métodos que se describieron en el Ejemplo 3 anterior. Este casete de expresión de la Levadura invertasa (suc2) se demostró anteriormente en el Ejemplo 3 que funcionaba en *Prototheca moriformis* (UTEX 1435). Los transformantes se detectaron mediante el uso de placas/medio que contenían sacarosa. Se seleccionó un subconjunto de los clones positivos para cada especie de *Prototheca* y se confirmó la presencia del transgén por análisis de transferencia de tipo Southern. Diez de los clones positivos confirmados de cada especie se seleccionaron para el análisis de la hidrólisis de sacarosa y la productividad de los lípidos. Los clones se cultivaron en medios que contenían sacarosa como única fuente de carbono y se compararon con su contraparte de tipo salvaje que creció en glucosa. Después de 6 días, los cultivos se cosecharon y se secaron y se evaluó el porcentaje de lípidos totales y el peso seco de las células. Los medios de cada cultivo también se analizaron para la hidrólisis de sacarosa mediante el uso de un analizador bioquímico YSI2700 para el contenido de sacarosa en el transcurso del experimento. Los clones de las tres especies hidrolizaron la sacarosa, los transformantes de *Prototheca stagnora* y *Prototheca zopfii* hirdrolizaron la sacarosa más eficientemente que la *Prototheca krugani.* La producción de lípidos totales y el peso seco de las células de las tres especies de transformantes fueron comparables con sus contrapartes de tipo salvaje que crecieron en glucosa. Estos datos demostraron el éxito de la transformación y la expresión de genes exógenos en varias especies del género *Prototheca*.

Ejemplo 11: Promotores y genes derivados de algas para su uso en microalgas

A. <u>Secuencias de promotores y 5'UTR de *Chlorella protothecoides*</u>

Una genoteca de ADNc se generó a partir de *Chlorella protothecoides* (UTEX 250) que creció mixotroficamente mediante el uso de técnicas estándar. En base a las secuencias de ADNc, las secuencias de los cebadores se diseñaron en ciertos genes domésticos conocidos, para "caminar" corriente arriba de las regiones de codificación mediante el uso del kit de ADN Walking de Seegene (Rockville, MD). Las secuencias aisladas incluyen una actina (SEQ ID NO: 31) y el promotor del factor de elongación-1a (EF1a) (SEQ ID NO: 32)/UTR, los cuales contienen intrones (como se muestra en minúscula) y exones (mayúsculas en cursiva) y el sitio de inicio previsto (en negrita) y dos elementos de la beta-tubulina, promotor/UTR: La isoforma A (SEQ ID NO: 33) y la isoforma B (SEQ ID NO: 34).

B. <u>Enzimas de la biosíntesis de lípidos y las secuencias dirigidas a los plastidios de *C. protothecoides*</u>

Desde la genoteca de ADNc que se describió anteriormente, tres ADNc que codifican proteínas funcionales en el metabolismo de los lípidos en *Chlorella protothecoides* (UTEX 250) se clonaron mediante el uso de los mismos métodos que se describieron anteriormente. Las secuencias de nucleótidos y aminoácidos para una acil ACP desaturasa (SEQ ID NO: 45 y 46) y dos geranil geranil difosfato sintasas (SEQ ID NO: 47-50) se incluyen en la lista

de secuencias a continuación. Además, tres ADNc con la secuencia señal putativa dirigida a los plastidios se clonaron también. Las secuencias de nucleótidos y aminoácidos para una gliceraldehído-3-fosfato deshidrogenasa (SEQ ID NO: 51 y 52), una proteína compleja de desprendimiento de oxígeno OEE33 (SEQ ID NO: 53 y 54) y una proteasa Clp (SEQ ID NO: 55 y 56) se incluyeron en la lista de secuencias a continuación. La secuencia putativa dirigida a plastidio se subrayó, tanto en la secuencia de nucleótidos como de aminoácidos. Las secuencias dirigidas a plastidios se pueden usar para dirigir los productos de los transgenes a los plastidios de los microbios, tales como las enzimas de modificación de lípidos.

Ejemplo 12: 5'UTR/promotores de *Prototheca moriformis* que son sensibles al nitrógeno

Una genoteca de ADNc se generó a partir *Prototheca moriformis* (UTEX 1435) mediante el uso de técnicas estándar. Las células de *Prototheca moriformis* se cultivaron durante 48 horas en condiciones abundantes de nitrógeno. Entonces un inóculo de 5% (v/v) se transfirió a bajo nitrógeno y las células se cosecharon cada 24 horas durante siete días. Después de aproximadamente 24 horas en cultivo, el suministro de nitrógeno en el medio se agotó por completo. Las muestras que se recolectaron se congelaron de inmediato mediante el uso de hielo seco e isopropanol. El ARN total se aisló posteriormente de las muestras de sedimento de células congeladas y una porción de cada muestra se mantuvo en reserva para los estudios de RT-PCR. El resto del ARN total que se cosechó de las muestras se sometió a la selección de poliA. Cantidades equimolares de ARN seleccionado de poliA de cada condición se agruparon entonces, y se usaron para generar una genoteca de ADNc en el vector pcDNA 3.0 (Invitrogen). Se seleccionaron al azar alrededor de 1200 clones de la genoteca de ADNc resultante que se agruparon y se sometieron a secuenciación en ambas cadenas. Aproximadamente 68 ADNc diferentes se seleccionaron entre las 1200 secuencias y se usaron para diseñar los cebadores específicos del ADNc para su uso en estudios de RT-PCR en tiempo real.

El ARN que se aisló de las muestras de sedimentos de células que se reservaron, se usó como sustrato en los estudios de RT-PCR en tiempo real que usaron el conjunto de cebadores específicos del ADNc que se generaron anteriormente. Esta reserva de ARN se convirtió en ADNc y se usó como sustrato para la RT-PCR para cada uno del conjunto de 68 cebadores específicos del gen. El ciclo umbral o los números C_T se usaron para indicar la relativa abundancia de transcripción de cada uno de los 68 ADNc dentro de cada muestra de ARN que se colectó a lo largo del tiempo. Los ADNc que mostraron un aumento significativo (más de tres veces) entre las condiciones de abundante de nitrógeno y nitrógeno empobrecido se identificaron como posibles genes cuya expresión se reguló de forma ascendente por el agotamiento de nitrógeno. Como se discutió en la especificación, el agotamiento/la limitación de nitrógeno, se conoce que es un inductor de la lipogénesis en microorganismos oleaginosos.

Con el fin de identificar las secuencias putativas de los promotores/5'UTR del ADNc cuya expresión se reguló de forma ascendente durante el agotamiento/limitación de nitrógeno, el ADN total se aisló de *Prototheca moriformis* (UTEX 1435) que crecieron en condiciones abundantes de nitrógeno y se sometieron a secuenciación mediante el uso de la tecnología de secuenciación 454 (Roche). Los ADNc marcados como que se regulan de forma ascendente según los resultados de RT-PCR anteriores, se compararon mediante el uso de BLAST contra los cóntigos ensamblados que provenían de lecturas de la secuenciación 454 del genoma. Los extremos 5 'del ADNc se correlacionaron con cóntigos específicos, y de ser posible, más de 500bp del flanco 5' de ADN se usó para identificar promotores/UTR putativos. La presencia de los promotores/5'UTR se confirmó posteriormente y se clonó mediante la amplificación por PCR del ADN genómico. Los extremos 5 ' individuales del ADNc se usaron para diseñar los cebadores 3' y los extremos 5' de los ensamblajes de los cóntigos 454 se usaron para diseñar los cebadores 5' específicos de genes.

Como un primer examen, uno de los promotores putativos, el 5'UTR/promotor aislado de Aat2 (transportador de amonio, SEQ ID NO: 99), se clonó en la construcción de la tioesterasa C14 de *Cinnamomum camphora* con el péptido de tránsito de la estearoil ACP desaturasa de *Chlorella protothecoides* que se describió en el Ejemplo 5 anterior, se sustituyó el promotor de la glutamato deshidrogenasa de *C.sorokinana*. Esta construcción se mostró como la SEQ ID NO: 112. Para probar el promotor putativo, la construcción de tioesterasa se transformó en las células de *Prototheca moriformis* para confirmar la real actividad del promotor por la detección de un aumento de los ácidos grasos C14/C12 bajo condiciones de bajo/sin nitrógeno, mediante el uso de los métodos que se describieron anteriormente. Las pruebas similares para los promotores putativos que se regulan con nitrógeno que se aíslan a partir de la detección del ADNc/genoma, se pueden hacer mediante el uso de los mismos métodos.
Otros 5'UTR/promotores regulados por nitrógeno putativos que se aislaron a partir de la detección del ADNc/genoma fueron:

5'UTR/promotor	SEQ ID NO	veces de aumento
5'UTR/promotor FatB/A	SEQ ID NO: 91	n/a
5'UTR/promotor NRAMP transportadora de metales	SEQ ID NO: 92	9.65
5'UTR/promotor de la proteína Flap Flagelar asociada	SEQ ID NO: 93	4.92

5'UTR/promotor de sulfito reductasa SulfRed	SEQ ID NO: 94	10.91
5'UTR/promotor del transportador de azúcar SugT	SEQ ID NO: 95	17.35
5'UTR/promotor del transportador de amonio 03 Amt03	SEQ ID NO: 96	10.1
5'UTR/promotor del transportador de amonio 02 Amt02	SEQ ID NO: 97	10.76
5'UTR/promotor del transportador de aminoácido 01 Aat01	SEQ ID NO: 98	6.21
5'UTR/promotor del transportador de aminoácido 02 Aat02	SEQ ID NO: 99	6.5
5'UTR/promotor del transportador de aminoácido 03 Aat03	SEQ ID NO: 100	7.87
5'UTR/promotor del transportador de aminoácido 04 Aat04	SEQ ID NO: 101	10.95
5'UTR/promotor del transportador de aminoácido 05 Aat05	SEQ ID NO: 102	6.71

Las veces de incremento se refiere a la veces de incremento en abundancia de ADNc después de 24 horas de cultivo en medio bajo en nitrógeno.

Ejemplo 13: Recombinación homóloga en las especies de *Prototheca*

La recombinación homóloga de transgenes tiene varias ventajas sobre los métodos de transformación que se describieron en los ejemplos anteriores. En primer lugar, la introducción de transgenes, sin recombinación homóloga puede ser impredecible, porque no hay control sobre el número de copias del plásmido que se introduce en la célula. Además, la introducción de transgenes, sin recombinación homóloga puede ser inestable debido a que el plásmido puede permanecer episomal y se pierde en las divisiones celulares posteriores. Otra de las ventajas de la recombinación homóloga es la capacidad de "desactivación" de genes dianas, de introducir epítopo etiquetas, cambiar promotores de genes endógenos y otro tipo de modificaciones de genes diana (por ejemplo, la introducción de mutaciones puntuales.

Se construyeron dos vectores con una región específica del genoma de *Prototheca moriformis* (UTEX 1435), que se designó KE858. La KE858 es un fragmento genómico 1.3 kb que abarca parte de la región que codifica para una proteína que comparte homología con el ARN de transferencia (ARNt) de la familia de proteínas. La transferencia de tipo Southern demostró que la secuencia KE858 estaba presente en una sola copia en el genoma de *Prototheca moriformis* (UTEX 1435). El primer tipo de vector que se construyó, designado SZ725 (SEQ ID NO: 103), consistía en el fragmento de 1.3 kb KE858 completo, clonado en un esqueleto del vector pUC19 que también contenía el gen optimizado (*suc2*) de la levadura invertasa que se usó en el Ejemplo 3 anterior. El fragmento KE858 contenía un sitio único SnaB1 que no aparecía en ningún otro lugar de la construcción diana. El segundo tipo de vector que se construyó, designado SZ726 (SEQ ID NO: 126), consistió en la secuencia KE858 que se interrumpió por la inserción del gen de la levadura invertasa *(suc2)* en el sitio SnaB1 dentro de la secuencia genómica KE858. El fragmento de ADN que contiene las secuencias KE858 que flanquean el gen levadura invertasa se puede escindir del esqueleto del vector mediante digestión con EcoRI, que corta en cada extremo de la región KE858.

Ambos vectores se usaron para dirigir la recombinación homóloga del gen (*suc2*) de la levadura invertasa en la región KE858 correspondiente del genoma de *Prototheca moriformis* (UTEX 1435). Los extremos del ADN lineal homólogo a la región genómica que era el objetivo de la recombinación homóloga se expusieron por digestión del vector de la construcción SZ725 con SnaB1 y el vector de la construcción SZ726 con EcoRI. Las construcciones de los vectores digeridos se introdujeron entonces en los cultivos de *Prototheca moriformis* mediante el uso de los métodos que se describieron anteriormente en el Ejemplo 3. Los transformantes de cada construcción del vector se seleccionaron entonces mediante el uso de placas de sacarosa. Diez transformantes independientes, clonalmente puros de cada vector de transformación se analizaron para la recombinación exitosa del gen de la levadura invertasa en el lugar del genoma que se deseaba (mediante el uso de la transferencia de tipo Southern) y para la estabilidad del transgén.

El análisis de transferencia de tipo Southern de los transformantes SZ725 mostró que 4 de los 10 transformantes recogidos para el análisis contenían las bandas recombinantes previstas, lo que indicaba que un único evento de cruce se produjo entre la secuencias KE858 en el vector y las secuencias KE858 en el genoma. Por el contrario, los diez transformantes SZ726 contenían las bandas recombinantes previstas, lo que indicaba que se produjeron eventos de doble cruce entre el fragmento EcoRI de pSZ726 que portaba la secuencia KE858 que flanqueaba el transgén de la levadura invertasa y la región KE858 correspondiente del genoma.

La expresión de sacarosa invertasa y la estabilidad del transgén se evaluaron por el crecimiento de los transformantes por más de 15 generaciones en ausencia de la selección. Los cuatro transformantes SZ725 y los diez transformantes SZ276que dieron positivo para el transgén por transferencia de tipo Southern se seleccionaron y 48 colonias individuales de cada uno de los transformantes se cultivaron en serie: en primer lugar sin selección en medio con glucosa y luego con la selección de los medios que contenían sacarosa como la única fuente de carbono. Los diez transformantes SZ276 (100%) mantuvieron su capacidad de crecer en sacarosa al cabo de 15 generaciones, mientras que aproximadamente 97% de los transformantes SZ725 conservaron su capacidad de crecer en sacarosa al cabo de 15 generaciones. Los transgenes que se introdujeron por un evento de cruce doble (vector SZ726) tienen una estabilidad muy alta durante el doblaje de las generaciones. En cambio, los transgenes

que se introdujeron por un solo cruce (vector SZ725) pueden dar lugar a cierta inestabilidad durante el doblaje de las generaciones porque se introdujeron copias en tándem de los transgenes, las regiones homólogas repetidas que flanquean los transgenes se pueden recombinar y eliminar del ADN transgénico que se encontraba entre ellos.
Estos experimentos demuestran el uso exitoso de la recombinación homóloga para generar transformantes de *Prototheca* que contengan un gen heterólogo de la sacarosa invertasa que esté integrado de forma estable en los cromosomas nucleares del organismo. El éxito de la recombinación homóloga permite otras alteraciones genómicas en *Prototheca*, incluidas las supresiones de genes, las mutaciones puntuales y el epítopo etiquetado de un producto génico deseado. Estos experimentos demuestran también el primer sistema documentado para la recombinación homóloga en el genoma nuclear de un microalga eucariota.

A. Uso de la recombinación homóloga para la desactivación de un gen endógeno de *Prototheca moriformis*

En la detección en el ADNc/ genoma de *Prototheca moriformis* que se describió en el Ejemplo 11, se identificó el ADNc de una estearoil ACP desaturasa endógena (SAPD). Las enzimas estearoil ACP desaturasas forman parte de la vía de síntesis de los lípidos y tienen la función de introducir dobles enlaces en las cadenas de acil graso. En algunos casos, puede ser una ventaja desactivar o reducir la expresión de enzimas de la ruta de los lípidos con el fin de modificar un perfil de ácidos grasos. Una construcción de recombinación homóloga se creó para evaluar si la expresión de una enzima endógena estearoil ACP desaturasa se podía reducir (o desactivar) y si se podía observar la correspondiente reducción de los ácidos grasos insaturados en el perfil lipídico de la célula hospedera.Una secuencia codificante de aproximadamente 1.5kb de un gen de estearoil ACP desaturasa de *Prototheca moriformis* (UTEX 1435) se identificó y se clonó (SEQ ID NO: 104). La construcción de recombinación homóloga se construyó mediante el uso de 0.5kb de la secuencia codificante de SAPD en el extremo 5' (sitio de dirección 5'), seguida del promotor de la β-tublina de *Chlamydomonas reinhardtii* que conducía un gen *suc2* de la sacarosa invertasa de levadura con codón optimizado, con 3'UTR *de Chlorella vulgaris*. El resto (~ 1 kb) de la secuencia codificante de la SAPD de *Prototheca moriformis* se insertó entonces, después de la 3'UTR de *C.vulgaris* para conformar el sitio diana 3'. La secuencia de este casete de recombinación homóloga se muestra en la SEQ ID NO: 105. Como se observó anteriormente, la tasa de éxito para la integración del casete de recombinación homóloga en el genoma nuclear se puede aumentar mediante la linealización del casete y dejar los extremos expuestos antes de la transformación de la microalga. El casete de recombinación homóloga dirigida a una enzima SAPD endógena en *Prototheca moriformis* se lineanlizó y luego se transformó en la célula hospedera (*Prototheca moriformis*, UTEX 1435). Una integración exitosa eliminaría del genoma del hospedero la región de codificación de la enzima SAPD endógena a través de un evento de doble recombinación recíproca, mientras que la expresión del nuevo gen *suc2* insertado se regularía por el promotor de la β-tubulina de *C.reinhardtii*. Los clones resultantes se pueden examinar mediante el uso placas/medios que contengan sacarosa como única fuente de carbono. Los clones que contengan una integración exitosa del casete de recombinación homóloga tendrán la capacidad de crecer en sacarosa como única fuente de carbono y los cambios en la saturación global de los ácidos grasos en el perfil lipídico servirá como factor de confirmación secundaria. Además, los ensayos de transferencia de tipo Southern que usen una sonda específica para el gen *suc2* de la sacarosa invertasa de levadura y la RT-PCR también podrán confirmar la presencia y expresión del gen de la invertasa en los clones positivos. Como alternativa, la misma construcción sin el promotor de la β-tubulina se puede usar para cortar la región de codificación de la enzima SAPD endógena. En este caso, el gen *suc2* de la sacarosa invertasa de levadura recién incorporado, será regulado por el promotor de la SAPD endógena/5'UTR.

Ejemplo 14: Producción de combustible

A. Extracción del aceite de microalgas, mediante el uso de una prensa expulsora y una presión auxiliar

La biomasa de microalgas que contenía 38% de aceite por PSC se secó con un secador de tambor lo que resultó en un contenido de humedad resultante de 5-5.5%. La biomasa se introdujo en una prensa francesa L250. Se introdujeron a través de la prensa 30.4 kg (67 lbs.) de la biomasa y no se recuperó aceite. La misma biomasa microbiana seca se combinó con distintos porcentajes de pasto forrajero como una presión auxiliar que se introdujo a través de la prensa. La combinación de la biomasa microbiana seca y el 20% p/p del pasto forrajero produjo el mejor porcentaje de rendimiento de aceite en general. Las tortas de prensado se sometieron a extracción con hexano y el rendimiento final de la condición de 20% de pasto forrajero fue de 61.6% del aceite total disponible (calculado por peso). La biomasa por encima de 50% del peso seco de las células de aceite no requiere el uso de una presión auxiliar, como el pasto forrajero para liberar el aceite.

B. Composición de los monosacáridos de la biomasa de *Prototheca moriformis* deslipidada

La Prototheca moriformis (UTEX 1435) se cultivó en condiciones y medio nutriente (con 4% de glucosa) como se describió en el Ejemplo 45 anterior. La biomasa de microalgas se cosechó a continuación, y se secó con un secador de tambor. La biomasa de algas secas se lisó y el aceite se extrajo por medio de una prensa expulsora como se describió en el Ejemplo 44 anterior. El aceite residual de la biomasa presionada entonces se extrajo mediante el uso de disolventes de éter de petróleo. El éter de petróleo residual se evaporó de la comida deslipidada mediante el uso de un rotoevaporador (Buchi Labortechnik AG, Suiza). El análisis de la composición de glicosil

(monosacáridos) se realizó entonces, en la comida deslipidada mediante el uso de la combinación de la cromatografía de gases con la espectrometría de masas (GC/MS) del per-O-trimetilsili (TMS) derivado de los monosacáridos metil glucósidos producidos a partir de la muestra por metanólisis ácida. Una muestra de comida deslipidada se sometió a metanólisis en HCl 1M en metanol a 80°C durante aproximadamente 20 horas, seguida de una nueva N-acetilación con piridina y anhídrido acético en metanol (para la detección de azúcares amino). Las muestras se sometieron entonces a per O-trimetilsililación por el tratamiento con Tri-Sil (Pierce) a 80°C durante 30 minutos (ver métodos en Merkle y Poppe (1994) *Methods Enzymol.* 230: 1-15 y York y otros, (1985) *Methods Enzymol.* 118:3-40). La GC/MS de los glucósidos de metilo TMS se llevó a cabo en un sistema HP 6890 GC conectado a un MSD 5975b, mediante el uso de una columna capilar de sílice fundida All Tech EC-1 (30m x 0.25 mm de DI). Los monosacáridos se identificaron por sus tiempos de retención con respecto a los estándares, y el carácter de estos carbohidratos es autenticado por sus espectros de masas. Se añadieron 20 microgramos por muestra de inositol a la muestra antes de la derivatización como un estándar interno. El perfil de monosacárido de la biomasa de *Prototheca moriformis* deslipidada (UTEX 1435) se resume en la Tabla 18 a continuación. El porcentaje de carbohidratos totales de la muestra se calculó en 28.7%.

Tabla 18. Análisis de la composición de monosacárido (glicosilo) de la biomasa deslipidada (UTEX 1435) de *Prototheca moriformis*.

	Masa (µg)	**% Molar (del total de carbohidrato)**
Arabinosa	0.6	1.2
Xilosa	n.d.	n.d.
Ácido galacturónico (GalUA)	n.d.	n.d.
Manosa	6.9	11.9
Galactosa	14.5	25.2
Glucosa	35.5	61.7
N acetil galactosamina (GalNAc)	n.d.	n.d.
N acetil glucosamina (GlcNAc)	n.d.	n.d.
Heptosa	n.d.	n.d.
ácido 3 deoxi--2-manno-2 octulsónico (KDO)	n.d.	n.d.
Suma	**57**	**100**

n.d. = ninguno detectado

El contenido de carbohidratos y monosacáridos de la composición de la comida deslipidada lo hace adecuado para uso como alimento para animales o como parte de una formulación de piensos para animales. Así, en un aspecto, la presente invención proporciona comida deslipidada que tiene el contenido de los productos establecidos en la tabla anterior.

C. <u>Producción de biodiesel a partir de aceite de *Prototheca*</u>

El aceite desgomado de *Prototheca moriformis* UTEX 1435, que se produjo de acuerdo a los métodos que se describieron anteriormente, se sometió a transesterificación para producir ésteres metílicos de ácidos grasos. Los resultados se muestran a continuación:
El perfil lipídico del aceite fue:

C10:0 0.02
C12:0 0.06
C14:0 1.81
C14.1 0.07
C16:0 24.53
C16:1 1.22
C18:0 2.34
C18:1 59.21
C18:2 8.91
C18:3 0.28
C20:0 0.23
C20:1 0.10
C20:1 0.08
C21:0 0.02
C22:0 0.06
C24:0 0.10

Tabla 19. Perfil de biodiesel del aceite de triglicérido de*Prototheca moriformis* .

Método	Prueba		Resultado	Unidades
ASTM D6751 A1	Filtrabilidad de impregnado en frío de mezclas combustible de biodiesel	Tiempo de filtración	120	seg.
		Volumen Filtrado	300	ml
ASTM D93	Punto de inflamación en vaso cerrado Pensky-Martens	Procedimiento usado	A	
		Punto de inflamación corregido	165.0	°C
ASTM D2709	Agua y sedimento en combustibles destilados intermedios (Método centrífuga)	Sedimento y agua	0.000	% en vol
EN 14538	Determinación del contenido de Ca y Mg por ICP OES	Suma de (Ca y Mg)	<1	mg/kg
EN 14538	Determinación del contenido de Ca y Mg por ICP OES	Suma de (Na y K)	<1	mg/kg
ASTM D445	Vicosidad cinemática / dinámica	viscosidad cinemática @ 104 ºF/ 40 ºC	4.873	mm²/s
ASTM D874	Ceniza sulfatada a partir de aceites lubricantes y aditivos	Ceniza sulfatada	< 0.005	%p
ASTM D5453	Determinación del azufre total en hidrocarburos ligeros, combustible para motor de ignición por chispa, combustible para motor diesel, y aceite de motor por fluorescencia ultravioleta.	Azufre, mg /kg	1.7	mg/kg
ASTM D130	Corrosión - tira de cobre	Corrosión Cu-Biodiesel 50ºC (122ºF)/3 h	1a	
ASTM D2500	Punto de turbidez	Punto de turbidez	6	°C
ASTM D4530	Residuo de micro carbono	Promedio del residuo de carbono del micro método	< 0.10	%p
ASTM D664	Índice de acidez de los productos de petróleo por valoración potenciométrica	Procedimiento usado	A	
		índice de acidez	0.20	mg KOH/g
ASTM D6584	Determinación de la glicerina libre y total en metil éster de Biodiesel B-100 por cromatografía de gases	Glicerina libre	< 0.005	%p
		Glicerina total	0.123	%p
ASTM D4951	Elementos aditivos en aceites lubricantespor ICP-AES	Fósforo	0.000200	%p
ASTM D1160	Destilación de productos de petróleo a presión reducida	IBP	248	°C
		AET @ 5% Recuperación	336	°C
		AET @ 10% Recuperación	338	°C
		AET @ 20% Recuperación	339	°C
		AET @ 30% Recuperación	340	°C
		AET @ 40% Recuperación	342	°C
		AET @ 50% Recuperación	344	°C
		AET @ 60% Recuperación	345	°C
		AET @ 70% Recuperación	347	°C
		AET @ 80% Recuperación	349	°C
		AET @ 90% Recuperación	351	°C
		AET @ 95% Recuperación	353	°C
		FBP	362	°C
		% Recuperado	98.5	%

(continuación)

Método	Prueba		Resultado	Unidades
		% Pérdidas	1.5	%
		% Residuo	0.0	%
		Volumen de la trampa para frío	0.0	ml
		IBP	248	°C
EN 14112	Determinación de la estabilidad de oxidación (prueba de oxidación acelerada)	Estabilidad de oxidación	>12	h
		Temp. de operación (usualmente 110 grados C)	110	°C
ASTM D4052	Densidad de líquidos por medidor de densidad digital	peso API @ 60°F	29.5	°API
ASTM D 6890	Determinación de retardo en la ignición (ID) y número de cetano derivado (DCN)	Número de cetano derivado (DCN)	> 61.0	

El perfil lipídico del biodiesel fue muy similar al perfil lipídico del aceite como materia prima. Los demás aceites que proporcionaron los métodos y las composiciones de la invención se pueden someter a transesterificación para rendir biodiesel con los perfiles de lípidos, que incluyen (a) al menos 4% C8-C14, (b) al menos 0.3% C8; (c) al menos 2% C10, (d) al menos 2% C12 y (3) al menos 30% C8-C14.

La filtrabilidad por inmersión en frío por el método ASTM D6751 A1 de producción de biodiesel, fue de 120 segundos para un volumen de 300 ml. Esta prueba consiste en la filtración de 300 ml de B100, refrigerados a 40°F durante 16 horas, se dejó calentar a temperatura ambiente, y se filtró al vacío con fibra de vidrio de 0.7 micras con soporte de acero inoxidable. Los aceites de la invención se pueden transesterificar para generar biodiesel con un tiempo de inmersión en frío de menos de 120 segundos, menos de 100 y menos de 90 segundos.

D. <u>La producción de diesel renovable</u>

El aceite desgomado de *Prototheca moriformis* UTEX 1435, elaborado de acuerdo a los métodos que se describieron anteriormente y que tenía el mismo perfil lipídico que el aceite usado para producir biodiesel en el Ejemplo X anterior, se sometió a la transesterificación para producir diesel renovable.

El aceite primero se trató con hidrógeno para eliminar el oxígeno y el esqueleto de glicerol, lo que produjo n-parafinas. Las n-parafinas se sometieron a craqueo e isomerización. El cromatograma del material se muestra en la Figura 13. El material se sometió a filtración en frío, que eliminó aproximadamente 5% del material C18. Tras la filtración en frío el volumen total de material se redujo a punto de inflamación y se evaluó el punto de inflamación, la distribución de la destilación ASTM D-86, el punto de turbidez y viscosidad. El punto de inflamación fue de 63°C, la viscosidad fue de 2.86 cSt (centistokes), el punto de turbidez fue de 4°C. Los valores de la destilación ASTM D86 se muetran en la Tabla 20:

Tabla 20. Lecturas en °C:

Volumen	Temperatura
IBP	173
5	217.4
10	242.1
15	255.8
20	265.6
30	277.3
40	283.5
50	286.6
60	289.4
70	290.9
80	294.3
90	300
95	307.7
FBP	331.5

La T10-T90 del material producido fue 57.9°C. Los métodos de tratamiento con hidrógeno, la isomerización, y otra modificación covalente de los aceites que se describen en la presente, así como los métodos de destilación y fraccionamiento (tales como la filtración en frío) que se describen en la presente, se pueden emplear para generar

composiciones de diesel renovables con otros intervalos T90 T10, tales como, 20 , 25, 30, 35, 40, 45, 50, 60 y 65°C mediante el uso de los aceites triglicéridos que se producen de acuerdo con los métodos que se describen en la presente.

La T10 del material producido fue 242.1 °C. Los métodos de tratamiento con hidrógeno, la isomerización, y otra modificación covalente de los aceites que se describen en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se describen en la presente, se pueden emplear para generar composiciones de diesel renovables con otros valores de T10, tales como T10 entre 180 y 295, entre 190 y 270, entre 210 y 250, entre 225 y 245, y al menos 290.

La T90 del material que se produjo fue de 300 °C. Los métodos de tratamiento con hidrógeno, la isomerización, y otra modificación covalente de los aceites que se describen en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se describen en la presente, se pueden emplear para generar composiciones de diesel renovables con otros valores de T90, tales como T90 entre 280 y 380, entre 290 y 360, entre 300 y 350, entre 310 y 340, y al menos 290.

La FBP del material producido fue de 300°C. Los métodos de tratamiento con hidrógeno, la isomerización, y otra modificación covalente de los aceites que se describen en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se describen en la presente, se pueden emplear para generar composiciones de diesel renovables con otros valores FBP, tales como FBP entre 290 y 400, entre 300 y 385, entre 310 y 370, entre 315 y 360, y al menos 300.

Los demás aceites que proporcionan los métodos y las composiciones de la invención se pueden someter a combinaciones de tratamiento con hidrógeno, isomerización, y otras modificaciones covalentes, que incluyen los aceites con los perfiles de lípidos, (a) al menos 4% C8-C14, (b) al menos el 0.3% C8, (c) al menos 2% C10; (d) al menos 2% C12, y (3) al menos 30% C8-C14.

Ejemplo 15: Utilización de sacarosa por la *Chlorella luteoviridis*

A. <u>Crecimiento de la SAG 2214 en glucosa y sacarosa</u>

La SAG 2214 (designada como *Chlorella luteoviridis*) se probó para el crecimiento en la oscuridad en medios que contenían glucosa o sacarosa. Los cultivos líquidos heterótrofos se iniciaron mediante el uso de un inóculo de un frasco congelado en medio ya sea que contenía 4% de glucosa o 4% de sacarosa como única fuente de carbono. Los cultivos se cultivaron en la oscuridad, con agitación a 200 rpm. Se tomaron muestras de los cultivos en los puntos de tiempo 0, 24, 48 y 72 horas y el crecimiento se midió por la absorbancia relativa a 750nm (UV Mini1240, Shimadzu). La SAG 2214 creció igual de bien en la glucosa como en la sacarosa, lo que muestra que esta microalga puede utilizar la sacarosa tan eficazmente como la glucosa como única fuente de carbono. El resultado de este experimento se representa gráficamente en la Figura 3.

B. <u>Productividad de lípidos y el perfil de ácidos grasos de la SAG 2214</u>

La cepa de microalgas SAG 2214 se cultivó en un medio líquido que contenía glucosa o sacarosa como única fuente de carbono en condiciones similares a las descritas en el Ejemplo 32. Después de 7 días, las células se cosecharon para el cálculo del peso seco de las células. Las células se centrifugaron y se liofilizaron durante 24 horas. Los sedimentos de células secas se pesaron y se calculó el peso seco de células por litro. Las células para el análisis de los lípidos también se cosecharon y se centrifugaron a 4000 x g durante 10 minutos a temperatura ambiente. Se desechó el sobrenadante y las muestras se procesaron para el análisis de los lípidos y el perfil de ácidos grasos mediante el uso de los procedimientos estándar de cromatografía de gas (GC/FID). Los resultados se resumen a continuación en la Tablas 21 y 22.

Tabla 21. Productividad de lípidos y PSC para SAG 2214.

Muestra	**Lípido (g/l)**	**PSC (g/l)**	**% lípido PSC**
SAG 2214 glucosa	2.43	5.73	42.44%
SAG 2214 sacarosa	0.91	2.00	45.56%

Tabla 22. Perfil de ácido graso para SAG 2214.

Ácido graso	Porcentaje (p/p)
C:16:0	21
C:18:1	38
C:18:2	41

C. Comparación genómica de la SAG 2214 con otras cepas de *Chlorella luteoviridis*

La cepa de microalgas SAG 2214 resultó ser de interés general debido a su capacidad de crecer en sacarosa como fuente de carbono (se ilustró anteriormente). Además de las características de crecimiento de esta cepa, sus relaciones taxonómicas con otras especies de microalgas también fueron de interés. Designada por la colección del SAG como una cepa de *Chlorella luteoviridis*, el gen ARNr 23S de la SAG 2214 se secuenció y se comparó con la secuencia ARNr 23S del genoma de otras nueve cepas también identificadas por las colecciones SAG y UTEX como la *Chlorella luteoviridis*. Estas cepas fueron UTEX 21, 22, 28, 257 y 258, y las cepas del SAG 2133, 2196, 2198 y 2203. Los métodos para determinar el genotipo del ADN que se usaron, fueron los mismos métodos que se describieron anteriormente en el Ejemplo 1. Las alineaciones de las secuencias y los árboles sin raíces se generaron mediante el uso del software de análisis de ADN Geneious. De las nueve cepas a las que se le determinó el genotipo, UTEX 21, 22, 28 y 257 tuvieron idénticas secuencias de ADN del ARNr 23S (SEQ ID NO: 106). Las otras cinco cepas de *Chlorella luteoviridis* tuvieron secuencias de ARNr 23S que fueron altamente homólogas a UTEX 21, 22, 28 y 257.

La secuencia del gen ARNr 23S de la SAG 2214 (SEQ ID NO: 30) es decididamente diferente del de las otras nueve cepas *C. luteoviridis*, éste tiene una gran inserción que no se encontró en las otras cepas. Un análisis más detallado de esta secuencia del gen ARNr 23S que usó el BLAST indicó que compartía la mayor homología con los miembros del género *Leptosira* y *Trebouxia* (miembros de la porción ficobionte de los líquenes). Estos resultados indican que la SAG 2214 no puede ser una cepa de *Chlorella luteoviridis* según la clasificación de la colección de cepas, pero comparten la identidad de nucleótidos de ARNr 23S de manera significativa con las algas simbiontes que se encontraron en líquenes. El análisis genómico junto con las características de crecimiento indicó que la SAG 2214 puede ser una fuente de genes y proteínas implicadas en el metabolismo de la sacarosa, así como en la señalización y el tránsito de los péptidos responsables de la correcta localización de tales enzimas. La SAG 2214 y otras cepas con un alto grado de similitud genómica también pueden ser cepas útiles para la producción de aceite mediante el uso de la sacarosa como fuente de carbono fijo.

A pesar de que esta invención se describió en relación con realizaciones específicas de ésta, se entenderá que es capaz de modificaciones adicionales. Esta solicitud pretende cubrir cualquier variación, los usos, o adaptaciones de la invención siguiendo, en general, los principios de la invención y se incluyen tales desviaciones de la presente descripción que vienen dentro de la práctica conocida o habitual en la materia a la que pertenece la invención y como se puede aplicar a las características esenciales anteriormente expuestas.

Las publicaciones mencionadas en este documento se citan con el fin de describir y divulgar los reactivos, metodologías y conceptos que se pueden usar en relación con la presente invención. Nada en la presente se debe interpretar como una admisión de que estas referencias son materia anterior en relación con las invenciones descritas en la presente. En particular, se hace referencia a las siguientes solicitudes de patente: la solicitud provisional de Estados Unidos núm. 60/941,581, presentada el 1 de junio de 2007, titulada "Producción de hidrocarburos en microorganismos"; la solicitud provisional de Estados Unidos núm. 60 / 959174, presentada el 10 de julio de 2007, titulada "Producción de hidrocarburos en microorganismos"; la solicitud provisional de Estados Unidos núm. 60/968, 291, presentada el 27 de agosto de 2007, titulada "Producción de hidrocarburos en microorganismos"; la solicitud provisional de Estados Unidos núm. 61/024, 069 presentada el 28 de enero de 2008, titulada "Producción de hidrocarburos en microorganismos"; la solicitud PCT núm. PCT/US08/65563, presentada el 2 de junio de 2008, titulada "La producción de aceite en microorganismos"; la solicitud de patente de Estados Unidos núm. 12/131,783, presentada el 2 de junio de 2008, titulada "Uso de material celulósico para el cultivo de microorganismos"; la solicitud de patente de Estados Unidos núm. 12/131,773, presentada el 2 de junio de 2008, titulada "Diesel y combustible de avión renovables de origen microbiano "; la solicitud de patente de Estados Unidos núm. 12/ 131,793, presentada el 2 de junio de 2008, titulada "Utilización de sacarosa como materia prima para la fabricación d combustible a base de aceite."; la solicitud de patente de Estados Unidos núm. 12/131,766, presentada el 2 de junio de 2008, titulada " Utilización del glicerol como materia prima para la fabricación del combustible a base de aceite."; la solicitud de patente de Estados Unidos núm. 12/131,804, presentada el 2 de junio de 2008, titulada "Modificación de la ruta de los lípidos en microorganismos oleginosos."; la solicitud de patente de Estados Unidos núm. 61/118,590, presentada el 28 de noviembre de 2008, titulada "Producción de aceite en microorganismos"; la solicitud de patente provisional núm. US 61/118,994, presentada 1 de diciembre de 2008, titulada "Producción de aceite en microorganismos"; la solicitud de patente provisional de Estados Unidos núm. 61/174,357, presentada 03

de abril de 2009, titulada "Producción de aceite en microorganismos"; la solicitud de patente provisional de Estados Unidos núm. 61/219,525, presentada el 23 de junio de 2009, titulada "Producción de aceite en microorganismos", y la solicitud PCT núm. _____ [[Expediente núm. 026172-005020PC], presentada el 30 de noviembre de 2009, titulada "Fabricación de aceites adaptados en microorganismos heterótrofos recombinantes".

La divulgación proporciona además las siguientes realizaciones enumeradas.

1. Una célula del género *Prototheca* que comprende un gen exógeno.

2. La célula de la realización 1, en la que el gen exógeno es secuencia de codificación y está en ligadura operable con un promotor.

3. La célula de la realización 2, en la que el promotor es de un gen endógeno a una especie del género *Prototheca*.

4.La célula de la realización 2, en la que la secuencia de codificación codifica una proteína seleccionada del grupo que consiste de una sacarosa invertasa, una acil-ACP tioesterasa grasa, una acil-CoA/aldehído reductasa grasa, una acil-CoA reductasa grasa, una aldehído reductasa grasa, una aldehído descarbonilasa grasa, una proteína portadora de acilo y una proteína que imparte resistencia a un antibiótico.

5. La célula de la realización 4, en la que la proteína es una acil-ACP tioesterasa grasa que tiene actividad de hidrólisis hacia uno o más sustratos de acil-ACP grasos de longitud de cadena C8, C10, C12 o C14.

6. La célula de la realización 5, en la que la acil-ACP tioesterasa tiene al menos 50% de identidad de aminoácidos con una o más secuencias seleccionadas del grupo que consiste de SEQ ID NOs: 59, 61, 63 y 138-140.

7.La célula de la realización 5, en la que la secuencia de codificación comprende la secuencia dirigida al plastidio de microalgas.

8. La célula de la realización 7, en la que la secuencia dirigida al plastidio es de la especie del género *Prototheca*.

9. La célula de la realización 7, en la que la secuencia dirigida al plástido comprende al menos 25% de identidad de secuencia de aminoácidos a una secuencia de aminoácidos de SEQ ID NOs: 127-133 y es capaz de dirigir una proteína codificada por un gen exógeno no localizado en el genoma del plastido al plástido.

10. La célula de la realización 2, en la que el promotor es regulado de forma ascendente al menos 3 veces en una célula del género *Prototheca* cuando el ambiente extracelular cambia de contener al menos 5mM nitrógeno a no contener nitrógeno.

11. La célula de la realización 10, en la que el promotor comprende un segmento de 50 o más nucleótidos de una de las SEQ ID NO: 91-102.

12. La célula de la realización 1, en la que el gen exógeno se integra en el cromosoma de la célula.

13. La célula de la realización 1, en la que la célula tiene una secuencia rARN 23S con al menos 75% de identidad de nucleótido para una secuencia de nucleótidos de SEQ ID NOs: 11-19.

14. La célula de la realización 1, en la que la célula es *Prototheca moriformis, Prototheca krugani, Prototheca stagnora*, o *Prototheca zopfii*.

15. Una célula del género *Prototheca*, en la que comprende un gen acil-ACP tioesterasa graso exógeno y un perfil de lípidos de al menos alrededor de 4% de C8-C14.

16. La célula de la realización 15, en la que la cantidad de C8 es al menos 0.3%.

17. La célula de la realización 15, en la que la cantidad de C10 es al menos 2%.

18. La célula de la realización 15, en la que la cantidad de C12 es al menos 2%.

19. La célula de la realización 15, en la que la cantidad de C14 es al menos 4%.

20. La célula de la realización 15, en la que la cantidad de C8-C14 es al menos 30%.

21. La célula de la realización 15, caracterizada además porque comprende un gen de sacarosa invertasa exógeno.

22. La célula de la realización 15, en la que la célula tiene una secuencia rARN 23S con al menos 75% de identidad de nucleótido para una o más de SEQ ID NOs: 11-19.

23.La célula de la realización 15, en la que la célula es *Prototheca moriformis, Prototheca krugani, Prototheca stagnora*, o *Prototheca zopfii*.

24. La célula de la realización 15, en la que el gen acil-ACP tioesterasa graso exógeno se integra en el cromosoma de la célula.

25.Un método para elaborar una composición de triglicéridos de al menos alrededor de 4% de C8-C14, caracterizado porque cultivar la célula de la realización 15.

26.El método de la realización 25, en el que la cantidad de C8 es al menos el 0.3%.

27.El método de la realización 25, en el que la cantidad de C10 es al menos el 2%.

28.El método de la realización 25, en el que la cantidad de C12 es al menos el 2%.

29.El método de la realización 25, en el que la cantidad de C14 es al menos el 4%.

30.El método de la realización 25, en el que la cantidad de C8-C14 es al menos el 30%.

31.Un método para hacer una composición de triglicéridos que comprende cultivar la célula de la realización 15, en el que la célula comprende además un gen exógeno que codifica una sacarosa invertasa y la sacarosa se proporciona como una fuente de carbono.

32.El método de la realización 31, en el que la sacarosa invertasa tiene al menos 50% de identidad de aminoácidos a una o más de SEQ ID Nos: 3, 20-29 y 90.

33.Una composición de aceite de triglicéridos, que comprende un perfil de lípidos de:

a. al menos 4% de C8-C14; y
b. uno o más de los siguientes atributos:

i.menos de 0.4 microgramos/ml total carotenoides;
ii.menos de 0.001 microgramos/ml licopeno;
iii.menos de 0.02 microgramos/ml beta-caroteno;
iv.menos de 0.02 miligramos de clorofila por kilogramo de aceite;
v.0.40-0.60 miligramos de gamma tocoferol por 100 gramos de aceite;
vi.3-9 mg campesterol por 100 gramos de aceite; y
vii.menos de 0.5 miligramos de tocotrienoles totales por gramo de aceite.

34.La composición de aceite de triglicéridos de la realización 33, en la que la cantidad de C8 es al menos 0.3%.

35.La composición de aceite de triglicéridos de la realización 33, en la que la cantidad de C10 es al menos 2%.

36.La composición de aceite de triglicéridos de la realización 33, en la que la cantidad de C12 es al menos 2%.

37.La composición de aceite de triglicéridos de la realización 33, en la que la cantidad de C12 es al menos 4%.

38.La composición de aceite de triglicéridos de la realización 33, en la que la cantidad de C10-C14 es al menos 30%.

39.Una mezcla del aceite de la realización 33 con al menos otra composición seleccionada del grupo que consiste de soya, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de vacuno, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuate, grasa blanca selecta, manteca de cerdo, *Camelina* sativa, semillas de mostaza,

nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, linaza (lino), avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez lisa, jojoba, jatrofa, macadamia, nueces del Brasil, aguacate, petróleo, o una fracción destilada de cualquiera de los aceites anteriores.

40.El método de procesamiento del aceite de la realización 33, en el que comprende realizar una o más reacciones químicas seleccionadas del grupo que consiste de transesterificación, hidrogenación, hidrocraqueo, desoxigenación, isomerización, interesterificación, hidroxilación, hidrólisis para proporcionar ácidos grasos libres, y saponificación.

41.Un combustible, en el que es hecho a partir de hidrogenación e isomerización del aceite de la realización 33.

42.Un combustible, en el que es hecho a partir de la transesterificación del aceite de la realización 33.

43.Una composición de alcano hecha de triglicérido aislado de células del género *Prototheca*, en la que el intervalo de destilación ASTM D86 T10-T90 es al menos 25°C.

44.Una composición de éster alquilo de ácido graso hecha de triglicéridos aislados de células del género *Prototheca*, en la que la composición tiene un tiempo de mojado en frío ASTM D6751 A1 de menos de 120 segundos.

45.Una composición que comprende:

(a) polisacáridos que comprenden uno o más monosacáridos de la lista que consiste de galactosa de 20-30 por ciento mol; glucosa 55-65 por ciento mol; y manosa 5-15 por ciento mol;
(b) proteína; y
(c) ADN que comprende una secuencia de rARN 23S con al menos 75% de identidad de secuencia con una o más de SEQ ID NOs: 11-19; y
(d) un gen exógeno.

46.La composición de la realización 45, en la que el gen exógeno se selecciona de una sacarosa invertasa y una acil-ACP tioesterasa grasa.

47.La composición de la realización 46, en la que el gen exógeno es una acil-ACP tioesterasa grasa y la composición comprende además lípido con un perfil lípido de al menos 4% de C8-C14.

48.La composición de la realización 45, en la que la composición se formula para consumo como un alimento animal.

LISTADO DE SECUENCIAS

<110> SOLAZYME, INC.

<120> PRODUCCIÓN DE ACEITES ADAPTADOS EN MICROORGANISMOS HETEROTRÓFICOS

<130> 026172-005010PC

<140> PCT/US2009/066142
<141> 2009-11-30

<150> 61/219,525
<151> 2009-06-23

<150> 61/174,357
<151> 2009-04-30

<150> 61/118,994
<151> 2008-12-01

<150> 61/118,590
<151> 2008-11-28

<160> 143

<170> PatentIn versión 3.5

<210> 1
<211> 1187
<212> ADN
<213> Chlorella sp.

<400> 1

gatcagacgg gcctgacctg cgagataatc aagtgctcgt aggcaaccaa ctcagcagct 60

gcttggtgtt gggtctgcag gatagtgttg cagggcccca aggacagcag gggaacttac 120

accttgtccc cgacccagtt ttatggagtg cattgcctca agagcctagc cggagcgcta 180

ggctacatac ttgccgcacc ggtatgaggg gatatagtac tcgcactgcg ctgtctagtg 240

agatgggcag tgctgcccat aaacaactgg ctgctcagcc atttgttggc ggaccattct 300

ggggggggcca gcaatgcctg actttcgggt agggtgaaaa ctgaacaaag actaccaaaa 360

cagaatttct tcctccttgg aggtaagcgc aggccggccc gcctgcgccc acatggcgct 420

ccgaacacct ccatagctgt aagggcgcaa acatggccgg actgttgtca gcactctttc 480

atggccatac aaggtcatgt cgagattagt gctgagtaag acactatcac cccatgttcg 540

attgaagccg tgacttcatg ccaacctgcc cctgggcgta gcagacgtat gccatcatga 600

ccactagccg acatgcgctg tcttttgcca ccaaaacaac tggtacaccg ctcgaagtcg 660

tgccgcacac ctccgggagt gagtccggcg actcctcccc ggcgggccgc ggccctacct 720

gggtagggtc gccatacgcc cacgaccaaa cgacgcagga ggggattggg gtagggaatc 780

ccaaccagcc taaccaagac ggcacctata ataataggtg gggggactaa cagccctata 840

tcgcaagctt tgggtgccta tcttgagaag cacgagttgg agtggctgtg tacggtcgac 900

cctaaggtgg gtgtgccgca gcctgaaaca aagcgtctag cagctgcttc tataatgtgt 960

cagccgttgt gtttcagtta tattgtatgc tattgtttgt tcgtgctagg gtggcgcagg 1020

cccacctact gtggcgggcc attggttggt gcttgaattg cctcaccatc taaggtctga 1080

acgctcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagtg 1140

tgcaaaccag cacatagcac tcccttacat cacccagcag tacaaca 1187

<210> 2
<211> 1414
<212> ADN
<213> Chlorella ellipsoidea

<400> 2

cgctgcgcac cagggccgcc agctcgctga tgtcgctcca aatgcggtcc cccgattttt 60

tgttcttcat cttctccacc ttggtggcct tcttggccag ggccttcagc tgcatgcgca 120

cagaccgttg agctcctgat cagcatcctc aggaggccct ttgacaagca agcccctgtg 180

caagcccatt cacggggtac cagtggtgct gaggtagatg ggtttgaaaa ggattgctcg 240

gtcgattgct gctcatggaa ttggcatgtg catgcatgtt cacaatatgc caccaggctt 300

tggagcaaga gagcatgaat gccttcaggc aggttgaaag ttcctggggg tgaagaggca 360

gggccgagga ttggaggagg aaagcatcaa gtcgtcgctc atgctcatgt tttcagtcag 420

agtttgccaa gctcacagga gcagagacaa gactggctgc tcaggtgttg catcgtgtgt 480

gtggtggggg gggggggttt aatacggtac gaaatgcact tggaattccc acctcatgcc 540

agcggaccca catgcttgaa ttcgaggcct gtggggtgag aaatgctcac tctgccctcg 600

ttgctgaggt acttcaggcc gctgagctca aagtcgatgc cctgctcgtc tatcagggcc 660

tgcacctctg ggctgaccgg ctcagcctcc ttcgcgggca tggagtaggc gccggcagcg 720

ttcatgtccg ggcccagggc agcggtggtg ccataaatgt cggtgatggt ggggaggggg 780

gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc 840

ggcagcactg gtctccagcc agccagcaag tggctgttca ggaaagcggc catgttgttg 900

gtccctgcgc atgtaattcc ccagatcaaa ggagggaaca gcttggattt gatgtagtgc 960

ccaaccggac tgaatgtgcg atggcaggtc cctttgagtc tcccgaatta ctagcagggc 1020

actgtgacct aacgcagcat gccaaccgca aaaaaatgat tgacagaaaa tgaagcggtg 1080

tgtcaatatt tgctgtattt attcgtttta atcagcaacc aagttcgaaa cgcaactatc 1140

gtggtgatca agtgaacctc atcagactta cctcgttcgg caaggaaacg gaggcaccaa 1200

attccaattt gatattatcg cttgccaagc tagagctgat ctttgggaaa ccaactgcca 1260

gacagtggac tgtgatggag tgccccgagt ggtggagcct cttcgattcg gttagtcatt 1320

actaacgtga accctcagtg aagggaccat cagaccagaa agaccagatc tcctcctcga 1380

caccgagaga gtgttgcggc agtaggacga caag 1414

<210> 3
<211> 512
<212> PRT
<213> Desconocido

<220>
<223> Descripción de Desconocido: Secuencia de levadura

<400> 3

```
Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His Phe Thr Pro Asn
1               5                  10                  15

Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr Asp Glu Lys Asp
            20                  25                  30

Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn Asp Thr Val Trp
        35                  40                  45

Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp Asp Leu Thr Asn
    50                  55                  60

Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg Asn Asp Ser Gly
65                  70                  75                  80

Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn Thr Ser Gly Phe
                85                  90                  95

Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val Ala Ile Trp Thr
            100                 105                 110

Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser Tyr Ser Leu Asp
        115                 120                 125

Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro Val Leu Ala Ala
    130                 135                 140

Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp Tyr Glu Pro Ser
145                 150                 155                 160

Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp Tyr Lys Ile Glu
                165                 170                 175

Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu Glu Ser Ala Phe
            180                 185                 190

Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys Pro Gly Leu Ile
        195                 200                 205
```

```
Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr Trp Val Met Phe
    210                 215                 220

Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser Phe Asn Gln Tyr
225                 230                 235                 240

Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala Phe Asp Asn Gln
                245                 250                 255

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

Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly Ile Ala Trp Ala
        275                 280                 285

Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn Pro Trp Arg Ser
    290                 295                 300

Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr Glu Tyr Gln Ala
305                 310                 315                 320

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

Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr Asn Thr Thr Leu
            340                 345                 350

Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn Ser Thr Gly Thr
        355                 360                 365

Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr Gln Thr Ile Ser
    370                 375                 380

Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys Gly Leu Glu Asp
385                 390                 395                 400

Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser Ala Ser Ser Phe
                405                 410                 415

Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val Lys Glu Asn Pro
            420                 425                 430

Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro Phe Lys Ser Glu
        435                 440                 445

Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu Asp Gln Asn Ile
    450                 455                 460
```

Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser Thr Asn Thr Tyr
465 470 475 480

Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn Met Thr Thr Gly
485 490 495

Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val Arg Glu Val Lys
500 505 510

<210> 4
<211> 20
<212> PRT
<213> Desconocido

<220>
<223> Descripción de Desconocido: Secuencia de levadura

<400> 4

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1 5 10 15

Ile Ser Ala Ser
20

<210> 5
<211> 18
<212> PRT
<213> Desconocido

<220>
<223> Descripción de Desconocido: Señal de secreción vegetal más alta

<400> 5

Met Ala Asn Lys Ser Leu Leu Leu Leu Leu Leu Leu Gly Ser Leu Ala
1 5 10 15

Ser Gly

<210> 6
<211> 10
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Secuencia de consenso sintética

<400> 6

Met Ala Arg Leu Pro Leu Ala Ala Leu Gly
1 5 10

<210> 7
<211> 20
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Péptido sintético

<400> 7

Met Ala Asn Lys Leu Leu Leu Leu Leu Leu Leu Leu Leu Pro Leu
1 5 10 15

Ala Ala Ser Gly
20

<210> 8
<211> 2615
<212> ADN
<213> Saccharomyces cerevisiae

<400> 8

gaattcccca acatggtgga gcacgacact ctcgtctact ccaagaatat caaagataca 60

gtctcagaag accaaagggc tattgagact tttcaacaaa gggtaatatc gggaaacctc 120

ctcggattcc attgcccagc tatctgtcac ttcatcaaaa ggacagtaga aaaggaaggt 180

ggcacctaca aatgccatca ttgcgataaa ggaaaggcta tcgttcaaga tgcctctgcc 240

gacagtggtc ccaaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt 300

ccaaccacgt cttcaaagca agtggattga tgtgaacatg gtggagcacg acactctcgt 360

ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca 420

acaaagggta atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat 480

caaaaggaca gtagaaaagg aaggtggcac ctacaaatgc catcattgcg ataaaggaaa 540

ggctatcgtt caagatgcct ctgccgacag tggtcccaaa gatggacccc cacccacgag 600

gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga 660

tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc 720

tatataagga agttcatttc atttggagag gacacgctga aatcaccagt ctctctctac 780

aaatctatct ctggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 840

gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgtgc 900

actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 960

atgccaagtg gcatctgtac tttcaataca acccgaacga tactgtctgg gggacgccat 1020

tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 1080

tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 1140

atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 1200

cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca 1260

cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 1320

cgaaggtctt ttggtacgag ccctcgcaga agtggatcat gacagcggca aagtcacagg 1380

actacaagat cgaaatttac tcgtctgacg accttaaatc ctggaagctc gaatccgcgt 1440

tcgcaaacga gggctttctc ggctaccaat acgaatgccc aggcctgata gaggtcccaa 1500

cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1560

cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg 1620

catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1680

tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg 1740

agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgctc gtgaggaaat 1800

tctctctcaa cactgagtac caggccaacc cggaaaccga actcataaac ctgaaagccg 1860

aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt 1920

tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccggtaca cttgaatttg 1980

aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct 2040

ccctctggtt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt 2100

ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttgtt aaggagaacc 2160

catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt 2220

cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg 2280

gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga 2340

acatgacgac gggtgtggat aacctgttct acatcgacaa attccaggtg agggaagtca 2400

agtgagatct gtcgatcgac aagctcgagt ttctccataa taatgtgtga gtagttccca 2460

gataagggaa ttagggttcc tatagggttt cgctcatgtg ttgagcatat aagaaaccct 2520

tagtatgtat ttgtatttgt aaaatacttc tatcaataaa atttctaatt cctaaaacca 2580

aaatccagta ctaaaatcca gatcccccga attaa 2615

<210> 9
<211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 9
tgttgaagaa tgagccggcg ac 22

<210> 10
<211> 20
<212> ADN

<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 10
cagtgagcta ttacgcactc 20

<210> 11
<211> 541
<212> ADN
<213> Prototheca kruegani

<400> 11

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag 60

ccatagcgaa agcaagttta gtaagcttag gtcattcttt ttagacccga aaccgagtga 120

tctacccatg atcagggtga agtgttagta aaataacatg gaggcccgaa ccgactaatg 180

ttgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcag tacaaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggtatgtca 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

g 541

<210> 12
<211> 573
<212> ADN
<213> Prototheca wickerhamii

<400> 12

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga 60

aacctaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat 120

aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360

ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 13
<211> 541
<212> ADN
<213> Prototheca stagnora

<400> 13

tgttgaagaa tgagccggcg agttaaaaaa aatggcatgg ttaaagatat ttctctgaag 60

ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagacccga aaccgagtga 120

tctacccatg atcagggtga agtgttggtc aaataacatg gaggcccgaa ccgactaatg 180

gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggacgt gagtatgtca 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

g 541

<210> 14
<211> 541
<212> ADN
<213> Prototheca moriformis

<400> 14

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag 60

ccatagcgaa agcaagttta acaagctaaa gtcacccttt ttagacccga aaccgagtga 120

tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180

gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggtatgtta 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

g 541

<210> 15
<211> 573
<212> ADN
<213> Prototheca moriformis

<400> 15

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga 60

aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120

aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360

ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 16
<211> 573
<212> ADN
<213> Prototheca wickerhamii

<400> 16

tgttgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga 60

aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120

aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360

ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 17
<211> 541
<212> ADN
<213> Prototheca moriformis

<400> 17

tgttgaagaa tgagccggcg agttaaaaag agtggcgtgg ttaaagaaaa ttctctggaa 60

ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120

tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180

gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300

actgtttctt ttgtgggctc cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagtgag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

g 541

<210> 18
<211> 541
<212> ADN
<213> Prototheca zopfii

<400> 18

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag 60

ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120

tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180

gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300

actgtttctt tcgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

g 541

<210> 19
<211> 565
<212> ADN
<213> Prototheca moriformis

<400> 19

tgttgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc 60

cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatatttaga atctagtcat 120

tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcggtg 240

aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300

acatctagtc tatctagggg taaagcactg tttcggtgcg ggctgtgaga acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaatacaa ccaggaggtt ggcttagaag cagccatcct 540

ttaaagagtg cgtaatagct cactg 565

<210> 20
<211> 550
<212> PRT
<213> Cichorium intybus

<400> 20

Met Ser Asn Ser Ser Asn Ala Ser Glu Ser Leu Phe Pro Ala Thr Ser
1 5 10 15

Glu Gln Pro Tyr Arg Thr Ala Phe His Phe Gln Pro Pro Gln Asn Trp
20 25 30

Met Asn Asp Pro Asn Gly Pro Met Cys Tyr Asn Gly Val Tyr His Leu
35 40 45

Phe Tyr Gln Tyr Asn Pro Phe Gly Pro Leu Trp Asn Leu Arg Met Tyr
50 55 60

Trp Ala His Ser Val Ser His Asp Leu Ile Asn Trp Ile His Leu Asp
65 70 75 80

Leu Ala Phe Ala Pro Thr Glu Pro Phe Asp Ile Asn Gly Cys Leu Ser
85 90 95

Gly Ser Ala Thr Val Leu Pro Gly Asn Lys Pro Ile Met Leu Tyr Thr
100 105 110

Gly Ile Asp Thr Glu Asn Arg Gln Val Gln Asn Leu Ala Val Pro Lys
115 120 125

Asp Leu Ser Asp Pro Tyr Leu Arg Glu Trp Val Lys His Thr Gly Asn
130 135 140

Pro Ile Ile Ser Leu Pro Glu Glu Ile Gln Pro Asp Asp Phe Arg Asp
145 150 155 160

Pro Thr Thr Thr Trp Leu Glu Glu Asp Gly Thr Trp Arg Leu Leu Val
165 170 175

Gly Ser Gln Lys Asp Lys Thr Gly Ile Ala Phe Leu Tyr His Ser Gly
180 185 190

Asp Phe Val Asn Trp Thr Lys Ser Asp Ser Pro Leu His Lys Val Ser
195 200 205

Gly Thr Gly Met Trp Glu Cys Val Asp Phe Phe Pro Val Trp Val Asp
210 215 220

```
Ser Thr Asn Gly Val Asp Thr Ser Ile Ile Asn Pro Ser Asn Arg Val
225                 230                 235                 240

Lys His Val Leu Lys Leu Gly Ile Gln Asp His Gly Lys Asp Cys Tyr
                245                 250                 255

Leu Ile Gly Lys Tyr Ser Ala Asp Lys Glu Asn Tyr Val Pro Glu Asp
            260                 265                 270

Glu Leu Thr Leu Ser Thr Leu Arg Leu Asp Tyr Gly Met Tyr Tyr Ala
        275                 280                 285

Ser Lys Ser Phe Phe Asp Pro Val Lys Asn Arg Arg Ile Met Thr Ala
    290                 295                 300

Trp Val Asn Glu Ser Asp Ser Glu Ala Asp Val Ile Ala Arg Gly Trp
305                 310                 315                 320

Ser Gly Val Gln Ser Phe Pro Arg Ser Leu Trp Leu Asp Lys Asn Gln
                325                 330                 335

Lys Gln Leu Leu Gln Trp Pro Ile Glu Glu Ile Glu Met Leu His Gln
            340                 345                 350

Asn Glu Val Ser Phe His Asn Lys Lys Leu Asp Gly Gly Ser Ser Leu
        355                 360                 365

Glu Val Leu Gly Ile Thr Ala Ser Gln Ala Asp Val Lys Ile Ser Phe
    370                 375                 380

Lys Leu Ala Asn Leu Glu Glu Ala Glu Glu Leu Asp Pro Ser Trp Val
385                 390                 395                 400

Asp Pro Gln Leu Ile Cys Ser Glu Asn Asp Ala Ser Lys Lys Gly Lys
                405                 410                 415

Phe Gly Pro Phe Gly Leu Leu Ala Leu Ala Ser Ser Asp Leu Arg Glu
            420                 425                 430

Gln Thr Ala Ile Phe Phe Arg Val Phe Arg Lys Asn Gly Arg Tyr Val
        435                 440                 445

Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Met Lys Asn Gly Ile
    450                 455                 460

Glu Lys Arg Thr Tyr Gly Ala Phe Val Asp Ile Asp Pro Gln Gln Asp
```

```
465                 470                 475                 480

Glu Ile Ser Leu Arg Thr Leu Ile Asp His Ser Ile Val Glu Ser Phe
                485                 490                 495

Gly Gly Arg Gly Lys Thr Cys Ile Thr Thr Arg Val Tyr Pro Thr Leu
            500                 505                 510

Ala Ile Gly Glu Gln Ala Arg Leu Phe Ala Phe Asn His Gly Thr Glu
        515                 520                 525

Ser Val Glu Ile Ser Glu Leu Ser Ala Trp Ser Met Lys Lys Ala Gln
    530                 535                 540

Met Lys Val Glu Glu Pro
545                 550
```

<210> 21
<211> 581
<212> PRT
<213> Schizosaccharomyces pombe

<400> 21

```
Met Phe Leu Lys Tyr Ile Leu Ala Ser Gly Ile Cys Leu Val Ser Leu
1               5                  10                  15

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

Pro Val Ile Tyr Asn Ala Ser Asn Ile Thr Glu Val Ser Asn Ser Thr
        35                  40                  45

Thr Val Pro Pro Pro Pro Phe Val Asn Thr Thr Ala Pro Asn Gly Thr
    50                  55                  60

Cys Leu Gly Asn Tyr Asn Glu Tyr Leu Pro Ser Gly Tyr Tyr Asn Ala
65                  70                  75                  80

Thr Asp Arg Pro Lys Ile His Phe Thr Pro Ser Ser Gly Phe Met Asn
                85                  90                  95

Asp Pro Asn Gly Leu Val Tyr Thr Gly Gly Val Tyr His Met Phe Phe
            100                 105                 110

Gln Tyr Ser Pro Lys Thr Leu Thr Ala Gly Glu Val His Trp Gly His
        115                 120                 125

Thr Val Ser Lys Asp Leu Ile His Trp Glu Asn Tyr Pro Ile Ala Ile
    130                 135                 140
```

```
Tyr Pro Asp Glu His Glu Asn Gly Val Leu Ser Leu Pro Phe Ser Gly
145                 150                 155                 160

Ser Ala Val Val Asp Val His Asn Ser Ser Gly Leu Phe Ser Asn Asp
                165                 170                 175

Thr Ile Pro Glu Glu Arg Ile Val Leu Ile Tyr Thr Asp His Trp Thr
            180                 185                 190

Gly Val Ala Glu Arg Gln Ala Ile Ala Tyr Thr Thr Asp Gly Gly Tyr
        195                 200                 205

Thr Phe Lys Lys Tyr Ser Gly Asn Pro Val Leu Asp Ile Asn Ser Leu
    210                 215                 220

Gln Phe Arg Asp Pro Lys Val Ile Trp Asp Phe Asp Ala Asn Arg Trp
225                 230                 235                 240

Val Met Ile Val Ala Met Ser Gln Asn Tyr Gly Ile Ala Phe Tyr Ser
                245                 250                 255

Ser Tyr Asp Leu Ile His Trp Thr Glu Leu Ser Val Phe Ser Thr Ser
            260                 265                 270

Gly Tyr Leu Gly Leu Gln Tyr Glu Cys Pro Gly Met Ala Arg Val Pro
        275                 280                 285

Val Glu Gly Thr Asp Glu Tyr Lys Trp Val Leu Phe Ile Ser Ile Asn
    290                 295                 300

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

Trp Asn Gly Thr Asn Phe Val Pro Asp Asp Gly Gln Thr Arg Phe Val
                325                 330                 335

Asp Leu Gly Lys Asp Phe Tyr Ala Ser Ala Leu Tyr His Ser Ser Ser
            340                 345                 350

Ala Asn Ala Asp Val Ile Gly Val Gly Trp Ala Ser Asn Trp Gln Tyr
        355                 360                 365

Thr Asn Gln Ala Pro Thr Gln Val Phe Arg Ser Ala Met Thr Val Ala
    370                 375                 380

Arg Lys Phe Thr Leu Arg Asp Val Pro Gln Asn Pro Met Thr Asn Leu
```

```
385                 390                 395                 400

Thr Ser Leu Ile Gln Thr Pro Leu Asn Val Ser Leu Leu Arg Asp Glu
                405                 410                 415

Thr Leu Phe Thr Ala Pro Val Ile Asn Ser Ser Ser Ser Leu Ser Gly
            420                 425                 430

Ser Pro Ile Thr Leu Pro Ser Asn Thr Ala Phe Glu Phe Asn Val Thr
        435                 440                 445

Leu Ser Ile Asn Tyr Thr Glu Gly Cys Thr Thr Gly Tyr Cys Leu Gly
    450                 455                 460

Arg Ile Ile Ile Asp Ser Asp Asp Pro Tyr Arg Leu Gln Ser Ile Ser
465                 470                 475                 480

Val Asp Val Asp Phe Ala Ala Ser Thr Leu Val Ile Asn Arg Ala Lys
                485                 490                 495

Ala Gln Met Gly Trp Phe Asn Ser Leu Phe Thr Pro Ser Phe Ala Asn
            500                 505                 510

Asp Ile Tyr Ile Tyr Gly Asn Val Thr Leu Tyr Gly Ile Val Asp Asn
        515                 520                 525

Gly Leu Leu Glu Leu Tyr Val Asn Asn Gly Glu Lys Thr Tyr Thr Asn
    530                 535                 540

Asp Phe Phe Phe Leu Gln Gly Ala Thr Pro Gly Gln Ile Ser Phe Ala
545                 550                 555                 560

Ala Phe Gln Gly Val Ser Phe Asn Asn Val Thr Val Thr Pro Leu Lys
                565                 570                 575

Thr Ile Trp Asn Cys
            580
```

<210> 22
<211> 550
<212> PRT
<213> Pichia anomala

<400> 22

```
Met Ile Gln Leu Ser Pro Leu Leu Leu Leu Pro Leu Phe Ser Val Phe
1               5                  10                  15

Asn Ser Ile Ala Asp Ala Ser Thr Glu Tyr Leu Arg Pro Gln Ile His
            20                  25                  30
```

```
Leu Thr Pro Asp Gln Gly Trp Met Asn Asp Pro Asn Gly Met Phe Tyr
        35                  40                  45

Asp Arg Lys Asp Lys Leu Trp His Val Tyr Phe Gln His Asn Pro Asp
    50                  55                  60

Lys Lys Ser Ile Trp Ala Thr Pro Val Thr Trp Gly His Ser Thr Ser
65                  70                  75                  80

Lys Asp Leu Leu Thr Trp Asp Tyr His Gly Asn Ala Leu Glu Pro Glu
                85                  90                  95

Asn Asp Asp Glu Gly Ile Phe Ser Gly Ser Val Val Val Asp Arg Asn
            100                 105                 110

Asn Thr Ser Gly Phe Phe Asn Asp Ser Thr Asp Pro Glu Gln Arg Ile
        115                 120                 125

Val Ala Ile Tyr Thr Asn Asn Ala Gln Leu Gln Thr Gln Glu Ile Ala
    130                 135                 140

Tyr Ser Leu Asp Lys Gly Tyr Ser Phe Ile Lys Tyr Asp Gln Asn Pro
145                 150                 155                 160

Val Ile Asn Val Asn Ser Ser Gln Gln Arg Asp Pro Lys Val Leu Trp
                165                 170                 175

His Asp Glu Ser Asn Gln Trp Ile Met Val Val Ala Lys Thr Gln Glu
            180                 185                 190

Phe Lys Val Gln Ile Tyr Gly Ser Pro Asp Leu Lys Lys Trp Asp Leu
        195                 200                 205

Lys Ser Asn Phe Thr Ser Asn Gly Tyr Leu Gly Phe Gln Tyr Glu Cys
    210                 215                 220

Pro Gly Leu Phe Lys Leu Pro Ile Glu Asn Pro Leu Asn Asp Thr Val
225                 230                 235                 240

Thr Ser Lys Trp Val Leu Leu Leu Ala Ile Asn Pro Gly Ser Pro Leu
                245                 250                 255

Gly Gly Ser Ile Asn Glu Tyr Phe Ile Gly Asp Phe Asp Gly Thr Thr
            260                 265                 270

Phe His Pro Asp Asp Gly Ala Thr Arg Phe Met Asp Ile Gly Lys Asp
```

275 280 285

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

Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asn Thr Val Pro
305 310 315 320

Thr Glu Asn Trp Arg Ser Ser Met Ser Leu Val Arg Asn Tyr Thr Leu
325 330 335

Lys Tyr Val Asp Val Asn Pro Glu Asn Tyr Gly Leu Thr Leu Ile Gln
340 345 350

Lys Pro Val Tyr Asp Thr Lys Glu Thr Arg Leu Asn Glu Thr Leu Lys
355 360 365

Thr Leu Glu Thr Ile Asn Glu Tyr Glu Val Asn Asp Leu Lys Leu Asp
370 375 380

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

Gly Val Phe Glu Phe Asp Leu Lys Phe Thr Gln Thr Asp Leu Lys Met
405 410 415

Gly Tyr Ser Asn Met Thr Thr Gln Phe Gly Leu Tyr Ile His Ser Gln
420 425 430

Thr Val Lys Gly Ser Gln Glu Thr Leu Gln Leu Val Phe Asp Thr Leu
435 440 445

Ser Thr Thr Trp Tyr Ile Asp Arg Thr Thr Gln His Ser Phe Gln Arg
450 455 460

Asn Ser Pro Val Phe Thr Glu Arg Ile Ser Thr Tyr Val Glu Lys Ile
465 470 475 480

Asp Thr Thr Asp Gln Gly Asn Val Tyr Thr Leu Tyr Gly Val Val Asp
485 490 495

Arg Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Ser Ile Ala Met Thr
500 505 510

Asn Thr Phe Phe Phe Arg Glu Gly Lys Ile Pro Thr Ser Phe Glu Val
515 520 525

```
Val Cys Asp Ser Glu Lys Ser Phe Ile Thr Ile Asp Glu Leu Ser Val
    530                 535                 540

Arg Glu Leu Ala Arg Lys
545                 550
```

<210> 23
<211> 533
<212> PRT
<213> Debaryomyces occidentalis

<400> 23

```
Met Val Gln Val Leu Ser Val Leu Val Ile Pro Leu Leu Thr Leu Phe
1               5                  10                  15

Phe Gly Tyr Val Ala Ser Ser Ser Ile Asp Leu Ser Val Asp Thr Ser
            20                  25                  30

Glu Tyr Asn Arg Pro Leu Ile His Phe Thr Pro Glu Lys Gly Trp Met
        35                  40                  45

Asn Asp Pro Asn Gly Leu Phe Tyr Asp Lys Thr Ala Lys Leu Trp His
    50                  55                  60

Leu Tyr Phe Gln Tyr Asn Pro Asn Ala Thr Ala Trp Gly Gln Pro Leu
65                  70                  75                  80

Tyr Trp Gly His Ala Thr Ser Asn Asp Leu Val His Trp Asp Glu His
                85                  90                  95

Glu Ile Ala Ile Gly Pro Glu His Asp Asn Glu Gly Ile Phe Ser Gly
            100                 105                 110

Ser Ile Val Val Asp His Asn Asn Thr Ser Gly Phe Phe Asn Ser Ser
        115                 120                 125

Ile Asp Pro Asn Gln Arg Ile Val Ala Ile Tyr Thr Asn Asn Ile Pro
    130                 135                 140

Asp Leu Gln Thr Gln Asp Ile Ala Phe Ser Leu Asp Gly Gly Tyr Thr
145                 150                 155                 160

Phe Thr Lys Tyr Glu Asn Asn Pro Val Ile Asp Val Ser Ser Asn Gln
                165                 170                 175

Phe Arg Asp Pro Lys Val Phe Trp His Glu Arg Phe Lys Ser Met Asp
            180                 185                 190

His Gly Cys Ser Glu Ile Ala Arg Val Lys Ile Gln Ile Phe Gly Ser
```

```
                195                 200                 205

Ala Asn Leu Lys Asn Trp Val Leu Asn Ser Asn Phe Ser Ser Gly Tyr
    210                 215                 220

Tyr Gly Asn Gln Tyr Gly Met Ser Arg Leu Ile Glu Val Pro Ile Glu
225                 230                 235                 240

Asn Ser Asp Lys Ser Lys Trp Val Met Phe Leu Ala Ile Asn Pro Gly
                245                 250                 255

Ser Pro Leu Gly Gly Ser Ile Asn Gln Tyr Phe Val Gly Asp Phe Asp
            260                 265                 270

Gly Phe Gln Phe Val Pro Asp Asp Ser Gln Thr Arg Phe Val Asp Ile
        275                 280                 285

Gly Lys Asp Phe Tyr Ala Phe Gln Thr Phe Ser Glu Val Glu His Gly
    290                 295                 300

Val Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asp Gln Val
305                 310                 315                 320

Pro Thr Asn Pro Trp Arg Ser Ser Thr Ser Leu Ala Arg Asn Tyr Thr
                325                 330                 335

Leu Arg Tyr Val Ile Gln Met Leu Lys Leu Thr Ala Asn Ile Asp Lys
            340                 345                 350

Ser Val Leu Pro Asp Ser Ile Asn Val Val Asp Lys Leu Lys Lys Lys
        355                 360                 365

Asn Val Lys Leu Thr Asn Lys Lys Pro Ile Lys Thr Asn Phe Lys Gly
    370                 375                 380

Ser Thr Gly Leu Phe Asp Phe Asn Ile Thr Phe Lys Val Leu Asn Leu
385                 390                 395                 400

Asn Val Ser Pro Gly Lys Thr His Phe Asp Ile Leu Ile Asn Ser Gln
                405                 410                 415

Glu Leu Asn Ser Ser Val Asp Ser Ile Lys Ile Gly Phe Asp Ser Ser
            420                 425                 430

Gln Ser Leu Phe Tyr Ile Asp Arg His Ile Pro Asn Val Glu Phe Pro
        435                 440                 445
```

```
Arg Lys Gln Phe Phe Thr Asp Lys Leu Ala Ala Tyr Leu Glu Pro Leu
 450                 455                 460

Asp Tyr Asp Gln Asp Leu Arg Val Phe Ser Leu Tyr Gly Ile Val Asp
465                 470                 475                 480

Lys Asn Ile Ile Glu Leu Tyr Phe Asn Asp Gly Thr Val Ala Met Thr
                485                 490                 495

Asn Thr Phe Phe Met Gly Glu Gly Lys Tyr Pro His Asp Ile Gln Ile
            500                 505                 510

Val Thr Asp Thr Glu Glu Pro Leu Phe Glu Leu Glu Ser Val Ile Ile
        515                 520                 525

Arg Glu Leu Asn Lys
    530
```

<210> 24
<211> 654
<212> PRT
<213> Oryza sativa

<400> 24

```
Met Ala Thr Ser Arg Leu Thr Pro Ala Tyr Asp Leu Lys Asn Ala Ala
1               5                  10                  15

Ala Ala Val Tyr Thr Pro Leu Pro Glu Gln Pro His Ser Ala Glu Val
            20                  25                  30

Glu Ile Arg Asp Arg Lys Pro Phe Lys Ile Ile Ser Ala Ile Ile Leu
        35                  40                  45

Ser Ser Leu Leu Leu Leu Ala Leu Ile Leu Val Ala Val Asn Tyr Gln
    50                  55                  60

Ala Pro Pro Ser His Ser Ser Gly Asp Asn Ser Gln Pro Ala Ala Val
65                  70                  75                  80

Met Pro Pro Ser Arg Gly Val Ser Gln Gly Val Ser Glu Lys Ala Phe
                85                  90                  95

Arg Gly Ala Ser Gly Ala Gly Asn Gly Val Ser Phe Ala Trp Ser Asn
            100                 105                 110

Leu Met Leu Ser Trp Gln Arg Thr Ser Tyr His Phe Gln Pro Val Lys
        115                 120                 125

Asn Trp Met Asn Asp Pro Asn Gly Pro Leu Tyr Tyr Lys Gly Trp Tyr
```

```
    130                 135                 140

His Leu Phe Tyr Gln Tyr Asn Pro Asp Ser Ala Val Trp Gly Asn Ile
145                 150                 155                 160

Thr Trp Gly His Ala Val Ser Thr Asp Leu Ile Asn Trp Leu His Leu
                165                 170                 175

Pro Phe Ala Met Val Pro Asp Gln Trp Tyr Asp Val Asn Gly Val Trp
            180                 185                 190

Thr Gly Ser Ala Thr Ile Leu Pro Asp Gly Arg Ile Val Met Leu Tyr
        195                 200                 205

Thr Gly Asp Thr Asp Asp Tyr Val Gln Asp Gln Asn Leu Ala Phe Pro
    210                 215                 220

Ala Asn Leu Ser Asp Pro Leu Leu Val Asp Trp Val Lys Tyr Pro Asn
225                 230                 235                 240

Asn Pro Val Ile Tyr Pro Pro Pro Gly Ile Gly Val Lys Asp Phe Arg
                245                 250                 255

Asp Pro Thr Thr Ala Gly Thr Ala Gly Met Gln Asn Gly Gln Arg Leu
            260                 265                 270

Val Thr Ile Gly Ser Lys Val Gly Lys Thr Gly Ile Ser Leu Val Tyr
        275                 280                 285

Glu Thr Thr Asn Phe Thr Thr Phe Lys Leu Leu Tyr Gly Val Leu His
    290                 295                 300

Ala Val Pro Gly Thr Gly Met Trp Glu Cys Val Asp Leu Tyr Pro Val
305                 310                 315                 320

Ser Thr Thr Gly Glu Asn Gly Leu Asp Thr Ser Val Asn Gly Leu Gly
                325                 330                 335

Val Lys His Val Leu Lys Thr Ser Leu Asp Asp Asp Lys His Asp Tyr
            340                 345                 350

Tyr Ala Leu Gly Thr Tyr Asp Pro Val Lys Asn Lys Trp Thr Pro Asp
        355                 360                 365

Asn Pro Asp Leu Asp Val Gly Ile Gly Leu Arg Leu Asp Tyr Gly Lys
    370                 375                 380
```

```
Tyr Tyr Ala Ala Arg Thr Phe Tyr Asp Gln Asn Lys Gln Arg Arg Ile
385                 390                 395                 400

Leu Trp Gly Trp Ile Gly Glu Thr Asp Leu Glu Ala Val Asp Leu Met
                405                 410                 415

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

Lys Lys Thr Gly Thr Asn Val Leu Gln Arg Pro Glu Glu Glu Val Glu
        435                 440                 445

Ser Trp Ser Ser Gly Asp Pro Ile Thr Gln Arg Arg Ile Phe Glu Pro
    450                 455                 460

Gly Ser Val Val Pro Ile His Val Ser Gly Ala Thr Gln Leu Asp Ile
465                 470                 475                 480

Thr Ala Ser Phe Glu Val Asp Glu Thr Leu Leu Glu Thr Thr Ser Glu
                485                 490                 495

Ser His Asp Ala Gly Tyr Asp Cys Ser Asn Ser Gly Gly Ala Gly Thr
            500                 505                 510

Arg Gly Ser Leu Gly Pro Phe Gly Leu Leu Val Val Ala Asp Glu Lys
        515                 520                 525

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

Gly Lys Ala Lys Ala His Leu Cys Ala Tyr Gln Thr Arg Ser Ser Met
545                 550                 555                 560

Ala Ser Gly Val Glu Lys Glu Val Tyr Gly Ser Ala Val Pro Val Leu
                565                 570                 575

Asp Gly Glu Asn Tyr Ser Ala Arg Ile Leu Ile Asp His Ser Ile Val
            580                 585                 590

Glu Ser Phe Ala Gln Ala Gly Arg Thr Cys Val Arg Ser Arg Asp Tyr
        595                 600                 605

Pro Thr Lys Asp Ile Tyr Gly Ala Ala Arg Cys Phe Phe Phe Asn Asn
    610                 615                 620

Ala Thr Glu Ala Ser Val Arg Ala Ser Leu Lys Ala Trp Gln Met Lys
625                 630                 635                 640
```

Ser Phe Ile Arg Pro Tyr Pro Phe Ile Pro Asp Gln Lys Ser
645 650

<210> 25
<211> 690
<212> PRT
<213> Allium cepa

<400> 25

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

Pro Pro Ser Asp Glu Phe His Asp Gln Pro Pro Pro Leu Arg Ser Trp
20 25 30

Leu Arg Leu Leu Ser Ile Pro Leu Ala Leu Met Phe Leu Leu Phe Leu
35 40 45

Ala Thr Phe Leu Ser Asn Leu Glu Ser Pro Pro Ser Asp Ser Gly Leu
50 55 60

Val Ser Asp Pro Val Thr Phe Asp Val Asn Pro Ala Val Val Arg Arg
65 70 75 80

Gly Lys Asp Ala Gly Val Ser Asp Lys Thr Ser Gly Val Asp Ser Gly
85 90 95

Phe Val Leu Asp Pro Val Ala Val Asp Ala Asn Ser Val Val Val His
100 105 110

Arg Gly Lys Asp Ala Gly Val Ser Asp Lys Thr Ser Gly Val Asp Ser
115 120 125

Gly Leu Leu Lys Asp Ser Pro Leu Gly Pro Tyr Pro Trp Thr Asn Gln
130 135 140

Met Leu Ser Trp Gln Arg Thr Gly Phe His Phe Gln Pro Val Lys Asn
145 150 155 160

Trp Met Asn Asp Pro Asn Gly Pro Leu Tyr Tyr Lys Gly Trp Tyr His
165 170 175

Phe Phe Tyr Gln Tyr Asn Pro Glu Gly Ala Val Trp Gly Asn Ile Ala
180 185 190

Trp Gly His Ala Val Ser Arg Asp Leu Val His Trp Thr His Leu Pro
195 200 205

```
Leu Ala Met Val Pro Asp Gln Trp Tyr Asp Ile Asn Gly Val Trp Thr
    210                 215                 220

Gly Ser Ala Thr Ile Leu Pro Asp Gly Gln Ile Val Met Leu Tyr Thr
225                 230                 235                 240

Gly Ala Thr Asn Glu Ser Val Gln Val Gln Asn Leu Ala Val Pro Ala
                245                 250                 255

Asp Gln Ser Asp Thr Leu Leu Leu Arg Trp Lys Lys Ser Glu Ala Asn
            260                 265                 270

Pro Ile Leu Val Pro Pro Pro Gly Ile Gly Asp Lys Asp Phe Arg Asp
        275                 280                 285

Pro Thr Thr Ala Trp Tyr Glu Pro Ser Asp Asp Thr Trp Arg Ile Val
    290                 295                 300

Ile Gly Ser Lys Asp Ser Ser His Ser Gly Ile Ala Ile Val Tyr Ser
305                 310                 315                 320

Thr Lys Asp Phe Ile Asn Tyr Lys Leu Ile Pro Gly Ile Leu His Ala
                325                 330                 335

Val Glu Arg Val Gly Met Trp Glu Cys Val Asp Phe Tyr Pro Val Ala
            340                 345                 350

Thr Ala Asp Ser Ser His Ala Asn His Gly Leu Asp Pro Ser Ala Arg
        355                 360                 365

Pro Ser Pro Ala Val Lys His Val Leu Lys Ala Ser Met Asp Asp Asp
    370                 375                 380

Arg His Asp Tyr Tyr Ala Ile Gly Thr Tyr Asp Pro Ala Gln Asn Thr
385                 390                 395                 400

Trp Val Pro Asp Asp Ala Ser Val Asp Val Gly Ile Gly Leu Arg Tyr
                405                 410                 415

Asp Trp Gly Lys Phe Tyr Ala Ser Lys Thr Phe Tyr Asp His Ala Lys
            420                 425                 430

Lys Arg Arg Ile Leu Trp Ser Trp Ile Gly Glu Thr Asp Ser Glu Thr
        435                 440                 445

Ala Asp Ile Ala Lys Gly Trp Ala Ser Leu Gln Gly Val Pro Arg Thr
    450                 455                 460
```

```
Val Leu Leu Asp Val Lys Thr Gly Ser Asn Leu Ile Thr Trp Pro Val
465                 470                 475                 480

Val Glu Ile Glu Ser Leu Arg Thr Arg Pro Arg Asp Phe Ser Gly Ile
                485                 490                 495

Thr Val Asp Ala Gly Ser Thr Phe Lys Leu Asp Val Gly Gly Ala Ala
            500                 505                 510

Gln Leu Asp Ile Glu Ala Glu Phe Lys Ile Ser Ser Glu Glu Leu Glu
        515                 520                 525

Ala Val Lys Glu Ala Asp Val Ser Tyr Asn Cys Ser Ser Ser Gly Gly
    530                 535                 540

Ala Ala Glu Arg Gly Val Leu Gly Pro Phe Gly Leu Leu Val Leu Ala
545                 550                 555                 560

Asn Gln Asp Leu Thr Glu Gln Thr Ala Thr Tyr Phe Tyr Val Ser Arg
                565                 570                 575

Gly Met Asp Gly Gly Leu Asn Thr His Phe Cys Gln Asp Glu Lys Arg
            580                 585                 590

Ser Ser Lys Ala Ser Asp Ile Val Lys Arg Ile Val Gly His Ser Val
        595                 600                 605

Pro Val Leu Asp Gly Glu Ser Phe Ala Leu Arg Ile Leu Val Asp His
    610                 615                 620

Ser Ile Val Glu Ser Phe Ala Gln Gly Gly Arg Ala Ser Ala Thr Ser
625                 630                 635                 640

Arg Val Tyr Pro Thr Glu Ala Ile Tyr Asn Asn Ala Arg Val Phe Val
                645                 650                 655

Phe Asn Asn Ala Thr Gly Ala Lys Val Thr Ala Gln Ser Leu Lys Val
            660                 665                 670

Trp His Met Ser Thr Ala Ile Asn Glu Ile Tyr Asp Pro Ala Thr Ser
        675                 680                 685

Val Met
    690
```

<210> 26
<211> 501

<212> PRT
<213> Beta vulgaris

<220>
<221> MOD_RES
<222> (382)..(382)
<223> Cualquier aminoácido

<400> 26

```
Leu Phe Tyr Gln Tyr Asn Pro Asn Gly Val Ile Trp Gly Pro Pro Val
1               5                  10                  15

Trp Gly His Ser Thr Ser Lys Asp Leu Val Asn Trp Val Pro Gln Pro
            20                  25                  30

Leu Thr Met Glu Pro Glu Met Ala Ala Asn Ile Asn Gly Ser Trp Ser
        35                  40                  45

Gly Ser Ala Thr Ile Leu Pro Gly Asn Lys Pro Ala Ile Leu Phe Thr
    50                  55                  60

Gly Leu Asp Pro Lys Tyr Glu Gln Val Gln Val Leu Ala Tyr Pro Lys
65                  70                  75                  80

Asp Thr Ser Asp Pro Asn Leu Lys Glu Trp Phe Leu Ala Pro Gln Asn
                85                  90                  95

Pro Val Met Phe Pro Thr Pro Gln Asn Gln Ile Asn Ala Thr Ser Phe
            100                 105                 110

Arg Asp Pro Thr Thr Ala Trp Arg Leu Pro Asp Gly Val Trp Arg Leu
        115                 120                 125

Leu Ile Gly Ser Lys Arg Gly Gln Arg Gly Leu Ser Leu Leu Phe Arg
    130                 135                 140

Ser Arg Asp Phe Val His Trp Val Gln Ala Lys His Pro Leu Tyr Ser
145                 150                 155                 160

Asp Lys Leu Ser Gly Met Trp Glu Cys Pro Asp Phe Phe Pro Val Tyr
                165                 170                 175

Ala Asn Gly Asp Gln Met Gly Val Asp Thr Ser Ile Ile Gly Ser His
            180                 185                 190

Val Lys His Val Leu Lys Asn Ser Leu Asp Ile Thr Lys His Asp Ile
        195                 200                 205

Tyr Thr Ile Gly Asp Tyr Asn Ile Lys Lys Asp Ala Tyr Thr Pro Asp
```

210 215 220

Ile Gly Tyr Met Asn Asp Ser Ser Leu Arg Tyr Asp Tyr Gly Lys Tyr
225 230 235 240

Tyr Ala Ser Lys Thr Phe Phe Asp Asp Ala Lys Lys Glu Arg Ile Leu
245 250 255

Leu Gly Trp Ala Asn Glu Ser Ser Ser Val Glu Asp Asp Ile Lys Lys
260 265 270

Gly Trp Ser Gly Ile His Thr Ile Pro Arg Lys Ile Trp Leu Asp Lys
275 280 285

Leu Gly Lys Gln Leu Ile Gln Trp Pro Ile Ala Asn Ile Glu Lys Leu
290 295 300

Arg Gln Lys Pro Val Asn Ile Tyr Arg Lys Val Leu Lys Gly Gly Ser
305 310 315 320

Gln Ile Glu Val Ser Gly Ile Thr Ala Ala Gln Ala Asp Val Glu Ile
325 330 335

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

Trp Thr Ser Pro Gln Leu Leu Cys Ser Lys Lys Gly Ala Ser Val Lys
355 360 365

Gly Gly Leu Gly Pro Phe Gly Leu Leu Thr Leu Ala Ser Xaa Gly Leu
370 375 380

Glu Glu Tyr Thr Ala Val Phe Phe Arg Ile Phe Lys Ala Tyr Asp Asn
385 390 395 400

Lys Phe Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Leu Asn
405 410 415

Pro Thr Asn Asp Lys Thr Thr Tyr Gly Thr Phe Val Asp Val Asn Pro
420 425 430

Ile Arg Glu Gly Leu Ser Leu Arg Val Leu Ile Asp His Ser Val Val
435 440 445

Glu Ser Phe Gly Ala Lys Gly Lys Asn Val Ile Thr Ala Arg Val Tyr
450 455 460

```
Pro Thr Leu Ala Ile Asn Glu Lys Ala His Leu Tyr Val Phe Asn Arg
465                 470                 475                 480

Gly Thr Ser Asn Val Glu Ile Thr Gly Leu Thr Ala Trp Ser Met Lys
                485                 490                 495

Lys Ala Asn Ile Ala
            500
```

<210> 27
<211> 518
<212> PRT
<213> Bifidobacterium breve

<400> 27

```
Met Thr Asp Phe Thr Pro Glu Thr Pro Val Leu Thr Pro Ile Arg Asp
1               5                  10                  15

His Ala Ala Glu Leu Ala Lys Ala Glu Ala Gly Val Ala Glu Met Ala
            20                  25                  30

Ala Lys Arg Asn Asn Arg Trp Tyr Pro Lys Tyr His Ile Ala Ser Asn
        35                  40                  45

Gly Gly Trp Ile Asn Asp Pro Asn Gly Leu Cys Phe Tyr Lys Gly Arg
    50                  55                  60

Trp His Val Phe Tyr Gln Leu His Pro Tyr Gly Thr Gln Trp Gly Pro
65                  70                  75                  80

Met His Trp Gly His Val Ser Ser Thr Asp Met Leu Asn Trp Lys Arg
                85                  90                  95

Glu Pro Ile Met Phe Ala Pro Ser Leu Glu Gln Glu Lys Asp Gly Val
            100                 105                 110

Phe Ser Gly Ser Ala Val Ile Asp Asp Asn Gly Asp Leu Arg Phe Tyr
        115                 120                 125

Tyr Thr Gly His Arg Trp Ala Asn Gly His Asp Asn Thr Gly Gly Asp
    130                 135                 140

Trp Gln Val Gln Met Thr Ala Leu Pro Asp Asn Asp Glu Leu Thr Ser
145                 150                 155                 160

Ala Thr Lys Gln Gly Met Ile Ile Asp Cys Pro Thr Asp Lys Val Asp
                165                 170                 175

His His Tyr Arg Asp Pro Lys Val Trp Lys Thr Gly Asp Thr Trp Tyr
```

```
                180                 185                 190

Met Thr Phe Gly Val Ser Ser Glu Asp Lys Arg Gly Gln Met Trp Leu
        195                 200                 205

Phe Ser Ser Lys Asp Met Val Arg Trp Glu Tyr Glu Arg Val Leu Phe
    210                 215                 220

Gln His Pro Asp Pro Asp Val Phe Met Leu Glu Cys Pro Asp Phe Phe
225                 230                 235                 240

Pro Ile Lys Asp Lys Asp Gly Asn Glu Lys Trp Val Ile Gly Phe Ser
                245                 250                 255

Ala Met Gly Ser Lys Pro Ser Gly Phe Met Asn Arg Asn Val Asn Asn
            260                 265                 270

Ala Gly Tyr Met Ile Gly Thr Trp Glu Pro Gly Gly Glu Phe Lys Pro
        275                 280                 285

Glu Thr Glu Phe Arg Leu Trp Asp Cys Gly His Asn Tyr Tyr Ala Pro
    290                 295                 300

Gln Ser Phe Asn Val Asp Gly Arg Gln Ile Val Tyr Gly Trp Met Ser
305                 310                 315                 320

Pro Phe Val Gln Pro Ile Pro Met Glu Asp Asp Gly Trp Cys Gly Gln
                325                 330                 335

Leu Thr Leu Pro Arg Glu Ile Thr Leu Asp Asp Asp Gly Asp Val Val
            340                 345                 350

Thr Ala Pro Val Ala Glu Met Glu Gly Leu Arg Glu Asp Thr Leu Asp
        355                 360                 365

His Gly Ser Ile Thr Leu Asp Met Asp Gly Glu Gln Val Ile Ala Asp
    370                 375                 380

Asp Ala Glu Ala Val Glu Ile Glu Met Thr Ile Asp Leu Ala Ala Ser
385                 390                 395                 400

Thr Ala Asp Arg Ala Gly Leu Lys Ile His Ala Thr Glu Asp Gly Ala
                405                 410                 415

Tyr Thr Tyr Val Ala Tyr Asp Asp Gln Ile Gly Arg Val Val Val Asp
            420                 425                 430
```

```
Arg Gln Ala Met Ala Asn Gly Asp His Gly Tyr Arg Ala Ala Pro Leu
        435                 440                 445

Thr Asp Ala Glu Leu Ala Ser Gly Lys Leu Asp Leu Arg Val Phe Val
    450                 455                 460

Asp Arg Gly Ser Val Glu Val Tyr Val Asn Gly Gly His Gln Val Leu
465                 470                 475                 480

Ser Ser Tyr Ser Tyr Ala Ser Glu Gly Pro Arg Ala Ile Lys Leu Val
                485                 490                 495

Ala Glu Phe Gly Asn Leu Lys Val Glu Ser Leu Lys Leu His His Met
            500                 505                 510

Lys Ser Ile Gly Leu Glu
        515
```

<210> 28
<211> 532
<212> PRT
<213> Saccharomyces cerevisiae

<400> 28

```
Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1               5                  10                  15

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His
            20                  25                  30

Phe Thr Pro Asn Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr
        35                  40                  45

Asp Glu Lys Asp Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn
    50                  55                  60

Asp Thr Val Trp Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp
65                  70                  75                  80

Asp Leu Thr Asn Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg
                85                  90                  95

Asn Asp Ser Gly Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn
            100                 105                 110

Thr Ser Gly Phe Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val
        115                 120                 125

Ala Ile Trp Thr Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser
```

130 135 140

Tyr Ser Leu Asp Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro
145 150 155 160

Val Leu Ala Ala Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp
165 170 175

Tyr Glu Pro Ser Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp
180 185 190

Tyr Lys Ile Glu Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu
195 200 205

Glu Ser Ala Phe Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys
210 215 220

Pro Gly Leu Ile Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr
225 230 235 240

Trp Val Met Phe Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser
245 250 255

Phe Asn Gln Tyr Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala
260 265 270

Phe Asp Asn Gln Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala
275 280 285

Leu Gln Thr Phe Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly
290 295 300

Ile Ala Trp Ala Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn
305 310 315 320

Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr
325 330 335

Glu Tyr Gln Ala Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu
340 345 350

Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr
355 360 365

Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn
370 375 380

```
Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr
385                 390                 395                 400

Gln Thr Ile Ser Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys
                405                 410                 415

Gly Leu Glu Asp Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser
            420                 425                 430

Ala Ser Ser Phe Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val
        435                 440                 445

Lys Glu Asn Pro Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro
    450                 455                 460

Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu
465                 470                 475                 480

Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser
                485                 490                 495

Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn
            500                 505                 510

Met Thr Thr Gly Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val
        515                 520                 525

Arg Glu Val Lys
    530
```

<210> 29
<211> 512
<212> PRT
<213> Zymomonas mobilis

<400> 29

```
Met Glu Ser Pro Ser Tyr Lys Asn Leu Ile Lys Ala Glu Asp Ala Gln
1               5                  10                  15

Lys Lys Ala Gly Lys Arg Leu Leu Ser Ser Glu Trp Tyr Pro Gly Phe
            20                  25                  30

His Val Thr Pro Leu Thr Gly Trp Met Asn Asp Pro Asn Gly Leu Ile
        35                  40                  45

Phe Phe Lys Gly Glu Tyr His Leu Phe Tyr Gln Tyr Tyr Pro Phe Ala
    50                  55                  60

Pro Val Trp Gly Pro Met His Trp Gly His Ala Lys Ser Arg Asp Leu
```

65 70 75 80

Val His Trp Glu Thr Leu Pro Val Ala Leu Ala Pro Gly Asp Leu Phe
85 90 95

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

Leu Thr Leu Ile Tyr Thr Gly His Ile Val Leu Ser Asn Asp Ser Pro
115 120 125

Asp Ala Ile Arg Glu Val Gln Cys Met Ala Thr Ser Ile Asp Gly Ile
130 135 140

His Phe Gln Lys Glu Gly Ile Val Leu Glu Lys Ala Pro Met Pro Gln
145 150 155 160

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

Phe Met Val Val Gly Tyr Arg Thr Asp Asp Glu Lys His Gln Gly Ile
180 185 190

Gly His Val Ala Leu Tyr Arg Ser Glu Asn Leu Lys Asp Trp Ile Phe
195 200 205

Val Lys Thr Leu Leu Gly Asp Asn Ser Gln Leu Pro Leu Gly Lys Arg
210 215 220

Ala Phe Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asn Arg Ser
225 230 235 240

Val Leu Met Phe Ser Pro Gln Gly Leu Lys Ala Ser Gly Tyr Lys Asn
245 250 255

Arg Asn Leu Phe Gln Asn Gly Tyr Ile Leu Gly Lys Trp Gln Ala Pro
260 265 270

Gln Phe Thr Pro Glu Thr Ser Phe Gln Glu Leu Asp Tyr Gly His Asp
275 280 285

Phe Tyr Ala Ala Gln Arg Phe Glu Ala Lys Asp Gly Arg Gln Ile Leu
290 295 300

Ile Ala Trp Phe Asp Met Trp Glu Asn Gln Lys Pro Ser Gln Arg Asp
305 310 315 320

```
Gly Trp Ala Gly Cys Met Thr Leu Pro Arg Lys Leu Asp Leu Ile Asp
                325                 330                 335

Asn Lys Ile Val Met Thr Pro Val Arg Glu Met Glu Ile Leu Arg Gln
            340                 345                 350

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

Phe Thr Met Asp Ser Pro Leu Gln Glu Ile Glu Leu Ile Phe Asp Leu
    370                 375                 380

Glu Lys Ser Ser Ala Tyr Gln Ala Gly Leu Ala Leu Arg Cys Asn Gly
385                 390                 395                 400

Lys Gly Gln Glu Thr Leu Leu Tyr Ile Asp Arg Ser Gln Asn Arg Ile
                405                 410                 415

Ile Leu Asp Arg Asn Arg Ser Gly Gln Asn Val Lys Gly Ile Arg Ser
            420                 425                 430

Cys Pro Leu Pro Asn Thr Ser Lys Val Arg Leu His Ile Phe Leu Asp
        435                 440                 445

Arg Ser Ser Ile Glu Ile Phe Val Gly Asp Asp Gln Thr Gln Gly Leu
    450                 455                 460

Tyr Ser Ile Ser Ser Arg Ile Phe Pro Asp Lys Asp Ser Leu Lys Gly
465                 470                 475                 480

Arg Leu Phe Ala Ile Glu Gly Tyr Ala Val Phe Asp Ser Phe Lys Arg
                485                 490                 495

Trp Thr Leu Gln Asp Ala Asn Leu Ala Ala Phe Ser Ser Asp Ala Cys
            500                 505                 510
```

<210> 30
<211> 987
<212> ADN
<213> Chlorella luteoviridis

<400> 30

tgttgaagaa tgagccggcg acttatagga agtggcttgg ttaaggatac tttccgaagc 60

ctaagcgaaa gcaagttgta acaatagcga tatacctctt tgtaggtcag tcacttctta 120

tggacccgaa cccgggtgat ctaaccatga ccaggatgaa gcttgggtaa caccaagtga 180

aggtccgaac tcttcgatct ttaaaaatcg tgagatgagt tatggttagg ggtaaatctg 240

gcagttttgc cccgcaaaag ggtaaccttt tgtaattact gactcataac ggtgaagcct 300

aaggcgttag ctatggtaat accgtgggaa gtttcaatac cttcttgcat atttttttatt 360

tgcaccttta gtgcaaacag tgtaaagaaa gcgttttgaa acccccttaac gactaatttt 420

ttgcttttgc aagaacgtca gcactcacca atacactttc cgtttttttc ttttattaat 480

taaagcaaca taaaaatata ttttatagct ttaatcataa aactatgtta gcacttcgtg 540

ctaatgtgct aatgtgctaa tcaaatgaaa agtgttctta aaagtgagtt gaaggtagag 600

tctaatcttg cctgaaaggg caagctgcac attttttttt gaatgtgcaa caatggaaat 660

gccaatcgaa ctcggagcta gctggttctc cccgaaatgt gttgaggcgc agcgattcat 720

gattagtacg gtgtaggggt aaagcactgt ttcggtgcgg gctgtgaaaa cggtaccaaa 780

tcgtggcaaa ctaagaatac tacgcttgta taccatggat cagtgagact atgggggata 840

agctccatag tcaagaggga aacagcccag atcaccagtt aaggccccaa aatgacagct 900

aagtggcaaa ggaggtgaaa gtgcagaaac aaccaggagg tttgcccaga agcagccatc 960

ctttaaagag tgcgtaatag ctcactg 987

<210> 31
<211> 1412
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 31

gaattcgagt ttaggtccag cgtccgtggg gggggacggg ctgggagctt gggccgggaa 60

gggcaagacg atgcagtccc tctggggagt cacagccgac tgtgtgtgtt gcactgtgcg 120

gcccgcagca ctcacacgca aaatgcctgg ccgacaggca ggccctgtcc agtgcaacat 180

ccacggtccc tctcatcagg ctcaccttgc tcattgacat aacggaatgc gtaccgctct 240

ttcagatctg tccatccaga gaggggagca ggctccccac cgacgctgtc aaacttgctt 300

cctgcccaac cgaaaacatt attgtttgag gggggggggg gggggcaga ttgcatggcg 360

ggatatctcg tgaggaacat cactgggaca ctgtggaaca cagtgagtgc agtatgcaga 420

gcatgtatgc taggggtcag cgcaggaagg gggcctttcc cagtctccca tgccactgca 480

ccgtatccac gactcaccag gaccagcttc ttgatcggct tccgctcccg tggacaccag 540

tgtgtagcct ctggactcca ggtatgcgtg caccgcaaag gccagccgat cgtgccgatt 600

cctggggtgg aggatatgag tcagccaact tggggctcag agtgcacact ggggcacgat 660

acgaaacaac atctacaccg tgtcctccat gctgacacac cacagcttcg ctccacctga 720

atgtgggcgc atgggcccga atcacagcca atgtcgctgc tgccataatg tgatccagac 780

cctctccgcc cagatgccga gcggatcgtg ggcgctgaat agattcctgt ttcgatcact 840

gtttgggtcc tttccttttc gtctcggatg cgcgtctcga aacaggctgc gtcgggcttt 900

cggatccctt ttgctccctc cgtcaccatc ctgcgcgcgg gcaagttgct tgaccctggg 960

ctggtaccag ggttggaggg tattaccgcg tcaggccatt cccagcccgg attcaattca 1020

aagtctgggc caccaccctc cgccgctctg tctgatcact ccacattcgt gcatacacta 1080

cgttcaagtc ctgatccagg cgtgtctcgg gacaaggtgt gcttgagttt gaatctcaag 1140

gacccactcc agcacagctg ctggttgacc ccgccctcgc aactccctac catgtctgct 1200

ggtaggtcca gggatctttg ccatgcacac aggaccccgt ttgtgggggt ccccggtgca 1260

tgctgtcgct gtgcaggcgc cggtgtgggg cctgggcccc gcgggagctc aactcctccc 1320

catatgcctg ccgtccctcc cacccaccgc gacctggccc cctttgcaga ggaaggcgaa 1380

gtcagcgcca tcgtgtgcga taatggatcc gg 1412

<210> 32
<211> 1627
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 32

gaattcgccc ttgagtttag gtccagcgtc cgtgggggggg gcgtgagact ccccccctgac 60

cttcgtatgg cagggactcc tacttgccaa gtaatcagtt gacaatgcca cttcaatgct 120

cgttgtggta cactgacgcg ggtctaacat actgggaagc atgaattgcc gacatggact 180

cagttggaga cagtaacagc tctttgtgtt ctatcttcag gaacacattt ggcagcgcac 240

ccatacagtg gcgcacacgc agctgtacct gatgtggctc tattcccaca tgtttcaact 300

tgatccaaaa gtcactcaga ctctcagcag ctagacttga tcgcatcttt ggccatgaag 360

atgcttgcgc aactctagga atgggacgag aaaagagcct gctctgatcg gatatttcca 420

ttctctggat gggactgaga tgattctgaa gaaatgctgc tcgacttatt tggaagaaca 480

gcacctgacg catgctttga ggctgctgtg gctgggatgt gctgtatttg tcagcattga 540

gcatctacgg gtagatggcc ataaccacgc gctgcctatc atgcggtggg ttgtgtggaa 600

aacgtacaat ggacagaaat caatcccatt gcgagcctag cgtgcagcca tgcgctccct 660

ctgtagcccc gctccaagac aaagccagcc aatgccagaa cccacataga gagggtatct 720

tcctaatgac ctcgcccatc atttcctcca aattaactat aatgccttga ttgtggagtt 780

ggctttggct tgcagctgct cgcgctggca cttttgtagg cagcacaggg tatgccagcg 840

ccgaactttg tgcccttgag caggccacaa gggcacaaga ctacaccatg cagctggtat 900

acttggaact gataccattc ttaccaagca aggcacagca cagcctgcac cgactcactt 960

tgcttgagcg gggcacagcg ccgcgactga tcctgcgagc tgtggggagt tccgactgtt 1020

ctggacctcg gtctctgaaa gatgtgtacg atgggatcaa gtcattcaag tatgctcttc 1080

acatgagcaa tcgggggaga cacggtggcc ctaaaggtgt tcatctgatt caagtgtagt 1140

ggggggggtgc tgtttgtccc ggggcgcccc ccgctccccg accccggaga agggccccag 1200

aggactcggc cgcccacaga ggaataaccg ggcgtggctc ggccctgcgc ctccctcttt 1260

caatatttca cctggtgttc agtgcacgga cacgtaaaga actagataca atggccgagg 1320

gaaagacggt gagagcttgg cgttggtgga ccgggcagca tcagaaactc ctcttccccg 1380

cccgccttga aactcactgt aactccctcc tcttcccct cgcagcatct gtctatcgtt 1440

atcgtgagtg aaagggactg ccatgtgtcg ggtcgttgac cacggtcggc tcgggcgctg 1500

ctgcccgcgt cgcgaacgtt ccctgcaaac gccgcgcagc cgtccctttt tctgccgccg 1560

ccccaccccc tcgctccccc cttcaatcac accgcagtgc ggacatgtcg attccggcaa 1620

gtccacc 1627

<210> 33
<211> 570
<212> ADN
<213> Chlorella protothecoides

<400> 33

gaattccctg caggaagaag gccggcagca gctggtactt gtccttcacc tccttgatcg 60

gctgggtgag cttggccggg tcgcagtcgt cgatgccggc atcgcccagc acgctgtgcg 120

gggagccggc atcgacaacc ttggcactgc tcaccttggt caccggcatg gggtcatggc 180

gctgcagacc agcggcctgt cagcatgctg caggcatctg tgttttgtag tagatacttt 240

ctgatgcatc accacacgtt tggaaggtcc ccaagcccct tcaacagtct cgacatatga 300

cactcgcgcc ctcttcctcg tcccgtggcc tgatgagggt acgcaggtac cgcagctgcg 360

ccccgtcccg ccagttgccc tggccccgcc gggcccaatc tgttcattgc cgctccctgg 420

cagccgtgaa cttcacacta ccgctctctg tgaccttcag cacagcagga atcgccattt 480

caccggcggt cgttgctgcg gagcctcagc tgatctcgcc tgcgagaccc cacagtttga 540

atttgcggtc cccacacaac ctctgacgcc 570

<210> 34
<211> 568
<212> ADN
<213> Chlorella protothecoides

<220>
<221> base_modificada
<222> (133)..(133)
<223> a, c, t, g, desconocido u otro

<400> 34

gaattccctc aggaagaagg ccggcagcag ctggtacttg tccttcacct ccttgatcgg 60

ctgggtgagc ttcgcaggat cgcagtcgtc gatgccggca tcgcccagca cgctgtgcgg 120

ggagccggca tcnacaacct tggcactgct ccccttggtc accggcatgg ggtcatggcg 180

ctgcagccca gcggcctgtc agcatgctgc aggcatctgt gtattgtagt aggtacttcc 240

tgatgcatca acacacgttt ggaagctccc caagcccctt caacagtctc gacgtatgac 300

actcgcgccc tcttcctcgc cccgtggcct gatgagggta cgcaggtacc acagctgcgc 360

cccgtcccgc cagttgccct ggcccggccg ggcccaatct gttcattgcc gctccctggt 420

agccgtgaac tcacattacc gctctctgtg accttcagca cagcaggaat cgccatttca 480

ccggcggtcg ttgctgcgga gcctcagctg atctcgcctg cgagacccca cagtttgaat 540

ttgcggtccc cacacaacct ctgacgcc 568

<210> 35
<211> 50
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 35
tgacctaggt gattaattaa ctcgaggcag cagcagctcg gatagtatcg 50

<210> 36
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 36
ctacgagctc aagctttcca tttgtgttcc catcccacta cttcc 45

<210> 37
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 37
gatcagaatt ccgcctgcaa cgcaagggca gc 32

<210> 38
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 38
gcatactagt ggcgggacgg agagagggcg 30

<210> 39
<211> 1568
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 39

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaacgg cgcgccatat caatgattga acaagatgga ttgcacgcag 360

gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg 420

gctgctctga tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca 480

agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc 540

tggccacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg 600

actggctgct attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg 660

ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta 720

cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag 780

ccggtcttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac 840

tgttcgccag gctcaaggcg cgcatgcccg acggcgagga tctcgtcgtg acccatggcg 900

atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg 960

gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg 1020

aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg 1080

attcgcagcg catcgccttc tatcgccttc ttgacgagtt cttctaagat ctgtcgatcg 1140

acaagtgact cgaggcagca gcagctcgga tagtatcgac acactctgga cgctggtcgt 1200

gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta 1260

tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt gctagctgct 1320

tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat cgcttgcatc 1380

ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctcctgctca 1440

ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctcctggtac tgcaacctgt 1500

aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca aatggaaagc 1560

ttgagctc 1568

<210> 40
<211> 2571
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<220>
<221> base_modificada
<222> (997)..(999)
<223> a, c, t, g, desconocido u otro

<400> 40

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa 540

gcatcctgca agacacatgc tgtgcgatgc tgcgctgctg cctgctgcgc acgccgttga 600

gttggcagca gctcagccat gcactggatc aggctgggct gccactgcaa tgtggtggat 660

aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc 720

accatcagca ggagcgcatg cgaagggact ggccccatgc acgccatgcc aaaccggagc 780

gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggacg 840

catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca 900

caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggcc acccttcttg 960

cggggtcgcg ccccagccag cggtgatgcg ctgatcnnnc caaacgagtt cacattcatt 1020

tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg 1080

gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcat 1140

tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaaac gcgccacgtc 1200

agcaacctgt cccgtgggtc ccccgtgccg atgaaatcgt gtgcacgccg atcagctgat 1260

tgcccggctc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag 1320

tggcgcgcca tatcaatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg 1380

gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc 1440

gtgttccggc tgtcagcgca ggggcgcccg gttctttttg tcaagaccga cctgtccggt 1500

gccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac gacgggcgtt 1560

ccttgcgcag ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc 1620

gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc 1680

atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgccc attcgaccac 1740

caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccggtct tgtcgatcag 1800

gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag 1860

gcgcgcatgc ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat 1920

atcatggtgg aaaatggccg cttttctgga ttcatcgact gtggccggct gggtgtggcg 1980

gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa 2040

tgggctgacc gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcatcgcc 2100

ttctatcgcc ttcttgacga gttcttctaa gatctgtcga tcgacaagtg actcgaggca 2160

gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 2220

ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt 2280

gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata 2340

ccacccccag catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct 2400

acgctgtcct gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc 2460

ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 2520

ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agcttgagct c 2571

<210> 41
<211> 2550
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<220>
<221> base_modificada

<222> (997)..(999)
<223> a, c, t, g, desconocido u otro

<400> 41

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa 540

gcatcctgca agacacatgc tgtgcgatgc tgcgctgctg cctgctgcgc acgccgttga 600

gttggcagca gctcagccat gcactggatc aggctgggct gccactgcaa tgtggtggat 660

aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc 720

accatcagca ggagcgcatg cgaagggact ggccccatgc acgccatgcc aaaccggagc 780

gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggacg 840

catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca 900

caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggcc acccttcttg 960

cggggtcgcg ccccagccag cggtgatgcg ctgatcnnnc caaacgagtt cacattcatt 1020

tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg 1080

gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcat 1140

tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaaac gcgccacgtc 1200

agcaacctgt cccgtgggtc ccccgtgccg atgaaatcgt gtgcacgccg atcagctgat 1260

tgcccggctc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag 1320

tggcgcgcca tatcaatgat cgagcaggac ggcctccacg ccggctcccc cgccgcctgg 1380

gtggagcgcc tgttcggcta cgactgggcc cagcagacca tcggctgctc cgacgccgcc 1440

gtgttccgcc tgtccgccca gggccgcccc gtgctgttcg tgaagaccga cctgtccggc 1500

gccctgaacg agctgcagga cgaggccgcc cgcctgtcct ggctggccac caccggcgtg 1560

ccctgcgccg ccgtgctgga cgtggtgacc gaggccggcc gcgactggct gctgctgggc 1620

gaggtgcccg gccaggacct gctgtcctcc cacctggccc ccgccgagaa ggtgtccatc 1680

atggccgacg ccatgcgccg cctgcacacc ctggaccccg ccacctgccc cttcgaccac 1740

caggccaagc accgcatcga gcgcgcccgc acccgcatgg aggccggcct ggtggaccag 1800

gacgacctgg acgaggagca ccagggcctg gcccccgccg agctgttcgc ccgcctgaag 1860

gcccgcatgc ccgacggcga ggacctggtg gtgacccacg gcgacgcctg cctgcccaac 1920

atcatggtgg agaacggccg cttctccggc ttcatcgact gcggccgcct gggcgtggcc 1980

gaccgctacc aggacatcgc cctggccacc cgcgacatcg ccgaggagct gggcggcgag 2040

tgggccgacc gcttcctggt gctgtacggc atcgccgccc ccgactccca gcgcatcgcc 2100

ttctaccgcc tgctggacga gttcttctga ctcgaggcag cagcagctcg gatagtatcg 2160

acacactctg gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct 2220

gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg 2280

cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cacccccagc atcccttcc 2340

ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca 2400

gcgctgctcc tgctcctgct cactgccccc cgcacagcct tggtttgggc tccgcctgta 2460

ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg 2520

gatgggaaca caaatggaaa gcttgagctc 2550

<210> 42
<211> 1547
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 42

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaacgg cgcgccatat caatgatcga gcaggacggc ctccacgccg 360

gctcccccgc cgcctgggtg gagcgcctgt tcggctacga ctgggcccag cagaccatcg 420

gctgctccga cgccgccgtg ttccgcctgt ccgcccaggg ccgccccgtg ctgttcgtga 480

agaccgacct gtccggcgcc ctgaacgagc tgcaggacga ggccgcccgc ctgtcctggc 540

tggccaccac cggcgtgccc tgcgccgccg tgctggacgt ggtgaccgag gccggccgcg 600

actggctgct gctgggcgag gtgcccggcc aggacctgct gtcctcccac ctggcccccg 660

ccgagaaggt gtccatcatg gccgacgcca tgcgccgcct gcacaccctg gaccccgcca 720

cctgcccctt cgaccaccag gccaagcacc gcatcgagcg cgcccgcacc cgcatggagg 780

ccggcctggt ggaccaggac gacctggacg aggagcacca gggcctggcc cccgccgagc 840

tgttcgcccg cctgaaggcc cgcatgcccg acggcgagga cctggtggtg acccacggcg 900

acgcctgcct gcccaacatc atggtggaga acggccgctt ctccggcttc atcgactgcg 960

gccgcctggg cgtggccgac cgctaccagg acatcgccct ggccacccgc gacatcgccg 1020

aggagctggg cggcgagtgg gccgaccgct tcctggtgct gtacggcatc gccgcccccg 1080

actcccagcg catcgccttc taccgcctgc tggacgagtt cttctgactc gaggcagcag 1140

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 1200

acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 1260

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 1320

ccccagcatc cccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc 1380

tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1440

tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 1500

tgcacgggaa gtagtgggat gggaacacaa atggaaagct tgagctc 1547

<210> 43
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 43
gccgcgactg gctgctgctg g 21

<210> 44
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 44
aggtcctcgc cgtcgggcat g 21

<210> 45
<211> 1292
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 45

```
atcaaaggca tagattcaca tttgttggca ttgcagagca atcatcgcgc aggacgaaca      60
tcgctcacca agcacgtact gggcatccgg aggcctccgc aaattcctgc aacaggactc     120
gctgatcagt tcgcccaagg tctacgacgc tccctatcgg cgctagactt caacacatat     180
ttcactgtca cagcctcggc atgcatcagg cctcagtctc caccatgaag accatccagt     240
ctcggcacgc cggtcccatc ggacatgtgc agtcgggtcg ccgatcggcg gggcgcgcgg     300
gatcccgcat ggcgaccccc gtggccgcag ctaccgtcgc agcccctcgc tcggccctca     360
acctctcccc caccatcatt cgacaggagg tgctccactc cgccagcgcc cagcaactag     420
actgcgtggc ctccctggcg cccgtcttcg agtcccagat cctccccctc ctgacgcccg     480
tggacgagat gtggcagccc accgacttcc tccccgcctc gaactcggag gcattcttcg     540
accagatcgg cgacctgcgg gcgcgatcgg cggccatccc cgacgacctg ctggtctgcc     600
tggtggggga catgatcacg gaggaggccc tgcccaccta catggccatg ctgaacaccc     660
tggacgtcgt gcgcgatgag acagggcaca gccagcaccc ctacgccaag tggaccaggg     720
cttggatcgc ggaggagaac cgccatggcg acctgctgaa caagtacatg tggctgacgg     780
ggcgggtggg acatgctggc ggtggagcgc accatccagc catgctggcg gtggagcgca     840
ccatccagcg cctcatctca tcgggcatgg acccgggcac ggagaaccac ccctaccacg     900
cctttgtgtt caccagcttc caggagcgcg ccaccaagct gagccacggc tccaccgccc     960
gcctggcggt cgccgccggg gacgaggccc tggccaagat ctgcgggacc attgcgcggg    1020
acgagtcgcg ccacgaggcg gcgtacacgc ggaccatgga tgccatcttc cagcgcgacc    1080
ccagcggggc catggtggcg tttgcgcaca tgatgatgcg caagatcacc atgcccgccc    1140
acctcatgga cgacggccag cacggcgcgc gcaacggggg ggcgcaactt gttcgacgac    1200
tttgcggcag tggcggagcg ggcaggggtg tacaccgccg gcgactacat cggcatcctg    1260
cgccacctca tccggcgctg ggacgtggag gg                                  1292
```

<210> 46
<211> 364
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipétido sintético

<400> 46

```
Met His Gln Ala Ser Val Ser Thr Met Lys Thr Ile Gln Ser Arg His
1               5                  10                  15
```

```
Ala Gly Pro Ile Gly His Val Gln Ser Gly Arg Arg Ser Ala Gly Arg
            20                  25                  30

Ala Gly Ser Arg Met Ala Thr Pro Val Ala Ala Ala Thr Val Ala Ala
        35                  40                  45

Pro Arg Ser Ala Leu Asn Leu Ser Pro Thr Ile Ile Arg Gln Glu Val
    50                  55                  60

Leu His Ser Ala Ser Ala Gln Gln Leu Asp Cys Val Ala Ser Leu Ala
65                  70                  75                  80

Pro Val Phe Glu Ser Gln Ile Leu Pro Leu Leu Thr Pro Val Asp Glu
                85                  90                  95

Met Trp Gln Pro Thr Asp Phe Leu Pro Ala Ser Asn Ser Glu Ala Phe
            100                 105                 110

Phe Asp Gln Ile Gly Asp Leu Arg Ala Arg Ser Ala Ala Ile Pro Asp
        115                 120                 125

Asp Leu Leu Val Cys Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu
    130                 135                 140

Pro Thr Tyr Met Ala Met Leu Asn Thr Leu Asp Val Val Arg Asp Glu
145                 150                 155                 160

Thr Gly His Ser Gln His Pro Tyr Ala Lys Trp Thr Arg Ala Trp Ile
                165                 170                 175

Ala Glu Glu Asn Arg His Gly Asp Leu Leu Asn Lys Tyr Met Trp Leu
            180                 185                 190

Thr Gly Arg Val Gly His Ala Gly Gly Gly Ala His His Pro Ala Met
        195                 200                 205

Leu Ala Val Glu Arg Thr Ile Gln Arg Leu Ile Ser Ser Gly Met Asp
    210                 215                 220

Pro Gly Thr Glu Asn His Pro Tyr His Ala Phe Val Phe Thr Ser Phe
225                 230                 235                 240

Gln Glu Arg Ala Thr Lys Leu Ser His Gly Ser Thr Ala Arg Leu Ala
                245                 250                 255

Val Ala Ala Gly Asp Glu Ala Leu Ala Lys Ile Cys Gly Thr Ile Ala
            260                 265                 270
```

Arg Asp Glu Ser Arg His Glu Ala Ala Tyr Thr Arg Thr Met Asp Ala
275 280 285

Ile Phe Gln Arg Asp Pro Ser Gly Ala Met Val Ala Phe Ala His Met
290 295 300

Met Met Arg Lys Ile Thr Met Pro Ala His Leu Met Asp Asp Gly Gln
305 310 315 320

His Gly Ala Arg Asn Gly Gly Ala Gln Leu Val Arg Arg Leu Cys Gly
325 330 335

Ser Gly Gly Ala Gly Arg Gly Val His Arg Arg Arg Leu His Arg His
340 345 350

Pro Ala Pro Pro His Pro Ala Leu Gly Arg Gly Gly
355 360

<210> 47
<211> 1395
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 47

attatacatc ggcatcgtct caggtttcac gatctgcatg ctatctatgg gactgtgact 60

ccgccggcca ggttgtggtg cgcgagaatc ctccccgctc ctgccttctc atttccctga 120

cgggagtcgc cgctgagcac cgggcggatc atgggcgtcg gcacactcca aaccccatat 180

acatgtggtc gtgcattcac gcatagcgca cggtatgtcc cgcgacgcgc ggctcgaagc 240

cgtggccatc cgacgcgctg cacggccgag gtgagggcac gcccctccgc caatggcgcg 300

cagcccatga ccgccttcga cttccggcag tacatgcagc agcgcgccgc gctggtggac 360

gcagcgctgg acctggcagt gccgctgcag taccccgaga agatcaacga ggccatgcgg 420

tacagcctgc tggccggggg caagcgcgtg cgccccgcgc tctgcctcgc tgcctgcgag 480

ctcgtgggcg gcccccctgga ggcggccatg cccgccgcct gcgccatgga gatgatccac 540

accatgagcc tcatccacga cgacctcccc gccatggaca acgacgactt ccggcgcggc 600

cagcccgcca accacaaggc ctatggcgag gagattgcga tcctggcggg cgacgcgctg 660

ctgtcgctga gctttgagca catcgcgcgc gagacgcgag gcgtggaccc ggtgcgcgtc 720

ctggccgcca tctcggagtg gcgcgcggtg ggcagccgcg ggctggtggc ggggcaggtg 780

gtggacctgg gtttcgaggg cggcggcgtg gggctggccc cgctgcgcta catccacgag 840

cacaaaaccg cggcgctgct ggaggcggcg gtggtgtccg gcgcgctgct gggcggcgcg 900

gaggaggcgg acctggagcg cctgcgcacc tacaaccgcg ccatcggcct cgctttccag 960

gtggtggggg acatcctgga catcccgggg accagcgagg agctgggcaa gaccgcgggc 1020

aaggacctga gctcccccaa aacccccctac ccgtccctgg tggggctggc caggtccaaa 1080

aaaattgcgg acgaactgat tgaggacgcg aaaacccaac tcacccagta cgagccggcc 1140

cgagcggcgc ccctcgtaac cctggccgaa aacatttgaa accggaagaa ctgactgggg 1200

gccccccctg cccccagata cggcggggct cctccatcca gttttgggat gggaggagcg 1260

acaaccgacc ccgtaaccct gtgacgcgtt tgccttgcat acgtacgcat gccttgaaac 1320

ccatccatga ccctcaacaa tacctggttg tgtgtagctt ggtcctgaaa aaaaaaaaaa 1380

aaaaaaaaaa aaaaa 1395

<210> 48
<211> 342
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 48

Met Gly Val Gly Thr Leu Gln Thr Pro Tyr Thr Cys Gly Arg Ala Phe
1 5 10 15

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

His Pro Thr Arg Cys Thr Ala Glu Val Arg Ala Arg Pro Ser Ala Asn
35 40 45

Gly Ala Gln Pro Met Thr Ala Phe Asp Phe Arg Gln Tyr Met Gln Gln
50 55 60

Arg Ala Ala Leu Val Asp Ala Ala Leu Asp Leu Ala Val Pro Leu Gln
65 70 75 80

Tyr Pro Glu Lys Ile Asn Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly
85 90 95

Gly Lys Arg Val Arg Pro Ala Leu Cys Leu Ala Ala Cys Glu Leu Val
100 105 110

Gly Gly Pro Leu Glu Ala Ala Met Pro Ala Ala Cys Ala Met Glu Met
115 120 125

```
Ile His Thr Met Ser Leu Ile His Asp Asp Leu Pro Ala Met Asp Asn
    130                 135                 140

Asp Asp Phe Arg Arg Gly Gln Pro Ala Asn His Lys Ala Tyr Gly Glu
145                 150                 155                 160

Glu Ile Ala Ile Leu Ala Gly Asp Ala Leu Leu Ser Leu Ser Phe Glu
                165                 170                 175

His Ile Ala Arg Glu Thr Arg Gly Val Asp Pro Val Arg Val Leu Ala
            180                 185                 190

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

Gln Val Val Asp Leu Gly Phe Glu Gly Gly Gly Val Gly Leu Ala Pro
    210                 215                 220

Leu Arg Tyr Ile His Glu His Lys Thr Ala Ala Leu Leu Glu Ala Ala
225                 230                 235                 240

Val Val Ser Gly Ala Leu Leu Gly Gly Ala Glu Glu Ala Asp Leu Glu
                245                 250                 255

Arg Leu Arg Thr Tyr Asn Arg Ala Ile Gly Leu Ala Phe Gln Val Val
            260                 265                 270

Gly Asp Ile Leu Asp Ile Pro Gly Thr Ser Glu Glu Leu Gly Lys Thr
        275                 280                 285

Ala Gly Lys Asp Leu Ser Ser Pro Lys Thr Pro Tyr Pro Ser Leu Val
    290                 295                 300

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

Lys Thr Gln Leu Thr Gln Tyr Glu Pro Ala Arg Ala Ala Pro Leu Val
                325                 330                 335

Thr Leu Ala Glu Asn Ile
            340
```

<210> 49
<211> 833
<212> ADN
<213> Secuencia artificial

<220>

<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 49

cagatgccat gcgccctcgg gccgcgggcc tgagggtcca cgcagcgtcc tcggtggccc 60

agacgcacca ggccgccccc ccggcggaca ggaggttcga cgactaccag ccccgcaccg 120

ccatcctctt ccccggccag ggcgcgcaga gcgtgggcat ggcgggagag ctggcgaagg 180

ccgtccccgc cgccgcggcg ctgttcgacg ccgcctccga ccagctcggc tatgacctgc 240

tccgcgtgtg cgttgagggc cccaaggcgc gcctggacag caccgccgtc agccagcccg 300

ccatctacgt ggccagcctg gcggcggtgg agaagctgcg cgcggagggc ggggaggagg 360

cactggccgc catcgacgtc gctgccggtc tgtccttggg cgagtacacc gcgctggcct 420

ttgccggcgc cttctccttc gccgacgggc tgcgcctggt ggccctgcgc ggcgccagca 480

tgcaggccgc cgccgacgcc gcaccctcgg gcatggtctc cgtcatcggt ctgccctccg 540

acgcggtggc cgcgctgtgc gaggccgcca acgcgcaggt ggcccccgac caggccgtgc 600

gcatcgccaa ctacctctgc gacggcaact acgccgtcag cggtgggctg gagggctgcg 660

cggcggtgga gggcctggcc aaggcccaca aggcgcgcat gacggtgcgc ctggcggtgg 720

cgggcgcctt ccacaccccc ttcatgcagc cggcggtgga ggcgctgagc gcgggcgctg 780

gcggacacgc cgctggtcgc gccgcgcatc cccgtggtca gcaacgggac gcc 833

<210> 50
<211> 275
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 50

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

Ala Gln Thr His Gln Ala Ala Pro Pro Ala Asp Arg Arg Phe Asp Asp
            20                  25                  30

Tyr Gln Pro Arg Thr Ala Ile Leu Phe Pro Gly Gln Gly Ala Gln Ser
        35                  40                  45

Val Gly Met Ala Gly Glu Leu Ala Lys Ala Val Pro Ala Ala Ala Ala
    50                  55                  60

Leu Phe Asp Ala Ala Ser Asp Gln Leu Gly Tyr Asp Leu Leu Arg Val
65                  70                  75                  80

Cys Val Glu Gly Pro Lys Ala Arg Leu Asp Ser Thr Ala Val Ser Gln
```

```
                    85                  90                  95

Pro Ala Ile Tyr Val Ala Ser Leu Ala Ala Val Glu Lys Leu Arg Ala
            100                 105                 110

Glu Gly Gly Glu Glu Ala Leu Ala Ala Ile Asp Val Ala Ala Gly Leu
        115                 120                 125

Ser Leu Gly Glu Tyr Thr Ala Leu Ala Phe Ala Gly Ala Phe Ser Phe
    130                 135                 140

Ala Asp Gly Leu Arg Leu Val Ala Leu Arg Gly Ala Ser Met Gln Ala
145                 150                 155                 160

Ala Ala Asp Ala Ala Pro Ser Gly Met Val Ser Val Ile Gly Leu Pro
                165                 170                 175

Ser Asp Ala Val Ala Ala Leu Cys Glu Ala Ala Asn Ala Gln Val Ala
            180                 185                 190

Pro Asp Gln Ala Val Arg Ile Ala Asn Tyr Leu Cys Asp Gly Asn Tyr
        195                 200                 205

Ala Val Ser Gly Gly Leu Glu Gly Cys Ala Ala Val Glu Gly Leu Ala
    210                 215                 220

Lys Ala His Lys Ala Arg Met Thr Val Arg Leu Ala Val Ala Gly Ala
225                 230                 235                 240

Phe His Thr Pro Phe Met Gln Pro Ala Val Glu Ala Leu Ser Ala Gly
                245                 250                 255

Ala Gly Gly His Ala Ala Gly Arg Ala Ala His Pro Arg Gly Gln Gln
            260                 265                 270

Arg Asp Ala
        275
```

<210> 51
<211> 787
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 51

tgtccatctc cccccaccct ccatccaacc atcgtcgacg gcatgcaggc gctgtgttct 60

caccccgcgt ccctcacggc gcgtgcggta ccccatgggc gggccagccc agcacagcgg 120

gtgtccagcg ccggcccggc ctacaccggc ctgtcccggc acaccctggg ctgccccagc 180

acccccaccc tccagtcccg cgccgcggtc cagacccgcg gctcctcctc cggctccacc 240

acgcgcatga ccaccaccgc ccagcgcaag atcaaggtgg ccatcaacgg gttcggccgc 300

atcggccgcc agttcctgcg ctgcgtggag gggcgcgagg actcgctgct ggagatcgtg 360

gccgtgaacg actccggcgg cgtgaagcag gccagccacc tgctcaagta cgactccacc 420

atgggcacct tcaacgccga catcaagatc tcgggcgagg gcaccttctc cgtcaacggc 480

cgcgacatcc gcgtcgtctc ctcccgcgac cccctggccc tgccctgggg cgagctgggc 540

gtggacctgg tgatcgaggg gacgggagtg tttgtggacc gcaagggtgc cagcaagcac 600

ctgcaggcgg gggccaagaa ggtcatcatc accgcgccgg ccaagggctc cgacgtgccc 660

acctacgtca tgggcgtgaa cgcggaccag tactccaact ccgacgacat catctccaac 720

gcctcctgca ccaccaactg cctggcgccc tttgtcaagg tgctcaacga ccgcttcggc 780

atcgtga 787

<210> 52
<211> 248
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 52

```
Met Gln Ala Leu Cys Ser His Pro Ala Ser Leu Thr Ala Arg Ala Val
1               5                  10                  15

Pro His Gly Arg Ala Ser Pro Ala Gln Arg Val Ser Ser Ala Gly Pro
            20                  25                  30

Ala Tyr Thr Gly Leu Ser Arg His Thr Leu Gly Cys Pro Ser Thr Pro
        35                  40                  45

Thr Leu Gln Ser Arg Ala Ala Val Gln Thr Arg Gly Ser Ser Ser Gly
    50                  55                  60

Ser Thr Thr Arg Met Thr Thr Thr Ala Gln Arg Lys Ile Lys Val Ala
65                  70                  75                  80

Ile Asn Gly Phe Gly Arg Ile Gly Arg Gln Phe Leu Arg Cys Val Glu
                85                  90                  95

Gly Arg Glu Asp Ser Leu Leu Glu Ile Val Ala Val Asn Asp Ser Gly
            100                 105                 110

Gly Val Lys Gln Ala Ser His Leu Leu Lys Tyr Asp Ser Thr Met Gly
        115                 120                 125

Thr Phe Asn Ala Asp Ile Lys Ile Ser Gly Glu Gly Thr Phe Ser Val
    130                 135                 140

Asn Gly Arg Asp Ile Arg Val Val Ser Ser Arg Asp Pro Leu Ala Leu
145                 150                 155                 160

Pro Trp Gly Glu Leu Gly Val Asp Leu Val Ile Glu Gly Thr Gly Val
                165                 170                 175

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

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

Val Met Gly Val Asn Ala Asp Gln Tyr Ser Asn Ser Asp Asp Ile Ile
    210                 215                 220

Ser Asn Ala Ser Cys Thr Thr Asn Cys Leu Ala Pro Phe Val Lys Val
225                 230                 235                 240

Leu Asn Asp Arg Phe Gly Ile Val
                245
```

<210> 53
<211> 860
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 53

gatgttgaga atagtagctt gctgccttgt cgccatgcag agcgtgtgcg cgcagtcggt 60

ttcatgcaag ggggccttca cccagtccct gcggaccccc cgatgcagca ggagccagct 120

cgtctgccgg gctgatggca aggccggagc cttcatcaag accgtaaaga gcggtgctgc 180

cgctctggct gcctccctcc tcctgtctgg gggtgcgggc gcactgacct ttgatgagct 240

gcagggcctg acctacctgc aggtgaaggg ctctggcatc gccaacacct gccccaccct 300

gtctggcggc tcctccaaca tcaaggacct gaagagcggg acctactccg tcaacaagat 360

gtgcctggag cccacgtcct tcaaggtcaa ggaggaggca cagttcaaga acggcgaggc 420

cgactttgtg cccaccaagc tcgtcacgcg tctgacctac accctggacg agatctctgg 480

ccagatgaag atcgacggca gcggcggcgt ggagttcaag gaggaggatg gcatcgacta 540

tgctgcagtc accgtgcagc ttccgggcgg ggagcgcgtg cccttcctct tcaccatcaa 600

ggagcttgac gccaagggga ctgccgacgg cttcaagggc gagttcaccg tgccctccta 660

ccgtgggtcc tccttcctgg accccaaggg ccgcggcgcc tccaccggct acgacaacgc 720

cgtggccctg cccgccgccg gcgattccga ggagttggag aaggagaaca acaagtccac 780

caaggctctg aaggggggagg ccatcttctc catcgccaag gtggacgccg ggacagggga 840

ggtggcgggc atctttgagt 860

<210> 54
<211> 275
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 54

```
Met Gln Ser Val Cys Ala Gln Ser Val Ser Cys Lys Gly Ala Phe Thr
1               5                  10                  15

Gln Ser Leu Arg Thr Pro Arg Cys Ser Arg Ser Gln Leu Val Cys Arg
            20                  25                  30

Ala Asp Gly Lys Ala Gly Ala Phe Ile Lys Thr Val Lys Ser Gly Ala
        35                  40                  45

Ala Ala Leu Ala Ala Ser Leu Leu Leu Ser Gly Gly Ala Gly Ala Leu
    50                  55                  60

Thr Phe Asp Glu Leu Gln Gly Leu Thr Tyr Leu Gln Val Lys Gly Ser
65                  70                  75                  80

Gly Ile Ala Asn Thr Cys Pro Thr Leu Ser Gly Gly Ser Ser Asn Ile
                85                  90                  95

Lys Asp Leu Lys Ser Gly Thr Tyr Ser Val Asn Lys Met Cys Leu Glu
            100                 105                 110

Pro Thr Ser Phe Lys Val Lys Glu Glu Ala Gln Phe Lys Asn Gly Glu
        115                 120                 125

Ala Asp Phe Val Pro Thr Lys Leu Val Thr Arg Leu Thr Tyr Thr Leu
    130                 135                 140

Asp Glu Ile Ser Gly Gln Met Lys Ile Asp Gly Ser Gly Gly Val Glu
```

```
145                 150                 155                 160

Phe Lys Glu Glu Asp Gly Ile Asp Tyr Ala Ala Val Thr Val Gln Leu
                165                 170                 175

Pro Gly Gly Glu Arg Val Pro Phe Leu Phe Thr Ile Lys Glu Leu Asp
            180                 185                 190

Ala Lys Gly Thr Ala Asp Gly Phe Lys Gly Glu Phe Thr Val Pro Ser
        195                 200                 205

Tyr Arg Gly Ser Ser Phe Leu Asp Pro Lys Gly Arg Gly Ala Ser Thr
    210                 215                 220

Gly Tyr Asp Asn Ala Val Ala Leu Pro Ala Ala Gly Asp Ser Glu Glu
225                 230                 235                 240

Leu Glu Lys Glu Asn Asn Lys Ser Thr Lys Ala Leu Lys Gly Glu Ala
                245                 250                 255

Ile Phe Ser Ile Ala Lys Val Asp Ala Gly Thr Gly Glu Val Ala Gly
            260                 265                 270

Ile Phe Glu
        275
```

<210> 55
<211> 818
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 55

ataatcggaa cccagctgca cgcaccatca gtgcggcagc atgcagaccg tcgcagccag 60

ctatggcgta ttggcgccct ccggctccag cgtgacccgg ggctcgacca gcagcaagca 120

gcacttcacc accctcactc ccttttccgg cttcaggcgc ctgaatcatg tggatcgggc 180

ggggcaggcg gggtctggga gcccccagac cctgcagcag gccgtgggca aggccgtgcg 240

ccggtcgcgg ggccgcacca ccagcgccgt gcgcgtgacc cgcatgatgt ttgagcggtt 300

caccgagaag gccatcaagg tggtcatgct cgcgcaggag gaggctcgcc gtctgggcca 360

caacttcgtg gggacggagc aaatcctgct ggggttgatt ggggagtcca caggcatcgc 420

cgccaaggtc ctcaagtcga tgggcgtcac gctgaaagat gcgcgtgtgg aggtcgagaa 480

gatcatcggc cgggggagcg gctttgtggc cgtggagatc cccttcaccc cccgcgccaa 540

gcgtgtgctg gagctgtccc tggaggaggc tcgccagctc ggccacaact acattggcac 600

ggagcacatc ctgctgggcc tgctgcgcga gggtgagggc gtggcctccc gcgtgctgga 660

gaccctgggc gccgaccccc agaagatccg cactcaggtg gtacgcatgg tgggtgagtc 720

gcaggagccc gtgggcacca cggtgggcgg agggtccacc ggctccaaca agatgcccac 780

cctggaggag tacggcacca acctgaccgc ccaggccg 818

<210> 56
<211> 259
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 56

```
Met Gln Thr Val Ala Ala Ser Tyr Gly Val Leu Ala Pro Ser Gly Ser
1               5                  10                  15

Ser Val Thr Arg Gly Ser Thr Ser Ser Lys Gln His Phe Thr Thr Leu
            20                  25                  30

Thr Pro Phe Ser Gly Phe Arg Arg Leu Asn His Val Asp Arg Ala Gly
        35                  40                  45

Gln Ala Gly Ser Gly Ser Pro Gln Thr Leu Gln Gln Ala Val Gly Lys
    50                  55                  60

Ala Val Arg Arg Ser Arg Gly Arg Thr Thr Ser Ala Val Arg Val Thr
65                  70                  75                  80

Arg Met Met Phe Glu Arg Phe Thr Glu Lys Ala Ile Lys Val Val Met
                85                  90                  95

Leu Ala Gln Glu Glu Ala Arg Arg Leu Gly His Asn Phe Val Gly Thr
            100                 105                 110

Glu Gln Ile Leu Leu Gly Leu Ile Gly Glu Ser Thr Gly Ile Ala Ala
        115                 120                 125

Lys Val Leu Lys Ser Met Gly Val Thr Leu Lys Asp Ala Arg Val Glu
    130                 135                 140

Val Glu Lys Ile Ile Gly Arg Gly Ser Gly Phe Val Ala Val Glu Ile
145                 150                 155                 160

Pro Phe Thr Pro Arg Ala Lys Arg Val Leu Glu Leu Ser Leu Glu Glu
                165                 170                 175
```

Ala Arg Gln Leu Gly His Asn Tyr Ile Gly Thr Glu His Ile Leu Leu
180 185 190

Gly Leu Leu Arg Glu Gly Glu Gly Val Ala Ser Arg Val Leu Glu Thr
195 200 205

Leu Gly Ala Asp Pro Gln Lys Ile Arg Thr Gln Val Val Arg Met Val
210 215 220

Gly Glu Ser Gln Glu Pro Val Gly Thr Thr Val Gly Gly Gly Ser Thr
225 230 235 240

Gly Ser Asn Lys Met Pro Thr Leu Glu Glu Tyr Gly Thr Asn Leu Thr
245 250 255

Ala Gln Ala

<210> 57
<211> 2357
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 57

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120

gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300

caacccgcaa acggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 360

gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgtgc 420

actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 480

atgccaagtg gcatctgtac tttcaataca acccgaacga tactgtctgg gggacgccat 540

tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 600

tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 660

atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 720

cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca 780

cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 840

cgaaggtctt ttggtacgag ccctcgcaga agtggatcat gacagcggca aagtcacagg 900

actacaagat cgaaatttac tcgtctgacg accttaaatc ctggaagctc gaatccgcgt 960

tcgcaaacga gggctttctc ggctaccaat acgaatgccc aggcctgata gaggtcccaa 1020

cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1080

cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg 1140

catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1200

tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg 1260

agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgctc gtgaggaaat 1320

tctctctcaa cactgagtac caggccaacc cggaaaccga actcataaac ctgaaagccg 1380

aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt 1440

tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccggtaca cttgaatttg 1500

aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct 1560

ccctctggtt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt 1620

ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttgtt aaggagaacc 1680

catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt 1740

cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg 1800

gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga 1860

acatgacgac gggtgtggat aacctgttct acatcgacaa attccaggtg agggaagtca 1920

agtgagatct gtcgatcgac aagctcgagg cagcagcagc tcggatagta tcgacacact 1980

ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata 2040

tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc 2100

gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt 2160

catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc 2220

tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct 2280

ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga 2340

acacaaatgg aaagctt 2357

<210> 58
<211> 2335
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 58

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120

gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300

caacccgcaa acggcgcgcc atgctgctgc aggccttcct gttcctgctg gccggcttcg 360

ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca 420

cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag aaggacgcca 480

agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg cccttgttct 540

ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc 600

cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct 660

ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc tggacctaca 720

acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca 780

ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg 840

tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca 900

agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca 960

acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc 1020

aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg 1080

gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc gaggccttcg 1140

acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag accttcttca 1200

acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac tgggagtact 1260

ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc aagttctccc 1320

tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga 1380

tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga 1440

aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg 1500

tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct 1560

ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgcgt 1620

cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aacccctact 1680

tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac ctgtcctact 1740

acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg 1800

tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga 1860

cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgat 1920

taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg 1980

tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat 2040

caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt 2100

gtgctatttg cgaataccac ccccagcatc cccttccctc gtttcatatc gcttgcatcc 2160

caaccgcaac ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac 2220

tgcccctcgc acagccttgg tttgggctcc gcctgtattc tcctggtact gcaacctgta 2280

aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atgga 2335

<210> 59
<211> 382
<212> PRT
<213> Cinnamomum camphorum

<400> 59

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu Pro Asp Trp Ser
50 55 60

Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln
65 70 75 80

Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Asn Pro Pro Gln Leu Leu
85 90 95

Asp Asp His Phe Gly Pro His Gly Leu Val Phe Arg Arg Thr Phe Ala
100 105 110

Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr Ser Ile Val Ala
115 120 125

Val Met Asn His Leu Gln Glu Ala Ala Leu Asn His Ala Lys Ser Val
130 135 140

Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu Met Ser Lys Arg
145 150 155 160

```
Asp Leu Ile Trp Val Val Lys Arg Thr His Val Ala Val Glu Arg Tyr
                165                 170                 175

Pro Ala Trp Gly Asp Thr Val Glu Val Glu Cys Trp Val Gly Ala Ser
            180                 185                 190

Gly Asn Asn Gly Arg Arg His Asp Phe Leu Val Arg Asp Cys Lys Thr
        195                 200                 205

Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val Met Met Asn Thr
    210                 215                 220

Arg Thr Arg Arg Leu Ser Lys Ile Pro Glu Glu Val Arg Gly Glu Ile
225                 230                 235                 240

Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp Glu Glu Ile Lys
                245                 250                 255

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

Leu Thr Pro Arg Trp Asn Asp Leu Asp Ile Asn Gln His Val Asn Asn
        275                 280                 285

Ile Lys Tyr Val Asp Trp Ile Leu Glu Thr Val Pro Asp Ser Ile Phe
    290                 295                 300

Glu Ser His His Ile Ser Ser Phe Thr Ile Glu Tyr Arg Arg Glu Cys
305                 310                 315                 320

Thr Met Asp Ser Val Leu Gln Ser Leu Thr Thr Val Ser Gly Gly Ser
                325                 330                 335

Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln Leu Glu Gly Gly
            340                 345                 350

Ser Glu Val Leu Arg Ala Lys Thr Glu Trp Arg Pro Lys Leu Thr Asp
        355                 360                 365

Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Ser Ser Val
    370                 375                 380
```

<210> 60
<211> 1240
<212> ADN
<213> Cinnamomum camphorum

<400> 60

ggcgcgccat ggccaccacc tccctggcct ccgccttctg cagcatgaag gccgtgatgc 60

tggcccgcga cggccgcggc atgaagcccc gctccagcga cctgcagctg cgcgccggca 120

acgcccagac ctccctgaag atgatcaacg gcaccaagtt ctcctacacc gagagcctga 180

agaagctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc tccgccgccg 240

agaagcagtg gaccaacctg gagtggaagc ccaagcccaa ccccccccag ctgctggacg 300

accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc agctacgagg 360

tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag gaggccgccc 420

tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc ctggagatgt 480

ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag cgctaccccg 540

cctggggcga caccgtggag gtggagtgct gggtgggcgc ctccggcaac aacggccgcc 600

gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc tgcacctccc 660

tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag gaggtgcgcg 720

gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgaggag atcaagaagc 780

cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga 840

acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg atcctggaga 900

ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc gagtaccgcc 960

gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc ggctcctccg 1020

aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag gtgctgcgcg 1080

ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg 1140

ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac cacgacatcg 1200

actacaagga cgacgacgac aagtgactcg agttaattaa 1240

<210> 61
<211> 415
<212> PRT
<213> Cuphea hookeriana

<400> 61

Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

```
Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65                  70                  75                  80

Pro Pro Arg Thr Phe Leu His Gln Leu Pro Asp Trp Ser Arg Leu Leu
                85                  90                  95

Thr Ala Ile Thr Thr Val Phe Val Lys Ser Lys Arg Pro Asp Met His
            100                 105                 110

Asp Arg Lys Ser Lys Arg Pro Asp Met Leu Val Asp Ser Phe Gly Leu
        115                 120                 125

Glu Ser Thr Val Gln Asp Gly Leu Val Phe Arg Gln Ser Phe Ser Ile
    130                 135                 140

Arg Ser Tyr Glu Ile Gly Thr Asp Arg Thr Ala Ser Ile Glu Thr Leu
145                 150                 155                 160

Met Asn His Leu Gln Glu Thr Ser Leu Asn His Cys Lys Ser Thr Gly
                165                 170                 175

Ile Leu Leu Asp Gly Phe Gly Arg Thr Leu Glu Met Cys Lys Arg Asp
            180                 185                 190

Leu Ile Trp Val Val Ile Lys Met Gln Ile Lys Val Asn Arg Tyr Pro
        195                 200                 205

Ala Trp Gly Asp Thr Val Glu Ile Asn Thr Arg Phe Ser Arg Leu Gly
    210                 215                 220

Lys Ile Gly Met Gly Arg Asp Trp Leu Ile Ser Asp Cys Asn Thr Gly
225                 230                 235                 240

Glu Ile Leu Val Arg Ala Thr Ser Ala Tyr Ala Met Met Asn Gln Lys
                245                 250                 255

Thr Arg Arg Leu Ser Lys Leu Pro Tyr Glu Val His Gln Glu Ile Val
            260                 265                 270

Pro Leu Phe Val Asp Ser Pro Val Ile Glu Asp Ser Asp Leu Lys Val
        275                 280                 285

His Lys Phe Lys Val Lys Thr Gly Asp Ser Ile Gln Lys Gly Leu Thr
    290                 295                 300

Pro Gly Trp Asn Asp Leu Asp Val Asn Gln His Val Ser Asn Val Lys
305                 310                 315                 320
```

```
Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Thr Glu Val Leu Glu Thr
                325                 330                 335

Gln Glu Leu Cys Ser Leu Ala Leu Glu Tyr Arg Arg Glu Cys Gly Arg
            340                 345                 350

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

Val Arg Ser Gln Tyr Gln His Leu Leu Arg Leu Glu Asp Gly Thr Ala
    370                 375                 380

Ile Val Asn Gly Ala Thr Glu Trp Arg Pro Lys Asn Ala Gly Ala Asn
385                 390                 395                 400

Gly Ala Ile Ser Thr Gly Lys Thr Ser Asn Gly Asn Ser Val Ser
                405                 410                 415
```

<210> 62
<211> 1339
<212> ADN
<213> Cuphea hookeriana

<400> 62

ggcgcgccat ggtggccgcc gccgcctcca gcgccttctt ccccgtgccc gcccccggcg 60

cctcccccaa gcccggcaag ttcggcaact ggccctccag cctgagcccc tccttcaagc 120

ccaagtccat ccccaacggc ggcttccagg tgaaggccaa cgacagcgcc caccccaagg 180

ccaacggctc cgccgtgagc ctgaagagcg gcagcctgaa cacccaggag gacacctcct 240

ccagcccccc cccccgcacc ttcctgcacc agctgcccga ctggagccgc ctgctgaccg 300

ccatcaccac cgtgttcgtg aagtccaagc gccccgacat gcacgaccgc aagtccaagc 360

gccccgacat gctggtggac agcttcggcc tggagtccac cgtgcaggac ggcctggtgt 420

tccgccagtc cttctccatc cgctcctacg agatcggcac cgaccgcacc gccagcatcg 480

agaccctgat gaaccacctg caggagacct ccctgaacca ctgcaagagc accggcatcc 540

tgctggacgg cttcggccgc accctggaga tgtgcaagcg cgacctgatc tgggtggtga 600

tcaagatgca gatcaaggtg aaccgctacc ccgcctgggg cgacaccgtg gagatcaaca 660

cccgcttcag ccgcctgggc aagatcggca tgggccgcga ctggctgatc tccgactgca 720

acaccggcga gatcctggtg cgcgccacca gcgcctacgc catgatgaac cagaagaccc 780

gccgcctgtc caagctgccc tacgaggtgc accaggagat cgtgcccctg ttcgtggaca 840

gccccgtgat cgaggactcc gacctgaagg tgcacaagtt caaggtgaag accggcgaca 900

gcatccagaa gggcctgacc cccggctgga acgacctgga cgtgaaccag cacgtgtcca 960

acgtgaagta catcggctgg atcctggaga gcatgcccac cgaggtgctg gagacccagg 1020

agctgtgctc cctggccctg gagtaccgcc gcgagtgcgg ccgcgactcc gtgctggaga 1080

gcgtgaccgc catggacccc agcaaggtgg gcgtgcgctc ccagtaccag cacctgctgc 1140

gcctggagga cggcaccgcc atcgtgaacg gcgccaccga gtggcgcccc aagaacgccg 1200

gcgccaacgg cgccatctcc accggcaaga ccagcaacgg caactccgtg tccatggact 1260

acaaggacca cgacggcgac tacaaggacc acgacatcga ctacaaggac gacgacgaca 1320

agtgactcga gttaattaa 1339

<210> 63
<211> 382
<212> PRT
<213> Umbellularia sp.

<400> 63

```
Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1               5                  10                  15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
            20                  25                  30

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
        35                  40                  45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu Pro Asp Trp Ser
    50                  55                  60

Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln
65                  70                  75                  80

Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Lys Leu Pro Gln Leu Leu
                85                  90                  95

Asp Asp His Phe Gly Leu His Gly Leu Val Phe Arg Arg Thr Phe Ala
            100                 105                 110

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

Val Met Asn His Met Gln Glu Ala Thr Leu Asn His Ala Lys Ser Val
    130                 135                 140

Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu Met Ser Lys Arg
145                 150                 155                 160

Asp Leu Met Trp Val Val Arg Arg Thr His Val Ala Val Glu Arg Tyr
```

```
                    165                 170                 175

Pro Thr Trp Gly Asp Thr Val Glu Val Glu Cys Trp Ile Gly Ala Ser
            180                 185                 190

Gly Asn Asn Gly Met Arg Arg Asp Phe Leu Val Arg Asp Cys Lys Thr
        195                 200                 205

Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val Leu Met Asn Thr
    210                 215                 220

Arg Thr Arg Arg Leu Ser Thr Ile Pro Asp Glu Val Arg Gly Glu Ile
225                 230                 235                 240

Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp Asp Glu Ile Lys
                245                 250                 255

Lys Leu Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr Ile Gln Gly Gly
            260                 265                 270

Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln His Val Asn Asn
        275                 280                 285

Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro Asp Ser Ile Phe
    290                 295                 300

Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr Arg Arg Glu Cys
305                 310                 315                 320

Thr Arg Asp Ser Val Leu Arg Ser Leu Thr Thr Val Ser Gly Gly Ser
                325                 330                 335

Ser Glu Ala Gly Leu Val Cys Asp His Leu Leu Gln Leu Glu Gly Gly
            340                 345                 350

Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro Lys Leu Thr Asp
        355                 360                 365

Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro Arg Val
    370                 375                 380
```

<210> 64
<211> 1240
<212> ADN
<213> Umbellularia sp.

<400> 64

ggcgcgccat ggccaccacc agcctggcct ccgccttctg ctccatgaag gccgtgatgc 60

tggcccgcga cggccgcggc atgaagcccc gcagctccga cctgcagctg cgcgccggca 120

acgcccccac ctccctgaag atgatcaacg gcaccaagtt cagctacacc gagagcctga 180

agcgcctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc agcgccgccg 240

agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag ctgctggacg 300

accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc tcctacgagg 360

tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc 420

tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc ctggagatgt 480

ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag cgctacccca 540

cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac aacggcatgc 600

gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc tgcacctccc 660

tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac gaggtgcgcg 720

gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag atcaagaagc 780

tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga 840

acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga 900

ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg gagtaccgcc 960

gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg 1020

aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag gtgctgcgcg 1080

cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg 1140

ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac cacgacatcg 1200

actacaagga cgacgacgac aagtgactcg agttaattaa 1240

<210> 65
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 65
ccgccgtgct ggacgtggtg 20

<210> 66
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 66

ggtggcgggg tccagggtgt 20

<210> 67
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 67
cggccggcgg ctccttcaac 20

<210> 68
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 68
ggcgctcccg taggtcgggt 20

<210> 69
<211> 1335
<212> ADN
<213> Chlorella sorokiniana

<400> 69

cgcctgcaac gcaagggcag ccacagccgc tcccacccgc cgctgaaccg acacgtgctt 60

gggcgcctgc cgcctgcctg ccgcatgctt gtgctggtga ggctgggcag tgctgccatg 120

ctgattgagg cttggttcat cgggtggaag cttatgtgtg tgctgggctt gcatgccggg 180

caatgcgcat ggtggcaaga gggcggcagc acttgctgga gctgccgcgg tgcctccagg 240

tggttcaatc gcggcagcca gagggatttc agatgatcgc gcgtacaggt tgagcagcag 300

tgtcagcaaa ggtagcagtt tgccagaatg atcggttcag ctgttaatca atgccagcaa 360

gagaaggggt caagtgcaaa cacgggcatg ccacagcacg ggcaccgggg agtggaatgg 420

caccaccaag tgtgtgcgag ccagcatcgc cgcctggctg tttcagctac aacggcagga 480

gtcatccaac gtaaccatga gctgatcaac actgcaatca tcgggcgggc gtgatgcaag 540

catgcctggc gaagacacat ggtgtgcgga tgctgccggc tgctgcctgc tgcgcacgcc 600

gttgagttgg cagcaggctc agccatgcac tggatggcag ctgggctgcc actgcaatgt 660

ggtggatagg atgcaagtgg agcgaatacc aaaccctctg gctgcttgct gggttgcatg 720

gcatcgcacc atcagcagga gcgcatgcga agggactggc cccatgcacg ccatgccaaa 780

ccggagcgca ccgagtgtcc acactgtcac caggcccgca agctttgcag aaccatgctc 840

atggacgcat gtagcgctga cgtcccttga cggcgctcct ctcgggtgtg ggaaacgcaa 900

tgcagcacag gcagcagagg cggcggcagc agagcggcgg cagcagcggc gggggccacc 960

cttcttgcgg ggtcgcgccc cagccagcgg tgatgcgctg atcccaaacg agttcacatt 1020

catttgcatg cctggagaag cgaggctggg gcctttgggc tggtgcagcc cgcaatggaa 1080

tgcgggaccg ccaggctagc agcaaaggcg cctcccctac tccgcatcga tgttccatag 1140

tgcattggac tgcatttggg tggggcggcc ggctgtttct ttcgtgttgc aaaacgcgcc 1200

agctcagcaa cctgtcccgt gggtcccccg tgccgatgaa atcgtgtgca cgccgatcag 1260

ctgattgccc ggctcgcgaa gtaggcgccc tcctttctgc tcgccctctc tccgtcccgc 1320

cactagtggc gcgcc 1335

<210> 70
<211> 1146
<212> ADN
<213> Umbellularia californica

<400> 70

atggccacca ccagcctggc ctccgccttc tgctccatga aggccgtgat gctggcccgc 60

gacggccgcg gcatgaagcc ccgcagctcc gacctgcagc tgcgcgccgg caacgccccc 120

acctccctga agatgatcaa cggcaccaag ttcagctaca ccgagagcct gaagcgcctg 180

cccgactggt ccatgctgtt cgccgtgatc accaccatct tcagcgccgc cgagaagcag 240

tggaccaacc tggagtggaa gcccaagccc aagctgcccc agctgctgga cgaccacttc 300

ggcctgcacg gcctggtgtt ccgccgcacc ttcgccatcc gctcctacga ggtgggcccc 360

gaccgcagca cctccatcct ggccgtgatg aaccacatgc aggaggccac cctgaaccac 420

gccaagagcg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcgc 480

gacctgatgt gggtggtgcg ccgcacccac gtggccgtgg agcgctaccc cacctggggc 540

gacaccgtgg aggtggagtg ctggatcggc gccagcggca acaacggcat gcgccgcgac 600

ttcctggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg 660

ctgatgaaca cccgcacccg ccgcctgagc accatccccg acgaggtgcg cggcgagatc 720

ggccccgcct tcatcgacaa cgtggccgtg aaggacgacg agatcaagaa gctgcagaag 780

ctgaacgact ccaccgccga ctacatccag ggcggcctga ccccccgctg gaacgacctg 840

gacgtgaacc agcacgtgaa caacctgaag tacgtggcct gggtgttcga gaccgtgccc 900

gacagcatct tcgagtccca ccacatcagc tccttcaccc tggagtaccg ccgcgagtgc 960

acccgcgact ccgtgctgcg cagcctgacc accgtgagcg gcggcagctc cgaggccggc 1020

ctggtgtgcg accacctgct gcagctggag ggcggcagcg aggtgctgcg cgcccgcacc 1080

gagtggcgcc ccaagctgac cgactccttc cgcggcatca gcgtgatccc cgccgagccc 1140

cgcgtg 1146

<210> 71
<211> 1146
<212> ADN
<213> Cinnamomum camphora

<400> 71

atggccacca cctccctggc ctccgccttc tgcagcatga aggccgtgat gctggcccgc 60

gacggccgcg gcatgaagcc ccgctccagc gacctgcagc tgcgcgccgg caacgcccag 120

acctccctga agatgatcaa cggcaccaag ttctcctaca ccgagagcct gaagaagctg 180

cccgactggt ccatgctgtt cgccgtgatc accaccatct tctccgccgc cgagaagcag 240

tggaccaacc tggagtggaa gcccaagccc aacccccccc agctgctgga cgaccacttc 300

ggcccccacg gcctggtgtt ccgccgcacc ttcgccatcc gcagctacga ggtgggcccc 360

gaccgctcca ccagcatcgt ggccgtgatg aaccacctgc aggaggccgc cctgaaccac 420

gccaagtccg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcgc 480

gacctgatct gggtggtgaa gcgcacccac gtggccgtgg agcgctaccc cgcctggggc 540

gacaccgtgg aggtggagtg ctgggtgggc gcctccggca acaacggccg ccgccacgac 600

ttcctggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg 660

atgatgaaca cccgcacccg ccgcctgagc aagatccccg aggaggtgcg cggcgagatc 720

ggccccgcct tcatcgacaa cgtggccgtg aaggacgagg agatcaagaa gccccagaag 780

ctgaacgact ccaccgccga ctacatccag ggcggcctga ccccccgctg gaacgacctg 840

gacatcaacc agcacgtgaa caacatcaag tacgtggact ggatcctgga gaccgtgccc 900

gacagcatct tcgagagcca ccacatctcc tccttcacca tcgagtaccg ccgcgagtgc 960

accatggaca gcgtgctgca gtccctgacc accgtgagcg gcggctcctc cgaggccggc 1020

ctggtgtgcg agcacctgct gcagctggag ggcggcagcg aggtgctgcg cgccaagacc 1080

gagtggcgcc ccaagctgac cgactccttc cgcggcatca gcgtgatccc cgccgagtcc 1140

agcgtg 1146

<210> 72
<211> 72
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Oligonucleótido sintético

<400> 72

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 60

gacgacaagt ga 72

<210> 73
<211> 408
<212> ADN
<213> Chlorella vulgaris

<400> 73

ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 60

ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 120

cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 180

ttgcgaatac caccccagc atccccttcc ctcgtttcat atcgcttgca tcccaaccgc 240

aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccccct 300

cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc 360

actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatgga 408

<210> 74
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 74
ctgggcgacg gcttcggcac 20

<210> 75
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 75
aagtcgcggc gcatgccgtt 20

<210> 76
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 76
taccccgcct ggggcgacac 20

<210> 77
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Cebador sintético

<400> 77

cttgctcagg cggcgggtgc 20

<210> 78
<211> 1317
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 78

atggtggccg ccgccgcctc cagcgccttc ttccccgtgc ccgcccccgg cgcctccccc 60

aagcccggca agttcggcaa ctggccctcc agcctgagcc cctccttcaa gcccaagtcc 120

atccccaacg gcggcttcca ggtgaaggcc aacgacagcg cccaccccaa ggccaacggc 180

tccgccgtga gcctgaagag cggcagcctg aacacccagg aggacacctc ctccagcccc 240

ccccccccgca ccttcctgca ccagctgccc gactggagcc gcctgctgac cgccatcacc 300

accgtgttcg tgaagtccaa gcgccccgac atgcacgacc gcaagtccaa gcgccccgac 360

atgctggtgg acagcttcgg cctggagtcc accgtgcagg acggcctggt gttccgccag 420

tccttctcca tccgctccta cgagatcggc accgaccgca ccgccagcat cgagaccctg 480

atgaaccacc tgcaggagac ctccctgaac cactgcaaga gcaccggcat cctgctggac 540

ggcttcggcc gcaccctgga gatgtgcaag cgcgacctga tctgggtggt gatcaagatg 600

cagatcaagg tgaaccgcta ccccgcctgg ggcgacaccg tggagatcaa cacccgcttc 660

agccgcctgg gcaagatcgg catgggccgc gactggctga tctccgactg caacaccggc 720

gagatcctgg tgcgcgccac cagcgcctac gccatgatga accagaagac ccgccgcctg 780

tccaagctgc cctacgaggt gcaccaggag atcgtgcccc tgttcgtgga cagccccgtg 840

atcgaggact ccgacctgaa ggtgcacaag ttcaaggtga agaccggcga cagcatccag 900

aagggcctga cccccggctg gaacgacctg gacgtgaacc agcacgtgtc caacgtgaag 960

tacatcggct ggatcctgga gagcatgccc accgaggtgc tggagaccca ggagctgtgc 1020

tccctggccc tggagtaccg ccgcgagtgc ggccgcgact ccgtgctgga gagcgtgacc 1080

gccatggacc ccagcaaggt gggcgtgcgc tcccagtacc agcacctgct gcgcctggag 1140

gacggcaccg ccatcgtgaa cggcgccacc gagtggcgcc ccaagaacgc cggcgccaac 1200

ggcgccatct ccaccggcaa gaccagcaac ggcaactccg tgtccatgga ctacaaggac 1260

cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtga 1317

<210> 79
<211> 1170
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 79

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcccagctg 120

cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc 180

gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag 240

tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc 300

ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg 360

aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc 420

aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc 480

tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggcatgggc 540

cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tggtgcgcgc caccagcgcc 600

tacgccatga tgaaccagaa gacccgccgc ctgtccaagc tgccctacga ggtgcaccag 660

gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac 720

aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac 780

ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg 840

cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag 900

tgcggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg 960

cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc 1020

accgagtggc gccccaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc 1080

aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1140

atcgactaca aggacgacga cgacaagtga 1170

<210> 80
<211> 1170
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 80

atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60

cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgcccagctg 120

cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc 180

gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag 240

tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc 300

ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg 360

aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc 420

aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc 480

tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggcatgggc 540

cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tggtgcgcgc caccagcgcc 600

tacgccatga tgaaccagaa gacccgccgc ctgtccaagc tgccctacga ggtgcaccag 660

gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac 720

aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac 780

ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg 840

cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag 900

tgcggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg 960

cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc 1020

accgagtggc gccccaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc 1080

aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1140

atcgactaca aggacgacga cgacaagtga 1170

<210> 81
<211> 1167
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 81

atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg 60

gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggcgcgc ccagctgccc 120

gactggagcc gcctgctgac cgccatcacc accgtgttcg tgaagtccaa gcgccccgac 180

atgcacgacc gcaagtccaa gcgccccgac atgctggtgg acagcttcgg cctggagtcc 240

accgtgcagg acggcctggt gttccgccag tccttctcca tccgctccta cgagatcggc 300

accgaccgca ccgccagcat cgagaccctg atgaaccacc tgcaggagac ctccctgaac 360

cactgcaaga gcaccggcat cctgctggac ggcttcggcc gcaccctgga gatgtgcaag 420

cgcgacctga tctgggtggt gatcaagatg cagatcaagg tgaaccgcta ccccgcctgg 480

ggcgacaccg tggagatcaa cacccgcttc agccgcctgg gcaagatcgg catgggccgc 540

gactggctga tctccgactg caacaccggc gagatcctgg tgcgcgccac cagcgcctac 600

gccatgatga accagaagac ccgccgcctg tccaagctgc cctacgaggt gcaccaggag 660

atcgtgcccc tgttcgtgga cagccccgtg atcgaggact ccgacctgaa ggtgcacaag 720

ttcaaggtga agaccggcga cagcatccag aagggcctga cccccggctg gaacgacctg 780

gacgtgaacc agcacgtgtc caacgtgaag tacatcggct ggatcctgga gagcatgccc 840

accgaggtgc tggagaccca ggagctgtgc tccctggccc tggagtaccg ccgcgagtgc 900

ggccgcgact ccgtgctgga gagcgtgacc gccatggacc ccagcaaggt gggcgtgcgc 960

tcccagtacc agcacctgct gcgcctggag gacggcaccg ccatcgtgaa cggcgccacc 1020

gagtggcgcc ccaagaacgc cggcgccaac ggcgccatct ccaccggcaa gaccagcaac 1080

ggcaactccg tgtccatgga ctacaaggac cacgacggcg actacaagga ccacgacatc 1140

gactacaagg acgacgacga caagtga 1167

<210> 82
<211> 1149
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 82

atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg 60

gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggcgcgc ccccgactgg 120

tccatgctgt tcgccgtgat caccaccatc ttcagcgccg ccgagaagca gtggaccaac 180

ctggagtgga agcccaagcc caagctgccc cagctgctgg acgaccactt cggcctgcac 240

ggcctggtgt tccgccgcac cttcgccatc cgctcctacg aggtgggccc cgaccgcagc 300

acctccatcc tggccgtgat gaaccacatg caggaggcca ccctgaacca cgccaagagc 360

gtgggcatcc tgggcgacgg cttcggcacc accctggaga tgtccaagcg cgacctgatg 420

tgggtggtgc gccgcaccca cgtggccgtg gagcgctacc ccacctgggg cgacaccgtg 480

gaggtggagt gctggatcgg cgccagcggc aacaacggca tgcgccgcga cttcctggtg 540

cgcgactgca agaccggcga gatcctgacc cgctgcacct ccctgagcgt gctgatgaac 600

acccgcaccc gccgcctgag caccatcccc gacgaggtgc gcggcgagat cggccccgcc 660

ttcatcgaca acgtggccgt gaaggacgac gagatcaaga agctgcagaa gctgaacgac 720

tccaccgccg actacatcca gggcggcctg acccccccgct ggaacgacct ggacgtgaac 780

cagcacgtga acaacctgaa gtacgtggcc tgggtgttcg agaccgtgcc cgacagcatc 840

ttcgagtccc accacatcag ctccttcacc ctggagtacc gccgcgagtg cacccgcgac 900

tccgtgctgc gcagcctgac caccgtgagc ggcggcagct ccgaggccgg cctggtgtgc 960

gaccacctgc tgcagctgga gggcggcagc gaggtgctgc gcgcccgcac cgagtggcgc 1020

cccaagctga ccgactcctt ccgcggcatc agcgtgatcc ccgccgagcc ccgcgtgatg 1080

gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac 1140

gacaagtga 1149

<210> 83
<211> 1146
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 83

atggcttccg cggcattcac catgtcggcg tgccccgcga tgactggcag ggcccctggg 60

gcacgtcgct ccggacggcc agtcgccacc cgcctgaggg ggcgcgcccc cgactggtcc 120

atgctgttcg ccgtgatcac caccatcttc agcgccgccg agaagcagtg gaccaacctg 180

gagtggaagc ccaagcccaa gctgccccag ctgctggacg accacttcgg cctgcacggc 240

ctggtgttcc gccgcacctt cgccatccgc tcctacgagg tgggccccga ccgcagcacc 300

tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc tgaaccacgc caagagcgtg 360

ggcatcctgg gcgacggctt cggcaccacc ctggagatgt ccaagcgcga cctgatgtgg 420

gtggtgcgcc gcacccacgt ggccgtggag cgctacccca cctggggcga caccgtggag 480

gtggagtgct ggatcggcgc cagcggcaac aacggcatgc gccgcgactt cctggtgcgc 540

gactgcaaga ccggcgagat cctgacccgc tgcacctccc tgagcgtgct gatgaacacc 600

cgcacccgcc gcctgagcac catccccgac gaggtgcgcg gcgagatcgg ccccgccttc 660

atcgacaacg tggccgtgaa ggacgacgag atcaagaagc tgcagaagct gaacgactcc 720

accgccgact acatccaggg cggcctgacc ccccgctgga acgacctgga cgtgaaccag 780

cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga ccgtgcccga cagcatcttc 840

gagtcccacc acatcagctc cttcaccctg gagtaccgcc gcgagtgcac ccgcgactcc 900

gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg aggccggcct ggtgtgcgac 960

cacctgctgc agctggaggg cggcagcgag gtgctgcgcg cccgcaccga gtggcgcccc 1020

aagctgaccg actccttccg cggcatcagc gtgatccccg ccgagccccg cgtgatggac 1080

tacaaggacc acgacggcga ctacaaggac cacgacatcg actacaagga cgacgacgac 1140

aagtga 1146

<210> 84
<211> 1155
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 84

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccccgac 120

tggtccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccaagctg ccccagctgc tggacgacca cttcggcctg 240

cacggcctgg tgttccgccg caccttcgcc atccgctcct acgaggtggg ccccgaccgc 300

agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag 360

agcgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420

atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgct accccacctg gggcgacacc 480

gtggaggtgg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg 540

gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg 600

aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc 660

gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720

gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg 780

aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgacagc 840

atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc 900

gactccgtgc tgcgcagcct gaccaccgtg agcggcggca gctccgaggc cggcctggtg 960

tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccg caccgagtgg 1020

cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg 1080

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140

gacgacaagt gatga 1155

<210> 85
<211> 1152
<212> ADN
<213> Secuencia artificial

<220>

<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 85

atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60

cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgccccgac 120

tggtccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccaagctg ccccagctgc tggacgacca cttcggcctg 240

cacggcctgg tgttccgccg caccttcgcc atccgctcct acgaggtggg ccccgaccgc 300

agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag 360

agcgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420

atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgct accccacctg gggcgacacc 480

gtggaggtgg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg 540

gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg 600

aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc 660

gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720

gactccaccg ccgactacat ccagggcggc ctgacccccc gctggaacga cctggacgtg 780

aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgacagc 840

atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc 900

gactccgtgc tgcgcagcct gaccaccgtg agcggcggca gctccgaggc cggcctggtg 960

tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccg caccgagtgg 1020

cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg 1080

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140

gacgacaagt ga 1152

<210> 86
<211> 1155
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 86

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccccgac 120

tggtccatgc tgttcgccgt gatcaccacc atcttctccg ccgccgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccaacccc ccccagctgc tggacgacca cttcggcccc 240

cacggcctgg tgttccgccg caccttcgcc atccgcagct acgaggtggg ccccgaccgc 300

tccaccagca tcgtggccgt gatgaaccac ctgcaggagg ccgccctgaa ccacgccaag 360

tccgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420

atctgggtgg tgaagcgcac ccacgtggcc gtggagcgct accccgcctg gggcgacacc 480

gtggaggtgg agtgctgggt gggcgcctcc ggcaacaacg gccgccgcca cgacttcctg 540

gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg 600

aacacccgca cccgccgcct gagcaagatc cccgaggagg tgcgcggcga gatcggcccc 660

gccttcatcg acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac 720

gactccaccg ccgactacat ccagggcggc ctgaccccgc gctggaacga cctggacatc 780

aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgacagc 840

atcttcgaga gccaccacat ctcctccttc accatcgagt accgccgcga gtgcaccatg 900

gacagcgtgc tgcagtccct gaccaccgtg agcggcggct cctccgaggc cggcctggtg 960

tgcgagcacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgccaa gaccgagtgg 1020

cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gtccagcgtg 1080

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140

gacgacaagt gatga 1155

<210> 87
<211> 1893
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 87

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 360

cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 420

cccgtgcgcg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 480

tccgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccccag 540

ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc 600

agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag 660

gaggccgccc tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc 720

ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag 780

cgctaccccg cctggggcga caccgtggag gtggagtgct gggtgggcgc ctccggcaac 840

aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 900

tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag 960

gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgaggag 1020

atcaagaagc cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 1080

ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg 1140

atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc 1200

gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc 1260

ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag 1320

gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 1380

gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac 1440

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 1500

atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 1560

ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 1620

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 1680

tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 1740

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 1800

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 1860

aagtagtggg atgggaacac aaatggaaag ctt 1893

<210> 88
<211> 1887
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 88

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcaccatgtc ggcgtgcccc 360

gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc cacccgcctg 420

aggggcgcg cccccgactg gtccatgctg ttcgccgtga tcaccaccat cttctccgcc 480

gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaacccccc ccagctgctg 540

gacgaccact tcggccccca cggcctggtg ttccgccgca ccttcgccat ccgcagctac 600

gaggtgggcc ccgaccgctc caccagcatc gtggccgtga tgaaccacct gcaggaggcc 660

gccctgaacc acgccaagtc cgtgggcatc ctgggcgacg gcttcggcac caccctggag 720

atgtccaagc gcgacctgat ctgggtggtg aagcgcaccc acgtggccgt ggagcgctac 780

cccgcctggg gcgacaccgt ggaggtggag tgctgggtgg gcgcctccgg caacaacggc 840

cgccgccacg acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgctgcacc 900

tccctgagcg tgatgatgaa cacccgcacc cgccgcctga gcaagatccc cgaggaggtg 960

cgcggcgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga ggagatcaag 1020

aagccccaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc 1080

tggaacgacc tggacatcaa ccagcacgtg aacaacatca agtacgtgga ctggatcctg 1140

gagaccgtgc ccgacagcat cttcgagagc caccacatct cctccttcac catcgagtac 1200

cgccgcgagt gcaccatgga cagcgtgctg cagtccctga ccaccgtgag cggcggctcc 1260

tccgaggccg gcctggtgtg cgagcacctg ctgcagctgg agggcggcag cgaggtgctg 1320

cgcgccaaga ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgatc 1380

cccgccgagt ccagcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1440

atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta 1500

tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga 1560

cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac 1620

gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatccccc 1680

tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc 1740

tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct 1800

gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag 1860

tgggatggga acacaaatgg aaagctt 1887

<210> 89
<211> 3631
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 89

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgccact agtatgctgc tgcaggcctt cctgttcctg ctggccggct tcgccgccaa 1380

gatcagcgcc tccatgacga acgagacgtc cgaccgcccc ctggtgcact tcacccccaa 1440

caagggctgg gggcgcgcca gccaccacgt gtacaagcgc ctgacccaga gcaccaacac 1500

caagtccccc agcgtgaacc agccctaccg caccggcttc cacttccagc cccccaagaa 1560

ctggatgaac gaccccaacg gccccatgat ctacaagggc atctaccacc tgttctacca 1620

gtggaacccc aagggcgccg tgtggggcaa catcgtgtgg gcccactcca ccagcaccga 1680

cctgatcaac tgggaccccc acccccccgc catcttcccc agcgcccccc tcgacatcaa 1740

cggctgctgg tccggcagcg ccaccatcct gcccaacggc aagcccgtga tcctgtacac 1800

cggcatcgac cccaagaacc agcaggtgca gaacatcgcc gagcccaaga acctgtccga 1860

cccctacctg cgcgagtgga agaagagccc cctgaacccc ctgatggccc ccgacgccgt 1920

gaacggcatc aacgcctcca gcttccgcga ccccaccacc gcctggctgg gccaggacaa 1980

gaagtggcgc gtgatcatcg gctccaagat ccaccgccgc ggcctggcca tcacctacac 2040

cagcaaggac ttcctgaagt gggagaagtc ccccgagccc ctgcactacg acgacggcag 2100

cggcatgtgg gagtgccccg acttcttccc cgtgacccgc ttcggcagca acggcgtgga 2160

gacctccagc ttcggcgagc ccaacgagat cctgaagcac gtgctgaaga tctccctgga 2220

cgacaccaag cacgactact acaccatcgg cacctacgac cgcgtgaagg acaagttcgt 2280

gcccgacaac ggcttcaaga tggacggcac cgcccccccgc tacgactacg gcaagtacta 2340

cgccagcaag accttcttcg actccgccaa gaaccgccgc atcctgtggg gctggaccaa 2400

cgagtcctcc agcgtggagg acgacgtgga gaagggctgg tccggcatcc agaccatccc 2460

ccgcaagatc tggctggacc gcagcggcaa gcagctgatc cagtggcccg tgcgcgaggt 2520

ggagcgcctg cgcaccaagc aggtgaagaa cctgcgcaac aaggtgctga agtccggcag 2580

ccgcctggag gtgtacggcg tgaccgccgc ccaggccgac gtggaggtgc tgttcaaggt 2640

gcgcgacctg gagaaggccg acgtgatcga gccctcctgg accgaccccc agctgatctg 2700

cagcaagatg aacgtgtccg tgaagtccgg cctgggcccc ttcggcctga tggtgctggc 2760

cagcaagaac ctggaggagt acacctccgt gtacttccgc atcttcaagg cccgccagaa 2820

cagcaacaag tacgtggtgc tgatgtgctc cgaccagtcc cgcagctccc tgaaggagga 2880

caacgacaag accacctacg gcgccttcgt ggacatcaac ccccaccagc ccctgagcct 2940

gcgcgccctg atcgaccact ccgtggtgga gagcttcggc ggcaagggcc gcgcctgcat 3000

cacctcccgc gtgtacccca agctggccat cggcaagtcc agccacctgt tcgccttcaa 3060

ctacggctac cagtccgtgg acgtgctgaa cctgaacgcc tggagcatga actccgccca 3120

gatcagcatg gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa 3180

ggacgacgac gacaagtgat taattaaccg gctcgaggca gcagcagctc ggatagtatc 3240

gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc tgccttgacc 3300

tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc 3360

gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag catcccccttc 3420

cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct gctatccctc 3480

agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg ctccgcctgt 3540

attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg 3600

ggatgggaac acaaatggaa agcttgagct c 3631

<210> 90
<211> 621
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polipéptido sintético

<400> 90

```
Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1               5                  10                  15

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His
            20                  25                  30

Phe Thr Pro Asn Lys Gly Trp Gly Arg Ala Ser His His Val Tyr Lys
        35                  40                  45

Arg Leu Thr Gln Ser Thr Asn Thr Lys Ser Pro Ser Val Asn Gln Pro
    50                  55                  60

Tyr Arg Thr Gly Phe His Phe Gln Pro Pro Lys Asn Trp Met Asn Asp
65                  70                  75                  80

Pro Asn Gly Pro Met Ile Tyr Lys Gly Ile Tyr His Leu Phe Tyr Gln
                85                  90                  95

Trp Asn Pro Lys Gly Ala Val Trp Gly Asn Ile Val Trp Ala His Ser
            100                 105                 110

Thr Ser Thr Asp Leu Ile Asn Trp Asp Pro His Pro Pro Ala Ile Phe
        115                 120                 125

Pro Ser Ala Pro Phe Asp Ile Asn Gly Cys Trp Ser Gly Ser Ala Thr
    130                 135                 140

Ile Leu Pro Asn Gly Lys Pro Val Ile Leu Tyr Thr Gly Ile Asp Pro
145                 150                 155                 160

Lys Asn Gln Gln Val Gln Asn Ile Ala Glu Pro Lys Asn Leu Ser Asp
                165                 170                 175

Pro Tyr Leu Arg Glu Trp Lys Lys Ser Pro Leu Asn Pro Leu Met Ala
            180                 185                 190

Pro Asp Ala Val Asn Gly Ile Asn Ala Ser Ser Phe Arg Asp Pro Thr
        195                 200                 205
```

```
Thr Ala Trp Leu Gly Gln Asp Lys Lys Trp Arg Val Ile Ile Gly Ser
    210                 215                 220

Lys Ile His Arg Arg Gly Leu Ala Ile Thr Tyr Thr Ser Lys Asp Phe
225                 230                 235                 240

Leu Lys Trp Glu Lys Ser Pro Glu Pro Leu His Tyr Asp Asp Gly Ser
                245                 250                 255

Gly Met Trp Glu Cys Pro Asp Phe Phe Pro Val Thr Arg Phe Gly Ser
            260                 265                 270

Asn Gly Val Glu Thr Ser Ser Phe Gly Glu Pro Asn Glu Ile Leu Lys
        275                 280                 285

His Val Leu Lys Ile Ser Leu Asp Asp Thr Lys His Asp Tyr Tyr Thr
    290                 295                 300

Ile Gly Thr Tyr Asp Arg Val Lys Asp Lys Phe Val Pro Asp Asn Gly
305                 310                 315                 320

Phe Lys Met Asp Gly Thr Ala Pro Arg Tyr Asp Tyr Gly Lys Tyr Tyr
                325                 330                 335

Ala Ser Lys Thr Phe Phe Asp Ser Ala Lys Asn Arg Arg Ile Leu Trp
            340                 345                 350

Gly Trp Thr Asn Glu Ser Ser Ser Val Glu Asp Asp Val Glu Lys Gly
        355                 360                 365

Trp Ser Gly Ile Gln Thr Ile Pro Arg Lys Ile Trp Leu Asp Arg Ser
    370                 375                 380

Gly Lys Gln Leu Ile Gln Trp Pro Val Arg Glu Val Glu Arg Leu Arg
385                 390                 395                 400

Thr Lys Gln Val Lys Asn Leu Arg Asn Lys Val Leu Lys Ser Gly Ser
                405                 410                 415

Arg Leu Glu Val Tyr Gly Val Thr Ala Ala Gln Ala Asp Val Glu Val
            420                 425                 430

Leu Phe Lys Val Arg Asp Leu Glu Lys Ala Asp Val Ile Glu Pro Ser
        435                 440                 445

Trp Thr Asp Pro Gln Leu Ile Cys Ser Lys Met Asn Val Ser Val Lys
```

```
          450                 455                 460

Ser Gly Leu Gly Pro Phe Gly Leu Met Val Leu Ala Ser Lys Asn Leu
465                 470                 475                 480

Glu Glu Tyr Thr Ser Val Tyr Phe Arg Ile Phe Lys Ala Arg Gln Asn
                485                 490                 495

Ser Asn Lys Tyr Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser
            500                 505                 510

Leu Lys Glu Asp Asn Asp Lys Thr Thr Tyr Gly Ala Phe Val Asp Ile
        515                 520                 525

Asn Pro His Gln Pro Leu Ser Leu Arg Ala Leu Ile Asp His Ser Val
    530                 535                 540

Val Glu Ser Phe Gly Gly Lys Gly Arg Ala Cys Ile Thr Ser Arg Val
545                 550                 555                 560

Tyr Pro Lys Leu Ala Ile Gly Lys Ser Ser His Leu Phe Ala Phe Asn
                565                 570                 575

Tyr Gly Tyr Gln Ser Val Asp Val Leu Asn Leu Asn Ala Trp Ser Met
            580                 585                 590

Asn Ser Ala Gln Ile Ser Met Asp Tyr Lys Asp His Asp Gly Asp Tyr
        595                 600                 605

Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
    610                 615                 620
```

<210> 91
<211> 997
<212> ADN
<213> Prototheca moriformis

<400> 91

cctgtcgatc gaagagaagg agacatgtgt acattattgg tgtgagggcg ctgaatcggc 60

catttttaa aatgatcacg ctcatgccaa tagacgcggc acataacgac gttcaaaccc 120

ccgccaaagc cgcggacaac cccatccctc cacacccccc acacaaagaa cccgccaccg 180

cttaccttgc ccacgaggta ggcctttcgt tgcgcaaaac cggcctcggt gatgaatgca 240

tgcccgttcc tgacgagcgc tgcccgggcc aacacgctct tttgctgcgt ctcctcaggc 300

ttgggggcct ccttgggctt gggtgccgcc atgatctgcg cgcatcagag aaacgttgct 360

ggtaaaaagg agcgcccggc tgcgcaatat atatataggc atgccaacac agcccaacct 420

cactcgggag cccgtcccac cacccccaag tcgcgtgcct tgacggcata ctgctgcaga 480

agcttcatga gaatgatgcc gaacaagagg ggcacgagga cccaatcccg gacatccttg 540

tcgataatga tctcgtgagt ccccatcgtc cgcccgacgc tccggggagc ccgccgatgc 600

tcaagacgag agggccctcg accaggaggg gctggcccgg gcgggcactg gcgtcgaagg 660

tgcgcccgtc gttcgcctgc agtcctatgc cacaaaacaa gtcttctgac ggggtgcgtt 720

tgctcccgtg cgggcaggca acagaggtat tcaccctggt catggggaga tcggcgatcg 780

agctgggata agagatactt ctggcaagca atgacaactt gtcaggaccg gaccgtgcca 840

tatatttctc acctagcgcc gcaaaaccta acaatttggg agtcactgtg ccactgagtt 900

cgactggtag ctgaatggag tcgctgctcc actaaacgaa ttgtcagcac cgccagccgg 960

ccgaggaccc gagtcatagc gagggtagta gcgcgcc 997

<210> 92
<211> 753
<212> ADN
<213> Prototheca moriformis

<400> 92

actaattgca atcgtgcagt aatcatcgat atggtcacaa gtagatcccc tactgacacc 60

ctctcgtaca tgtaggcaat gtcatcggcg ccgtcctgct gaccgatgcc gacgtagcag 120

agcagacccg ggccgatctg ggatacgagc cggccctcca cctgcgctcg aggtggaatc 180

aagtaaataa ccaatacact tttcgacacc acacagagtt gcacggacgg tggcgtacct 240

ctacgctcgc gctcttcacg cgctggacga ccgcacgcat gagcccgggt ggcttggtct 300

gggctgcaaa aatgcacaac aaacaagtat cagacgctca tggatgcaca cgcgctccca 360

agcacgctca gactaaatat tacagtagct cgtatctgat aagatatcga gacataccgc 420

tcaactcacc cgcaaactgc gccccgccag gtgatgcgca cagggcccca ccatgcgatc 480

catcgcatcg ctcctcgagg gcgctatcac gtggccggag agcgttcaca gcgtacgcca 540

ctgtatctgg gcggtatgcg gtccgtcaac atggagacag atacccgcac caccaccttg 600

caagctcttc catattggaa gtagaaaatt gtaattgtat catcgcacga ggggccaact 660

tgccgtcggc gagctgggcg acgaacacca cctggacgtt gtcgagactc gctcgtgccg 720

tgcgccgggc cgctgggtat ccagaccgtc gcc 753

<210> 93
<211> 1122
<212> ADN
<213> Prototheca moriformis

<400> 93

caacgacaac cagcaggcaa ctcggtcagc gacccaacac gcgagtcaaa ttgttgcgtg 60

ttcttgcctt gtctatttac tgtgatagca agactgtcgg tcagtcaata ccgcggtgcg 120

cacgtcgggg tgccaagcct agcagagcac gggacggctg gtgctgtgcg ccagctcagc 180

tcgcttcgcg accaattgta ggaccggcaa agtcaccaaa acatgccagc ggtgcgattc 240

aattggtcat gagctctaca aaattgtttt gtgcgtcgcg caggtatcca acggcgcggc 300

agagaaagtt tgacagctct cgatttcatc tcggaaaaat ggggagaatt tatgacacac 360

aagtgcgcag gcggcccagg cggccagcat attctggcgt gacctgggcc gcccacaaaa 420

tgcttggatg cactctaaaa taattatatt tgccatgaac aagggaagag ttaccgcacc 480

cagccctaga cttgggcgcc cgagcaaggt tacgtcaagc caccttcgcc catcgcccaa 540

ctccgtattc cccgacagcc gcacgtggcc ctcgccggaa tgaaccctga atcggcatca 600

cgccacgcgt tcgccaatcg ttccgctctc tggcttcatc ggcctgcgcc ttcacgtcgt 660

ggtcacgaca gtgcattcat acttccattt gcacctcggc acacactttt acgcatcgcc 720

taccttgct gcggcagtct agggtcactt tgcagccatg ggacagtgct acaccaccgt 780

cggtgcgcaa agctatttca agtgaaccgt gggcggaaaa aaggaatgta cactgtctca 840

accgactcct acaattgttt accatgcaga tcagagctcg acggccatca tcgagcaggt 900

gtggggcctt ggtggcgcgg cgcggggccc cagggcgtcg caggcattga tggcactctg 960

agactttcgc acgcgcatga gggaccccat caagagaaga gtgtgtcttt atgtccccat 1020

tcatgatgat gtatcttgtg attgtcgcag tttggcaagt ttaaccggat cgccgctcca 1080

ggtgtggcgt ggcggatttt tctaggggtg cttgagcagt cg 1122

<210> 94
<211> 574
<212> ADN
<213> Prototheca moriformis

<400> 94

ggcccagggc cctgcggatg gcccacacca gatctagcct ctcttatgcc atgcccgcct 60

cgctgcccgt cgtatccccc cgccgatccg cgcgtagggg accgcggcct gacccacgcc 120

acgaaagagc tttgctcctc aatttctcgc caacagaacc gtatcaaacg ctcaacgcct 180

atcccgaaca atccgtattc acaccaaatc gagtataccg gactggtttg cctagtcttg 240

aaggaaatga tcccgtccat gctcggaagg gggagcgggc ggaggatcct actcatctct 300

gaaatgggat tggtccgaag atgggttggg caagcacgtg ccaaacccca gcgagttgct 360

gacgagcagg ctcatccaat cccccggcga atcctccctc acgccccgca tgcatacaag 420

tccctcccac acgccccctc ccatccattt tcgcctggtc cgaacgcgag cggcgtcgag 480

gcggaccact tgctccgcag cgccgtctgg gtctccaccc cacagcggct ttgctgccag 540

aggcaccccc cttgccccac ctcctcttgc agcc 574

<210> 95
<211> 1096
<212> ADN
<213> Prototheca moriformis

<400> 95

ccaggcaggc ggtagggttg ccgattgctt gagcgaattg gaagatataa ttttttgtgg 60

tgtccctgga cgctgtttgt ggcgctcctt tttggagaag attgcgtggg ggagctttcc 120

atgtaccacg cttccttctg aaaggattct ggccgagtcc tgatgagccc aaagaaaaca 180

cctgcctttc agtgctggca ctctgaaaac gtcaacagat gattatacat gtcacaaaag 240

gcagccgatt aggaacggga gctctggccg ttcgtttggc tgcctgggct gattgaagtg 300

atccaccctg ttcgaatgaa ggcggtcgag tcgaattatc gaccggagct gtcgggaagg 360

cgtccggggc agagtgaggt gctgcggcct ggttgtcgtt caaaaagacc ccggtagccc 420

aacaatcacg aacgaaagga atataattgc ttgcatacta tacattcagt ttctatgtgg 480

cgggtagaca agtctcatgg gcttctaaag gctgtccctt gaaggctact tataaaaact 540

tgctgcgcca tggcacggat cgcgcttgcg caggctgcaa ccctgcgcgc aaggtcaaat 600

acacagcaaa agatactaac agaatttcta aaaacattta aatatttgtt tcgaccagcc 660

aattgtggtc gtaggcacgc aaaagacttt gttttgcgcc caccgagcat ccacgctggc 720

agtcaagcca gtccgatgtg cattgcgtgg cagcatcgag gagcatcaaa aacctcgtgc 780

acgcttttct gtcaatcatc atcaaccact ccaccatgta tacccgatgc atcgcggtgc 840

gcagcgcgcc acgcgtccca gacccgccca aaaacccagc agcggcgaaa gcaaatcttc 900

acttgcccga aaccccgagc agcggcattc acacgtgggc gaaaacccca cttgccctaa 960

caggcgtatg tctgctgtca cgatgcctga caacggtatt atagatatac actgattaat 1020

gtttgagtgt gtgcgagtcg cgaatcagga atgaattgct agtaggcact ccgaccgggc 1080

gggggccgag ggacca 1096

<210> 96
<211> 1075
<212> ADN
<213> Prototheca moriformis

<400> 96

ggccgacagg acgcgcgtca aaggtgctgg gcgtgtatgc cctggtcggc aggtcgttgc 60

tgttgctgcg ctcgtggttc cgcaaccctg attttggcgt cttattctgg cgtggcaagc 120

gctgacgccc gcgagccggg ccggcggcga tgcggtgtct cacggctgcc gagctccaag 180

ggaggcaaga gcgcccggat cagctgaagg gctttacacg caaggtacag ccgctcctgc 240

aaggctgcgt ggtggacttg aacctgtagg tcctctgctg aagttcctcc actacctcac 300

caggcccagc agaccaaagc acaggctttt caggtccgtg tcatccactc taaaacactc 360

gactacgacc tactgatggc cctagattct tcatcaacaa tgcctgagac acttgctcag 420

aattgaaact ccctgaaggg accaccagag gccctgagtt gttccttccc cccgtggcga 480

gctgccagcc aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca 540

ggtcatggga ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc 600

taatcagcta tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt 660

gatccttcgt gtacgggccc ttccctcaac cctaggtatg cgcgcatgcg gtcgccgcgc 720

aactcgcgcg agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat 780

gcgcggcgcc tttcttgcga taatttatgc aatggactgc tctgcaaatt tctgggtctg 840

tcgccaaccc taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta 900

ccgactaccc taatatcagc ccggctgcct gacgccagcg tccacttttg cgtacacatt 960

ccattcgtgc ccaagacatt tcattgtggt gcgaagcgtc cccagttacg ctcacctgtt 1020

tcccgacctc cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagcc 1075

<210> 97
<211> 772
<212> ADN
<213> Prototheca moriformis

<400> 97

tcaccagcgg acaaagcacc ggtgtatcag gtccgtgtca tccactctaa agagctcgac 60

tacgacctac tgatggccct agattcttca tcaaaaacgc ctgagacact tgcccaggat 120

tgaaactccc tgaagggacc accaggggcc ctgagttgtt ccttcccccc gtggcgagct 180

gccagccagg ctgtacctgt gatcggggct ggcgggaaaa caggcttcgt gtgctcaggt 240

tatgggaggt gcaggacagc tcattaaacg ccaacaatcg cacaattcat ggcaagctaa 300

tcagttattt cccattaacg agctataatt gtcccaaaat tctggtctac cgggggtgat 360

ccttcgtgta cgggcccttc cctcaaccct aggtatgcgc acatgcggtc gccgcgcaac 420

gcgcgcgagg gccgagggtt tgggacgggc cgtcccgaaa tgcagttgca cccggatgcg 480

tggcaccttt tttgcgataa tttatgcaat ggactgctct gcaaaattct ggctctgtcg 540

ccaaccctag gatcagcggt gtaggatttc gtaatcattc gtcctgatgg ggagctaccg 600

actgccctag tatcagcccg actgcctgac gccagcgtcc acttttgtgc acacattcca 660

ttcgtgccca agacatttca ttgtggtgcg aagcgtcccc agttacgctc acctgatccc 720

caacctcctt attgttctgt cgacagagtg ggcccagagg ccggtcgcag cc 772

<210> 98
<211> 991
<212> ADN
<213> Protoheca moriformis

<400> 98

cgaaggggtc tgcatcgatt cgcgcggtct ggaggccagc gtgactgctc gcgaaaatgc 60

tctgccgtgt cgggctctgg ctggggcggc cagagatctc accgtgccac acgcaactgc 120

cgcactctgt gcccgccacc tggcgcgcac atgcgacctc ttccccgtca taccctctcc 180

tcatgtgatc tttccacacg agtgacgcag gtgcgcggag tggagggaat caggacgttt 240

tcaaggtacc tgctcgagcc gtaccaacag ctgccgcccg gcaaggaaga gatcgaggca 300

gagattgccc ggctggaggc ccggataacg gagctcaaga gcaagctgtc cgagtgagac 360

cgcccaggtg cacgtgtcga ctcgctatga catgtactcg acacaacatg aggaattcat 420

cgaatttgta ggaagcgggc attggtacgg gagtgggaaa gcgaaaaaac ctccctccgg 480

cagtgccatc tgccggagtc gaacgttgat agggttctcg tgacagggtg tgacctctca 540

gccttgcatc aattaaacgc tatagacatt atcagtaacc gtgaatcccg cattggatgc 600

cacccgcgcg accattgggg acctgcatta cagatctagg tgagatgaca gcgaggcaac 660

ttcggcccgc ggcccagctt gcggcgcacc aatattggtc acgggaagcc acacaccgac 720

cataaatgaa tacttgtaag ctatgtcaac cgatcaatgg cgtcgaaagt gtgccacgag 780

gatccatctg gcggggcggc gtggcgcaca agcgcagtcg caatttctcg gacccatctg 840

acctaggccc agcgccgcgg gagaaatccc cggcgggtcc tccacgcagt aaccctaatg 900

agtatcgagc gccgaccatt tacaccatcg cccccgaaat ccttccgaca ttattattat 960

cttttagatc ttggaacaga ctctgccaac c 991

<210> 99
<211> 1347
<212> ADN
<213> Prototheca moriformis

<400> 99

agagagcgga ggtggggttg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg 60

cgactcgcac acggtccagt tacccccccc tccgcccaaa cgcaagcctc ccatcttgat 120

gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata 180

ataacaatgc cctgtgtcaa gtgcattcct aaaaaaattc tgtcccaacc aacaatccca 240

cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg 300

cgcaagcgac ccccatgcgc gaccaggctc gaaagtgatt tatgacttga gacgagcgag 360

tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt 420

gcacggcaac gcacggcttc gccaacccga ccagccagga cctcgactac tctaccgcga 480

attcgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact 540

tgatgccaat tttgacattc gggttgctcg ttggctgcgc ccacatcggc cgtgagtgca 600

gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta 660

catcaacacg acgtgttgcg taatcccgcc cggctgcgca tcgtgccaac ccattcgcga 720

tggatggtcg gaaaatggtg tgccaactgc cctgagggag gctctcgcga aacgggcacg 780

tccctgaaac cgaaactgtg gccttgtcgt cggccacgca agcacgtgga ccctaaacac 840

caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc 900

atcgcacact gccattataa tgcatctagc tcggcgacaa gtttagaaaa ggcaggctgc 960

attgttccat ttcgccgtgg cggcgtgggt gcccatttta cgaggtttgg gctcccgggc 1020

agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080

gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140

acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200

ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260

cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat 1320

attttccagt cgtcagcgaa atcaagt 1347

<210> 100
<211> 1180
<212> ADN
<213> Prototheca moriformis

<400> 100

gatggtgggg tgtctgcctt gggctgggtg atggaggctg gtggtgcgcg ggtttcctga 60

tgcattctat ctacgcagtg tcatggtgtc cattccacac accagtacac ccttacacta 120

aggatccatc cctccttccc tcttcaggac tacatggacc ccacgagcta ccgaccgggc 180

tttctcaaaa acgtcaaggt catgtttgac atgcgggacg tggtggacga cgtgcaaggt 240

gcgtccggag tgcgcgcaaa tgagcaagtc gggcaatgtg tcggggtggg caccggggct 300

ggagatccgc gatccccgag aaaacgccgt accacccccc gcgctattcc ctcgattgcg 360

cgcagatgtg gtgaccgaca cgggggacaa cctggcggac atggggcgcc ggacctggaa 420

gcacgccaag tcgcacacgg ggaggctcgt gcagtccccc ccatcgtacc tcaagggtct 480

ctttggtcgc gatccaaagt acgctggtgg catggcatgc ccgaaatgaa catcatgtgt 540

gatctccgat tgccaatggc cacctccacg gaccaccttg caggcggaag cgcaatccag 600

ggcccgagcc tgacgaggac ggagactcct cgtccagcgc ggggtccccg acccgacgca 660

gcagccgacc cctgctaacc cggcaacgat cggaccagca accttgctgt agttccgatc 720

cgtgatgacg ggcattgccg ccgctcgatc cgctttgatg actgtctatt atttgcgcgg 780

agccccctcg gaaccctacc ccgctcttgc aagccccttg catcggagat cctcgtgcgc 840

ccgccatgac cccactggat tgcccaacat ccttctttat cgtgtaaaat gtgattcctc 900

ggctgcaatc gactggcctt cgcttctggc cccaagaggg ctcgaacgtg cggcagcgag 960

ggcgctgaca cacccaagcc ctagggcttt caacgtcggc tgccaggccg gataggggga 1020

tcgcctcctt tccaccaccc acctacgagg gattcgagtc ggcttccagc tcagctattc 1080

ggccgcgccc ccggccctgc agacgtcctc cagtttccga acaggtcgct ctcagaacac 1140

ctgccgcggc tgcgatacgg caggctctca aagcgtcgac 1180

<210> 101
<211> 1263
<212> ADN
<213> Prototheca moriformis

<400> 101

cgcgtggagc ggtgcgtgcg gatgccgcgc gcctgccaag gccttttgta tgcctggcct 60

gggaagtttc ctgactgaag catcttcaag atgctctctc acgaccagcg acaccaacac 120

cgtcactttt tgcccctcct gccgcaggtg ccactttcta ctttgacgtc ttctccaggc 180

ggtacattgc gggactgagc gccaattcgg ccaagaacag cgctgtcgac ttgaggaggc 240

aggggtccgt cgactctgcc gagtgacacg ccttcgaccc gactgtacta cggcctgctg 300

aagagtgggt ctcgccggcc ggcgtgaccg gccctgtgcc cacaatcgac catctattcg 360

ctccttgtca tctggcgccg tcaattgccc gcgacttgac ggcaactggc tcgatcgagt 420

cgtattgaaa aagcacgttt tgtcctacag ggccgcggtc cgttaccaac gtggttctcg 480

ttaggttttc gtcgggcggt ggtgcgcgaa ctgtccgatg ccatcccggc aaaccccagc 540

aaggtcgcca gtctggttct gacgcaatag agtgcgtttt gggccagtct aaaaattcgt 600

ctggcatgac gtggctccac atcgtacccg gagcctgcct tggtaatgtg aggcaccggt 660

gccaactcca ttatggcagg catcgagcgc gcaggtgagt acatgacctt ccgtgaattg 720

ggaaggcgag cttgtgtaac gcctgcgatc gtgccagtga ggcatcgtaa actcaaaata 780

ttttgtagaa agtgtctgat gcctggtgag gctgcgtagg gcaagggcaa gcccttggca 840

gatgggtaat gggtccggac ctcacaacag caaccccgcg tccccttag ggcccctgag 900

gctcgatggc agggccagcg agcccgcggc caaagggcgc catcccacgg tcgcccaacg 960

actccacggg tcctatacct catcttgaat ggcactaaaa actatagaat atcgggcact 1020

ggtgggcgtc tggggtacag ctggccgagc gcagtggcaa accctaggtc ccgcctcaag 1080

ggcgattccc gggtcaatga cacgcaagca agatcacatg gcgcggtccg cctcgcggct 1140

ccacacccag gccctagttt cgcaacccat aaatatcgcc ccgataccat cataagccag 1200

caaataattt tttatcagag ttccaaacct cctcagctgt gggaaaccag cccactctga 1260

acg 1263

<210> 102
<211> 1400
<212> ADN
<213> Prototheca moriformis

<400> 102

ccgagcagtt catggccaag tacaaggact agagaccgga ggtcggtagg ctgaatggag 60

ctggcgtcgt cgtgcgcgac gtgcacgcga tgcgatacta cgaccccaca aacgcatgcc 120

tcccatcttg atgcctttcc ggccatttat actatttctc atttcgctgt aacatcttga 180

ataatagaat tgccctgtgt caagtggatt ccaagaaata ttctgtccca acaaaacaac 240

ccaacctgaa aacaacctca aataccacca gccctgccca cctgcccagt acacttttcc 300

aataccatct ccctaccttc acgcgcaagc ggcacccatg cgcgaccagg ctcgaaagga 360

tttcacgact caggacgagc gagtggcggc gcgaccgcct gcctgttcgt cacgtgccgt 420

acgtcggcga ccgctagagc tttgcctggc aacccccggc ttcgtcaacc cggccagcca 480

ggatctcgac cactctaccg cgaaatcgcc tcaagaagtc gccaaaagtg ccgtacacca 540

tgcttcgcag cgctgttcaa acttgatgcc aatcttgaca atcaggttgc tcgttggctg 600

cgtccacatc ggccgtgatt gcagcaggcg gggatcggac acggaggacg cggcgtcacg 660

ccgcgaacgc agcccgtaac tctacatcaa cgcgatatgt tgcgtaatcc cgcccggctg 720

cgcattgtga caacccattc gcgatggatg gtcggaaaat ggtgtgccaa ctgccctgag 780

ggactctctc gcgaaacggg cacgtccctg tatccgaaac tgtggcatgg ccttgtcgac 840

cacgcaagca cgtggaccct aacaccacga aaataagtaa aaaaggttga catcctctac 900

gagcgaattg tttgctcgac ccttcatcgc acactgtcat tataatgcat ctagctcggc 960

gacaagttta aaaaaggcag gctgcattat tccattttgc cgtggcggca tgggtgccca 1020

ttttatgagg tttgggctct tgggcagcga ccgagccagg ttgagtccct ctcgcccgtc 1080

gacaacgttc caaagcccat aagtggctaa taaacaactt gatggtacct gtacactgcc 1140

agttccttct tccccggccg aggtttacac gtgatggcca tggcttcgcg tttcaggctg 1200

acttcccatt ccgactttcc agagggtccg cggacgccgg gggttggctg cgtgaggccc 1260

acccccttgtt ccccgcgtcc cgacaaacac aattgcgtta cataaggggg aagccgcccc 1320

ccgttcagag tgcaaacatc tttcattata tttttcagtc gtcagcgaaa tcaagtatgt 1380

cgctgacagg catgaaggcc 1400

<210> 103
<211> 3681
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 103

gccctttgtc atcgttggca tgctttttgc gtatgtacca tatgttgaat gtataatacg 60

aacggttgac cgtctgagat gcgagctttg ggtcttgtca aatgcgtggc cgcacggctc 120

cctcgcaccc agccccgagg cgtcgcgcac ctggcgagga gcagacccac gccaagaaag 180

tctagtccag catgtaacaa catcaggcaa tgtgacgttt tcggttcccg atttctctgc 240

cgctctttga cggcaggcac gggcgagcaa ccggcggcgc tcgcgtcagg cacgatggat 300

gcggcgctgc ccacctgtca atgtacccca ccagtctgtc gatcgctaca agcaaccttg 360

tgctccacat tcccacttgc agacagtcta gtcgattttg ccaagctgga tgtgaggatt 420

ggccatatct tggaggccaa gattcacccg gatgctgatg ggtacgtacg cgagccaggc 480

aggcagctgc gttgactttc tgattggcac aaagctttgg ctactctcaa taccaaccac 540

gtgccccttc tgcacacctg cttccttctg atgaccactc gccacgcatg tcgcagtctg 600

tacgtcgagc agatcgacct cggcgaggag gggggccctc gcaccatcgt gagtggcctg 660

gtccggcacg tgaccctgga ggaccttgtc ggccggcggg tggtggtgct ggccaacctc 720

aagcctcgga gcatgcgcgg ggtcaaatcg gctgggatgc tgctctgcgc cgccaacgcg 780

gatcacaccg cggtggagcc gctgcgggtc ccggacgccg ccgtgacggg ggagcgggtc 840

tgggcggggg acgaggcact cctgtccacg gagcctgcca cacccaatca ggtaaggaca 900

cgttattggt gcgcatggtg cgaatgcgtg gtctgacctg ctgtgggtat gtgttgtggg 960

attggaaacc gaatgagggc cgttcaggat tgagcccttg gccccaccct gctcatcctc 1020

tcacgcccgc aggtccagaa gaagaaaatc tgggaggcag tacagccgct gctgagagtg 1080

aacgcccagg ggatcgctac tgtggcagga gaggctatgg tgaccagtgc ggggccactg 1140

accgcgccca cgctggttga cgccgcgatt tcctgacgcg agcgactgat tcttgacctt 1200

tgagaagcca ccacagcacc attttcattg ttcatccttg atttcagtac gacttctcac 1260

catttcagta ctgtaggacc cccaaaatag tgtgatcacg ctcgcaaggc acctgtgtga 1320

tcacggggaa gggcgaattc ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg 1380

gctgcgagac ggcttcccgg cgctgcatgc aacaccgatg atgcttcgac cccccgaagc 1440

tccttcgggg ctgcatgggc gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc 1500

aggcccccga ttgcaaagac attatagcga gctaccaaag ccatattcaa acacctagat 1560

cactaccact tctacacagg ccactcgagc ttgtgatcgc actccgctaa gggggcgcct 1620

cttcctcttc gtttcagtca caacccgcaa acggcgcgcc atgctgctgc aggccttcct 1680

gttcctgctg gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga 1740

ccgccccctg gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg 1800

gtacgacgag aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt 1860

ctgggggacg cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga 1920

ccagcccatc gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt 1980

ggtggactac aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg 2040

cgtggccatc tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct 2100

ggacggcggc tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac 2160

ccagttccgc gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc 2220

ggccaagtcc caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa 2280

gctggagtcc gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct 2340

gatcgaggtc cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat 2400

caaccccggc gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg 2460

cacccacttc gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta 2520

cgccctgcag accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg 2580

ggcctccaac tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc 2640

cctcgtgcgc aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat 2700

caacctgaag gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc 2760

caccaacacc acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg 2820

caccctggag ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt 2880

gttcgcggac ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat 2940

gggcttcgag gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt 3000

cgtgaaggag aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag 3060

cgagaacgac ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct 3120

gtacttcaac gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc 3180

cctgggctcc gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca 3240

ggtgcgcgag gtcaagtgat taattaactc gaggcagcag cagctcggat agtatcgaca 3300

cactctggac gctggtcgtg tgatggactg ttgccgccac acttgctgcc ttgacctgtg 3360

aatatccctg ccgcttttat caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt 3420

ttgcgagttg ctagctgctt gtgctatttg cgaataccac ccccagcatc cccttccctc 3480

gtttcatatc gcttgcatcc caaccgcaac ttatctacgc tgtcctgcta tccctcagcg 3540

ctgctcctgc tcctgctcac tgcccctcgc acagccttgg tttgggctcc gcctgtattc 3600

tcctggtact gcaacctgta aaccagcact gcaatgctga tgcacgggaa gtagtgggat 3660

gggaacacaa atggaaagct t 3681

<210> 104

<211> 3850

<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 104

tttggccccg ctttccagct ccggatctgc tggcgtccgc cgcgagacgt gacatcgcac 60

gtcgccggga gcgccagctt gatcacttgg caggggggccg tgctctacaa ataccaggcc 120

ccgcggcggt cagttcgcac atccaatacc tgccgagcca tcttgcctac acttttatc 180

gactcctcta ctctgttcgc gagagcgctc ggtccaggct tggaattcgc cgaattcagc 240

tcgatcagtc gcttctgcaa ctgatctcgg ccgttcgcag actgcctttt ctcagcttgt 300

caggtagcga gttgttgttt tatatttatt cgatttcatc tgtgttgcat gtcttgttcg 360

tgctgtgcgt tctttctggg ccgcgctgtc gggtcgcatg ggctagctgt actcatgtta 420

gtcatgccgg tccgaccttg ttcgaggaag gccccacact gagcgtgccc tctttctaca 480

ccccttgtgc agaaattaga tagaaagcag aattcctttc ttgcgctatg acacttccag 540

caaaaggtag ggcgggctgc gagacggctt cccggcgctg catgcaacac cgatgatgct 600

tcgacccccc gaagctcctt cggggctgca tgggcgctcc gatgccgctc cagggcgagc 660

gctgtttaaa tagccaggcc cccgattgca aagacattat agcgagctac caaagccata 720

ttcaaacacc tagatcacta ccacttctac acaggccact cgagcttgtg atcgcactcc 780

gctaaggggg cgcctcttcc tcttcgtttc agtcacaacc cgcaaacggc gcgccatgct 840

gctgcaggcc ttcctgttcc tgctggccgg cttcgccgcc aagatcagcg cctccatgac 900

gaacgagacg tccgaccgcc ccctggtgca cttcaccccc aacaagggct ggatgaacga 960

ccccaacggc ctgtggtacg acgagaagga cgccaagtgg cacctgtact tccagtacaa 1020

cccgaacgac accgtctggg ggacgccctt gttctggggc cacgccacgt ccgacgacct 1080

gaccaactgg gaggaccagc ccatcgccat cgccccgaag cgcaacgact ccggcgcctt 1140

ctccggctcc atggtggtgg actacaacaa cacctccggc ttcttcaacg acaccatcga 1200

cccgcgccag cgctgcgtgg ccatctggac ctacaacacc ccggagtccg aggagcagta 1260

catctcctac agcctggacg gcggctacac cttcaccgag taccagaaga accccgtgct 1320

ggccgccaac tccacccagt tccgcgaccc gaaggtcttc tggtacgagc cctcccagaa 1380

gtggatcatg accgcggcca agtcccagga ctacaagatc gagatctact cctccgacga 1440

cctgaagtcc tggaagctgg agtccgcgtt cgccaacgag ggcttcctcg gctaccagta 1500

cgagtgcccc ggcctgatcg aggtccccac cgagcaggac cccagcaagt cctactgggt 1560

gatgttcatc tccatcaacc ccggcgcccc ggccggcggc tccttcaacc agtacttcgt 1620

cggcagcttc aacggcaccc acttcgaggc cttcgacaac cagtcccgcg tggtggactt 1680

cggcaaggac tactacgccc tgcagacctt cttcaacacc gacccgacct acgggagcgc 1740

cctgggcatc gcgtgggcct ccaactggga gtactccgcc ttcgtgccca ccaacccctg 1800

gcgctcctcc atgtccctcg tgcgcaagtt ctccctcaac accgagtacc aggccaaccc 1860

ggagacggag ctgatcaacc tgaaggccga gccgatcctg aacatcagca acgccggccc 1920

ctggagccgg ttcgccacca acaccacgtt gacgaaggcc aacagctaca acgtcgacct 1980

gtccaacagc accggcaccc tggagttcga gctggtgtac gccgtcaaca ccacccagac 2040

gatctccaag tccgtgttcg cggacctctc cctctggttc aagggcctgg aggaccccga 2100

ggagtacctc cgcatgggct tcgaggtgtc cgcgtcctcc ttcttcctgg accgcgggaa 2160

cagcaaggtg aagttcgtga aggagaaccc ctacttcacc aaccgcatga gcgtgaacaa 2220

ccagcccttc aagagcgaga acgacctgtc ctactacaag gtgtacggct tgctggacca 2280

gaacatcctg gagctgtact tcaacgacgg cgacgtcgtg tccaccaaca cctacttcat 2340

gaccaccggg aacgccctgg gctccgtgaa catgacgacg ggggtggaca acctgttcta 2400

catcgacaag ttccaggtgc gcgaggtcaa gtgattaatt aactcgaggc agcagcagct 2460

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 2520

ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 2580

tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accaccccca 2640

gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 2700

tgctatccct cagcgctgct cctgctcctg ctcactgccc ctcgcacagc cttggtttgg 2760

gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac 2820

gggaagtagt gggatgggaa cacaaatgga ccgacacgcc cccggcccag gtccagttct 2880

cctgggtctt ccagaggccc gtcgccatgt aaagtggcag agattggcgc ctgattcgat 2940

ttggatccaa ggatctccaa tcggtgatgg ggactgagtg cccaactacc acccttgcac 3000

tatcgtcctc gcactattta ttcccacctt ctgctcgccc tgccgggcga ttgcgggcgt 3060

ttctgccctt gacgtatcaa tttcgccccc gctggcgcga ggattcttca ttctaataag 3120

aactcactcc cgccagctct gtacttttcc tgcggggccc ctgcatggct tgttcccaat 3180

gcttgctcga tcgacggcgc ccattgccca cggcgctgcc gcatccatgt gaagaaacac 3240

ggaagagtgc gaagactgga agtgaattaa gagtataaga agaggtacca agggattctc 3300

aggtgctctt aggaacggct tttccttcgc caagagaaac tgctactgct cgtgtcgcca 3360

cggtggtcaa gccgccccat ctgcgatcca ccaggcccat ccgcggactc gcgatcagcc 3420

tgctggatcc ggactgccga cctgaccgct cgcatccacc attacaaccc tccaattgga 3480

caccactccc acgtcctaaa gttcaccatg caagctgatc gatcgcattc gccgatgcac 3540

tcgcctgcca cagaggtgtg cgcttcggac tagcgtgcag gcgccccgag gccaccagca 3600

tgcaccgatg gaagcgggca cggccgctgc tccaggtcgc tggctcgctc agacccatag 3660

caacctccgc tgcgtcccta aatgtcacac agagcgtctt tgatgggtac ggatgggaga 3720

gaatctgatt gggcattgct ggtgcagtgc aggaagatgg caagtgcaca gtcagtcatg 3780

ctgtacaaac tggtgcctcg tagtattgac tcgtatagtg catagtatca tgcatggtcg 3840

ttacttgcaa 3850

<210> 105
<211> 3108
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 105

tttggccccg ctttccagct ccggatctgc tggcgtccgc cgcgagacgt gacatcgcac 60

gtcgccggga gcgccagctt gatcacttgg caggggggccg tgctctacaa ataccaggcc 120

ccgcggcggt cagttcgcac atccaatacc tgccgagcca tcttgcctac acttttatc 180

gactcctcta ctctgttcgc gagagcgctc ggtccaggct tggaattcgc cgaattcagc 240

tcgatcagtc gcttctgcaa ctgatctcgg ccgttcgcag actgcctttt ctcagcttgt 300

caggtagcga gttgttgttt tatatttatt cgatttcatc tgtgttgcat gtcttgttcg 360

tgctgtgcgt tctttctggg ccgcgctgtc gggtcgcatg ggctagctgt actcatgtta 420

gtcatgccgg tccgaccttg ttcgaggaag gccccacact gagcgtgccc tctttctaca 480

ccccttgtgc agaaattaga tagaaagcaa tgctgctgca ggccttcctg ttcctgctgg 540

ccggcttcgc cgccaagatc agcgcctcca tgacgaacga gacgtccgac cgccccctgg 600

tgcacttcac ccccaacaag ggctggatga acgaccccaa cggcctgtgg tacgacgaga 660

aggacgccaa gtggcacctg tacttccagt acaacccgaa cgacaccgtc tgggggacgc 720

ccttgttctg gggccacgcc acgtccgacg acctgaccaa ctgggaggac cagcccatcg 780

ccatcgcccc gaagcgcaac gactccggcg ccttctccgg ctccatggtg gtggactaca 840

acaacacctc cggcttcttc aacgacacca tcgacccgcg ccagcgctgc gtggccatct 900

ggacctacaa caccccggag tccgaggagc agtacatctc ctacagcctg gacggcggct 960

acaccttcac cgagtaccag aagaaccccg tgctggccgc caactccacc cagttccgcg 1020

acccgaaggt cttctggtac gagccctccc agaagtggat catgaccgcg gccaagtccc 1080

aggactacaa gatcgagatc tactcctccg acgacctgaa gtcctggaag ctggagtccg 1140

cgttcgccaa cgagggcttc ctcggctacc agtacgagtg ccccggcctg atcgaggtcc 1200

ccaccgagca ggaccccagc aagtcctact gggtgatgtt catctccatc aaccccggcg 1260

ccccggccgg cggctccttc aaccagtact tcgtcggcag cttcaacggc acccacttcg 1320

aggccttcga caaccagtcc cgcgtggtgg acttcggcaa ggactactac gccctgcaga 1380

ccttcttcaa caccgacccg acctacggga gcgccctggg catcgcgtgg gcctccaact 1440

gggagtactc cgccttcgtg cccaccaacc cctggcgctc ctccatgtcc ctcgtgcgca 1500

agttctccct caacaccgag taccaggcca acccggagac ggagctgatc aacctgaagg 1560

ccgagccgat cctgaacatc agcaacgccg gcccctggag ccggttcgcc accaacacca 1620

cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa cagcaccggc accctggagt 1680

tcgagctggt gtacgccgtc aacaccaccc agacgatctc caagtccgtg ttcgcggacc 1740

tctccctctg gttcaagggc ctggaggacc ccgaggagta cctccgcatg ggcttcgagg 1800

tgtccgcgtc ctccttcttc ctggaccgcg ggaacagcaa ggtgaagttc gtgaaggaga 1860

acccctactt caccaaccgc atgagcgtga acaaccagcc cttcaagagc gagaacgacc 1920

tgtcctacta caaggtgtac ggcttgctgg accagaacat cctggagctg tacttcaacg 1980

acggcgacgt cgtgtccacc aacacctact tcatgaccac cgggaacgcc ctgggctccg 2040

tgaacatgac gacgggggtg gacaacctgt tctacatcga caagttccag gtgcgcgagg 2100

tcaagtgacc gacacgcccc cggcccaggt ccagttctcc tgggtcttcc agaggcccgt 2160

cgccatgtaa agtggcagag attggcgcct gattcgattt ggatccaagg atctccaatc 2220

ggtgatgggg actgagtgcc caactaccac ccttgcacta tcgtcctcgc actatttatt 2280

cccaccttct gctcgccctg ccgggcgatt gcgggcgttt ctgcccttga cgtatcaatt 2340

tcgcccctgc tggcgcgagg attcttcatt ctaataagaa ctcactcccg ccagctctgt 2400

acttttcctg cggggccccc gcatggcttg ttcccaatgc ttgctcgatc gacggcgccc 2460

attgcccacg gcgctgccgc atccatgtga agaaacacgg aagagtgcga agactggaag 2520

tgaattaaga gtataagaag aggtaccaag ggattctcag gtgctcttag gaacggcttt 2580

tccttcgcca agagaaactg ctactgctcg tgtcgccacg gtggtcaagc cgccccatct 2640

gcgatccacc aggcccatcc gcggactcgc gatcagcctg ctggatccgg actgccgacc 2700

tgaccgctcg catccaccat tacaaccctc caattggaca ccactcccac gtcctaaagt 2760

tcaccatgca agctgatcga tcgcattcgc cgatgcactc gcctgccaca gaggtgtgcg 2820

cttcggacta gcgtgcaggc gccccgaggc caccagcatg caccgatgga agcgggcacg 2880

gccgctgctc caggtcgctg gctcgctcag acccatagca acctccgctg cgtccctaaa 2940

tgtcacacag agcgtctttg atgggtacgg atgggagaga atctgattgg gcattgctgg 3000

tgcagtgcag gaagatggca agtgcacagt cagtcatgct gtacaaactg gtgcctcgta 3060

gtattgactc gtatagtgca tagtatcatg catggtcgtt acttgcaa 3108

<210> 106
<211> 559
<212> ADN
<213> Chlorella luteoviridis

<400> 106

tgttgaagaa tgagccggcg acttatagga agtggcgtgg ttaaggaatt ttccgaagcc 60

caagcgaaag caagttttaa aaatagcgat atttgtcact ttttatggac ccgaacccgg 120

gtgatctaac cgtgaccagg atgaagcttg ggtaacacca agtgaaggtc cgaactcttc 180

gatctttaaa aatcgtgaga tgagttgcgg ttagtaggtg aaatgccaat cgaactcgga 240

gctagctggt tctccccgaa atgtgttgag gcgcagcgat gaatgacaaa acaaatagta 300

cggtgtaggg gtaaagcact gtttcggtgc gggctgcgaa agcggtacca aatcgtggca 360

aactcagaat actacgcttg tataccattc atcagtgaga ctgtggggga taagctccat 420

agtcaagagg gaaacagccc agatcaccag ttaaggcccc aaaatgacag ctaagtggca 480

aaggaggtga aagtgcagaa acaaccagga ggtttgccca gaagcagcca tcctttaaag 540

agtgcgtaat agctcactg 559

<210> 107
<211> 1841
<212> ADN
<213> Cuphea palustris

<400> 107

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 360

cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 420

cccgtgcgcg ggcgcgccag catgctgctg tcggcggtga ccacggtctt cggcgtggcc 480

gagaagcagt ggcccatgct ggaccgcaag tccaagcgcc ccgacatgct ggtcgagccc 540

ctgggcgtgg accgcatcgt ctacgacggc gtgagcttcc gccagtcgtt ctccatccgc 600

agctacgaga tcggcgccga ccgcaccgcc tcgatcgaga cgctgatgaa catgttccag 660

gagacctccc tgaaccactg caagatcatc ggcctgctga acgacggctt cggccgcacg 720

cccgagatgt gcaagcgcga cctgatctgg gtcgtgacca agatgcagat cgaggtgaac 780

cgctacccca cgtggggcga caccatcgag gtcaacacgt gggtgagcgc ctcgggcaag 840

cacggcatgg gccgcgactg gctgatctcc gactgccaca ccggcgagat cctgatccgc 900

gcgacgagcg tctgggcgat gatgaaccag aagacccgcc gcctgtcgaa gatcccctac 960

gaggtgcgcc aggagatcga gccccagttc gtcgactccg cccccgtgat cgtggacgac 1020

cgcaagttcc acaagctgga cctgaagacg ggcgacagca tctgcaacgg cctgaccccc 1080

cgctggacgg acctggacgt gaaccagcac gtcaacaacg tgaagtacat cggctggatc 1140

ctgcagtcgg tccccaccga ggtgttcgag acgcaggagc tgtgcggcct gaccctggag 1200

taccgccgcg agtgcggccg cgactccgtg ctggagagcg tcacggccat ggacccctcg 1260

aaggagggcg accgctccct gtaccagcac ctgctgcgcc tggaggacgg cgcggacatc 1320

gtgaagggcc gcaccgagtg gcgccccaag aacgccggcg ccaagggcgc catcctgacg 1380

ggcaagacca gcaacggcaa ctcgatctcc tgactcgagt taattaactc gaggcagcag 1440

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 1500

acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 1560

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 1620

ccccagcatc cccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc 1680

tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1740

tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 1800

tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 1841

<210> 108
<211> 1010
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 108

ggcgcgccca gctgcccgac tggagcatgc tgctggccgc gatcaccacc ctgttcctgg 60

cggccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgac atgctggtgg 120

acccccttcgg cctgggccgc ttcgtgcagg acggcctggt gttccgcaac aacttcagca 180

tccgcagcta cgagatcggc gcggaccgca ccgccagcat cgagaccctg atgaaccacc 240

tgcaggagac cgccctgaac cacgtgaaga gcgtgggcct gctggaggac ggcctgggca 300

gcacccgcga gatgagcctg cgcaacctga tctgggtggt gaccaagatg caggtggcgg 360

tggaccgcta ccccacctgg ggcgacgagg tgcaggtgag cagctgggcg accgccatcg 420

gcaagaacgg catgcgccgc gagtggatcg tgaccgactt ccgcaccggc gagaccctgc 480

tgcgcgccac cagcgtgtgg gtgatgatga acaagctgac ccgccgcatc agcaagatcc 540

ccgaggaggt gtggcacgag atcggcccca gcttcatcga cgcgcccccc ctgcccaccg 600

tggaggacga cggccgcaag ctgacccgct tcgacgagag cagcgccgac ttcatccgca 660

agggcctgac cccccgctgg agcgacctgg acatcaacca gcacgtgaac aacgtgaagt 720

acatcggctg gctgctggag agcgcgcccc ccgagatcca cgagagccac gagatcgcca 780

gcctgaccct ggagtaccgc cgcgagtgcg gccgcgacag cgtgctgaac agcgccacca 840

aggtgagcga cagcagccag ctgggcaaga gcgccgtgga gtgcaaccac ctggtgcgcc 900

tgcagaacgg cggcgagatc gtgaagggcc gcaccgtgtg gcgccccaag cgccccctgt 960

acaacgacgg cgccgtggtg gacgtgcccg ccaagaccag ctgactcgag 1010

<210> 109
<211> 5472
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 109

ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 360

gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 420

gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 480

aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 540

cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 600

gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 660

aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 720

tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 780

tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 840

gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 900

caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 960

gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1020

cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1080

gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1140

gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1200

accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1260

tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 1320

aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 1380

gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 1440

acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 1500

ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 1560

ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 1620

gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 1680

aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 1740

ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 1800

gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 1860

gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 1920

gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 1980

tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2040

atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2100

ttgtgctatt tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat 2160

cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2220

actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 2280

taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 2340

tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 2400

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 2460

aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 2520

gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca 2580

gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct 2640

gccgcatgct tgtgctggtg aggctgggca gtgctgccat gctgattgag gcttggttca 2700

tcgggtggaa gcttatgtgt gtgctgggct tgcatgccgg gcaatgcgca tggtggcaag 2760

agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc 2820

agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt 2880

ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa 2940

acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga 3000

gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg 3060

agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca 3120

tggtgtgcgg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct 3180

cagccatgca ctggatggca gctgggctgc cactgcaatg tggtggatag gatgcaagtg 3240

gagcgaatac caaaccctct ggctgcttgc tgggttgcat ggcatcgcac catcagcagg 3300

agcgcatgcg aagggactgg ccccatgcac gccatgccaa accggagcgc accgagtgtc 3360

cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca tgtagcgctg 3420

acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag 3480

gcggcggcag cagagcggcg gcagcagcgg cgggggccac ccttcttgcg gggtcgcgcc 3540

ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcat gcctggagaa 3600

gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag 3660

cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg 3720

gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccg 3780

tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga 3840

agtaggcgcc ctcctttctg ctcgccctct ctccgtcccg ccactagtat ggccaccgca 3900

tccactttct cggcgttcaa tgcccgctgc ggcgacctgc gtcgctcggc gggctccggg 3960

ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg cccagctgcc cgactggagc 4020

cgcctgctga ccgccatcac caccgtgttc gtgaagtcca agcgccccga catgcacgac 4080

cgcaagtcca agcgccccga catgctggtg gacagcttcg gcctggagtc caccgtgcag 4140

gacggcctgg tgttccgcca gtccttctcc atccgctcct acgagatcgg caccgaccgc 4200

accgccagca tcgagaccct gatgaaccac ctgcaggaga cctccctgaa ccactgcaag 4260

agcaccggca tcctgctgga cggcttcggc cgcaccctgg agatgtgcaa gcgcgacctg 4320

atctgggtgg tgatcaagat gcagatcaag gtgaaccgct accccgcctg gggcgacacc 4380

gtggagatca acacccgctt cagccgcctg ggcaagatcg gcatgggccg cgactggctg 4440

atctccgact gcaacaccgg cgagatcctg gtgcgcgcca ccagcgccta cgccatgatg 4500

aaccagaaga cccgccgcct gtccaagctg ccctacgagg tgcaccagga gatcgtgccc 4560

ctgttcgtgg acagccccgt gatcgaggac tccgacctga aggtgcacaa gttcaaggtg 4620

aagaccggcg acagcatcca gaagggcctg acccccggct ggaacgacct ggacgtgaac 4680

cagcacgtgt ccaacgtgaa gtacatcggc tggatcctgg agagcatgcc caccgaggtg 4740

ctggagaccc aggagctgtg ctccctggcc ctggagtacc gccgcgagtg cggccgcgac 4800

tccgtgctgg agagcgtgac cgccatggac cccagcaagg tgggcgtgcg ctcccagtac 4860

cagcacctgc tgcgcctgga ggacggcacc gccatcgtga acggcgccac cgagtggcgc 4920

cccaagaacg ccggcgccaa cggcgccatc tccaccggca agaccagcaa cggcaactcc 4980

gtgtccatgg actacaagga ccacgacggc gactacaagg accacgacat cgactacaag 5040

gacgacgacg acaagtgact cgaggcagca gcagctcgga tagtatcgac acactctgga 5100

cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct 5160

gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt 5220

gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat 5280

cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg 5340

ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctcctggtac 5400

tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca 5460

aatggaaagc tt 5472

<210> 110
<211> 5451
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 110

ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 360

gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 420

gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 480

aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 540

cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 600

gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 660

aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 720

tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 780

tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 840

gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 900

caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 960

gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1020

cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1080

gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1140

gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1200

accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1260

tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 1320

aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 1380

gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 1440

acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 1500

ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 1560

ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 1620

gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 1680

aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 1740

ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 1800

gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 1860

gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 1920

gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 1980

tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2040

atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2100

ttgtgctatt tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat 2160

cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2220

actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 2280

taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 2340

tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 2400

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 2460

aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 2520

gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca 2580

gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct 2640

gccgcatgct tgtgctggtg aggctgggca gtgctgccat gctgattgag gcttggttca 2700

tcgggtggaa gcttatgtgt gtgctgggct tgcatgccgg gcaatgcgca tggtggcaag 2760

agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc 2820

agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt 2880

ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa 2940

acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga 3000

gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg 3060

agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca 3120

tggtgtgcgg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct 3180

cagccatgca ctggatggca gctgggctgc cactgcaatg tggtggatag gatgcaagtg 3240

gagcgaatac caaaccctct ggctgcttgc tgggttgcat ggcatcgcac catcagcagg 3300

agcgcatgcg aagggactgg ccccatgcac gccatgccaa accggagcgc accgagtgtc 3360

cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca tgtagcgctg 3420

acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag 3480

gcggcggcag cagagcggcg gcagcagcgg cgggggccac ccttcttgcg gggtcgcgcc 3540

ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcat gcctggagaa 3600

gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag 3660

cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg 3720

gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccg 3780

tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga 3840

agtaggcgcc ctcctttctg ctcgccctct ctccgtcccg ccactagtat gacgttcggg 3900

gtcgccctcc cggccatggg ccgcggtgtc tcccttcccc ggcccagggt cgcggtgcgc 3960

gcccagtcgg cgagtcaggt tttggagagc gggcgcgccc ccgactggtc catgctgttc 4020

gccgtgatca ccaccatctt cagcgccgcc gagaagcagt ggaccaacct ggagtggaag 4080

cccaagccca agctgcccca gctgctggac gaccacttcg gcctgcacgg cctggtgttc 4140

cgccgcacct tcgccatccg ctcctacgag gtgggccccg accgcagcac ctccatcctg 4200

gccgtgatga accacatgca ggaggccacc ctgaaccacg ccaagagcgt gggcatcctg 4260

ggcgacggct tcggcaccac cctggagatg tccaagcgcg acctgatgtg ggtggtgcgc 4320

cgcacccacg tggccgtgga gcgctacccc acctggggcg acaccgtgga ggtggagtgc 4380

tggatcggcg ccagcggcaa caacggcatg cgccgcgact tcctggtgcg cgactgcaag 4440

accggcgaga tcctgacccg ctgcacctcc ctgagcgtgc tgatgaacac ccgcacccgc 4500

cgcctgagca ccatccccga cgaggtgcgc ggcgagatcg gccccgcctt catcgacaac 4560

gtggccgtga aggacgacga gatcaagaag ctgcagaagc tgaacgactc caccgccgac 4620

tacatccagg gcggcctgac cccccgctgg aacgacctgg acgtgaacca gcacgtgaac 4680

aacctgaagt acgtggcctg ggtgttcgag accgtgcccg acagcatctt cgagtcccac 4740

cacatcagct ccttcaccct ggagtaccgc cgcgagtgca cccgcgactc cgtgctgcgc 4800

agcctgacca ccgtgagcgg cggcagctcc gaggccggcc tggtgtgcga ccacctgctg 4860

cagctggagg gcggcagcga ggtgctgcgc gcccgcaccg agtggcgccc caagctgacc 4920

gactccttcc gcggcatcag cgtgatcccc gccgagcccc gcgtgatgga ctacaaggac 4980

cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtgactc 5040

gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg 5100

ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct 5160

cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg 5220

cgaataccac ccccagcatc cccttccctc gtttcatatc gcttgcatcc caaccgcaac 5280

ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc 5340

acagccttgg tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact 5400

gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 5451

<210> 111
<211> 5454
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 111

ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300

cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 360

gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 420

gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 480

aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 540

cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 600

gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 660

aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 720

tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 780

tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 840

gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 900

caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 960

gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1020

cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1080

gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1140

gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1200

accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1260

tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 1320

aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 1380

gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 1440

acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 1500

ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 1560

ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 1620

gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 1680

aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 1740

ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 1800

gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 1860

gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 1920

gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 1980

tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2040

atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2100

ttgtgctatt tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat 2160

cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2220

actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 2280

taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 2340

tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 2400

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 2460

aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 2520

gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca 2580

gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct 2640

gccgcatgct tgtgctggtg aggctgggca gtgctgccat gctgattgag gcttggttca 2700

tcgggtggaa gcttatgtgt gtgctgggct tgcatgccgg gcaatgcgca tggtggcaag 2760

agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc 2820

agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt 2880

ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa 2940

acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga 3000

gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg 3060

agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca 3120

tggtgtgcgg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct 3180

cagccatgca ctggatggca gctgggctgc cactgcaatg tggtggatag gatgcaagtg 3240

gagcgaatac caaaccctct ggctgcttgc tgggttgcat ggcatcgcac catcagcagg 3300

agcgcatgcg aagggactgg ccccatgcac gccatgccaa accggagcgc accgagtgtc 3360

cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca tgtagcgctg 3420

acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag 3480

gcggcggcag cagagcggcg gcagcagcgg cgggggccac ccttcttgcg gggtcgcgcc 3540

ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcat gcctggagaa 3600

gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag 3660

cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg 3720

gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccg 3780

tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga 3840

agtaggcgcc ctcctttctg ctcgccctct ctccgtcccg ccactagtat ggccaccgca 3900

tccactttct cggcgttcaa tgcccgctgc ggcgacctgc gtcgctcggc gggctccggg 3960

ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg cccccgactg gtccatgctg 4020

ttcgccgtga tcaccaccat cttctccgcc gccgagaagc agtggaccaa cctggagtgg 4080

aagcccaagc ccaacccccc ccagctgctg gacgaccact tcggccccca cggcctggtg 4140

ttccgccgca ccttcgccat ccgcagctac gaggtgggcc ccgaccgctc caccagcatc 4200

gtggccgtga tgaaccacct gcaggaggcc gccctgaacc acgccaagtc cgtgggcatc 4260

ctgggcgacg gcttcggcac caccctggag atgtccaagc gcgacctgat ctgggtggtg 4320

aagcgcaccc acgtggccgt ggagcgctac cccgcctggg gcgacaccgt ggaggtggag 4380

tgctgggtgg gcgcctccgg caacaacggc cgccgccacg acttcctggt gcgcgactgc 4440

aagaccggcg agatcctgac ccgctgcacc tccctgagcg tgatgatgaa cacccgcacc 4500

cgccgcctga gcaagatccc cgaggaggtg cgcggcgaga tcggccccgc cttcatcgac 4560

aacgtggccg tgaaggacga ggagatcaag aagccccaga agctgaacga ctccaccgcc 4620

gactacatcc agggcggcct gacccccgc tggaacgacc tggacatcaa ccagcacgtg 4680

aacaacatca agtacgtgga ctggatcctg gagaccgtgc ccgacagcat cttcgagagc 4740

caccacatct cctccttcac catcgagtac cgccgcgagt gcaccatgga cagcgtgctg 4800

cagtccctga ccaccgtgag cggcggctcc tccgaggccg gcctggtgtg cgagcacctg 4860

ctgcagctgg agggcggcag cgaggtgctg cgcgccaaga ccgagtggcg ccccaagctg 4920

accgactcct tccgcggcat cagcgtgatc cccgccgagt ccagcgtgat ggactacaag 4980

gaccacgacg gcgactacaa ggaccacgac atcgactaca aggacgacga cgacaagtga 5040

ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 5100

ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 5160

cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 5220

ttgcgaatac caccccagc atccccttcc ctcgtttcat atcgcttgca tcccaaccgc 5280

aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccccc 5340

cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc 5400

actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggaaa gctt 5454

<210> 112
<211> 2933
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 112

agagagcgga ggtggggttg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg 60

cgactcgcac acggtccagt tacccccccc tccgcccaaa cgcaagcctc ccatcttgat 120

gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata 180

ataacaatgc cctgtgtcaa gtgcattcct aaaaaaattc tgtcccaacc aacaatccca 240

cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg 300

cgcaagcgac ccccatgcgc gaccaggctc gaaagtgatt tatgacttga gacgagcgag 360

tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt 420

gcacggcaac gcacggcttc gccaacccga ccagccagga cctcgactac tctaccgcga 480

attcgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact 540

tgatgccaat tttgacattc gggttgctcg ttggctgcgc ccacatcggc cgtgagtgca 600

gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta 660

catcaacacg acgtgttgcg taatcccgcc cggctgcgca tcgtgccaac ccattcgcga 720

tggatggtcg gaaaatggtg tgccaactgc cctgagggag gctctcgcga aacgggcacg 780

tccctgaaac cgaaactgtg gccttgtcgt cggccacgca agcacgtgga ccctaaacac 840

caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc 900

atcgcacact gccattataa tgcatctagc tcggcgacaa gtttagaaaa ggcaggctgc 960

attgttccat ttcgccgtgg cggcgtgggt gcccatttta cgaggtttgg gctcccgggc 1020

agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080

gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140

acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200

ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260

cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat 1320

attttccagt cgtcagcgaa atcaagtact agtatggcca ccgcatccac tttctcggcg 1380

ttcaatgccc gctgcggcga cctgcgtcgc tcggcgggct ccgggccccg gcgcccagcg 1440

aggcccctcc ccgtgcgcgg gcgcgccccc gactggtcca tgctgttcgc cgtgatcacc 1500

accatcttct ccgccgccga gaagcagtgg accaacctgg agtggaagcc caagcccaac 1560

ccccccagc tgctggacga ccacttcggc ccccacggcc tggtgttccg ccgcaccttc 1620

gccatccgca gctacgaggt gggccccgac cgctccacca gcatcgtggc cgtgatgaac 1680

cacctgcagg aggccgccct gaaccacgcc aagtccgtgg gcatcctggg cgacggcttc 1740

ggcaccaccc tggagatgtc caagcgcgac ctgatctggg tggtgaagcg cacccacgtg 1800

gccgtggagc gctaccccgc ctggggcgac accgtggagg tggagtgctg ggtgggcgcc 1860

tccggcaaca acggccgccg ccacgacttc ctggtgcgcg actgcaagac cggcgagatc 1920

ctgacccgct gcacctccct gagcgtgatg atgaacaccc gcacccgccg cctgagcaag 1980

atccccgagg aggtgcgcgg cgagatcggc cccgccttca tcgacaacgt ggccgtgaag 2040

gacgaggaga tcaagaagcc ccagaagctg aacgactcca ccgccgacta catccagggc 2100

ggcctgaccc cccgctggaa cgacctggac atcaaccagc acgtgaacaa catcaagtac 2160

gtggactgga tcctggagac cgtgcccgac agcatcttcg agagccacca catctcctcc 2220

ttcaccatcg agtaccgccg cgagtgcacc atggacagcg tgctgcagtc cctgaccacc 2280

gtgagcggcg gctcctccga ggccggcctg gtgtgcgagc acctgctgca gctggagggc 2340

ggcagcgagg tgctgcgcgc caagaccgag tggcgcccca agctgaccga ctccttccgc 2400

ggcatcagcg tgatccccgc cgagtccagc gtgatggact acaaggacca cgacggcgac 2460

tacaaggacc acgacatcga ctacaaggac gacgacgaca agtgactcga gttaattaac 2520

tcgaggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg tgtgatggac 2580

tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc 2640

ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt 2700

tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat cccaaccgca 2760

acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc actgcccctc 2820

gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca 2880

ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaag ctt 2933

<210> 113
<211> 4817
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 113

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacgg cgcgccatgg tggccgccgc cgcctccagc gccttcttcc 3120

ccgtgcccgc ccccggcgcc tcccccaagc ccggcaagtt cggcaactgg ccctccagcc 3180

tgagcccctc cttcaagccc aagtccatcc ccaacggcgg cttccaggtg aaggccaacg 3240

acagcgccca ccccaaggcc aacggctccg ccgtgagcct gaagagcggc agcctgaaca 3300

cccaggagga cacctcctcc agccccccccc cccgcacctt cctgcaccag ctgcccgact 3360

ggagccgcct gctgaccgcc atcaccaccg tgttcgtgaa gtccaagcgc cccgacatgc 3420

acgaccgcaa gtccaagcgc cccgacatgc tggtggacag cttcggcctg gagtccaccg 3480

tgcaggacgg cctggtgttc cgccagtcct tctccatccg ctcctacgag atcggcaccg 3540

accgcaccgc cagcatcgag accctgatga accacctgca ggagacctcc ctgaaccact 3600

gcaagagcac cggcatcctg ctggacggct tcggccgcac cctggagatg tgcaagcgcg 3660

acctgatctg ggtggtgatc aagatgcaga tcaaggtgaa ccgctacccc gcctggggcg 3720

acaccgtgga gatcaacacc cgcttcagcc gcctgggcaa gatcggcatg ggccgcgact 3780

ggctgatctc cgactgcaac accggcgaga tcctggtgcg cgccaccagc gcctacgcca 3840

tgatgaacca gaagacccgc cgcctgtcca agctgcccta cgaggtgcac caggagatcg 3900

tgcccctgtt cgtggacagc cccgtgatcg aggactccga cctgaaggtg cacaagttca 3960

aggtgaagac cggcgacagc atccagaagg gcctgacccc cggctggaac gacctggacg 4020

tgaaccagca cgtgtccaac gtgaagtaca tcggctggat cctggagagc atgcccaccg 4080

aggtgctgga gacccaggag ctgtgctccc tggccctgga gtaccgccgc gagtgcggcc 4140

gcgactccgt gctggagagc gtgaccgcca tggaccccag caaggtgggc gtgcgctccc 4200

agtaccagca cctgctgcgc ctggaggacg gcaccgccat cgtgaacggc gccaccgagt 4260

ggcgccccaa gaacgccggc gccaacggcg ccatctccac cggcaagacc agcaacggca 4320

actccgtgtc catggactac aaggaccacg acggcgacta caaggaccac gacatcgact 4380

acaaggacga cgacgacaag tgactcgagg cagcagcagc tcggatagta tcgacacact 4440

ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata 4500

tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc 4560

gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatccccct tccctcgttt 4620

catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc 4680

tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct 4740

ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga 4800

acacaaatgg aaagctt 4817

<210> 114
<211> 4665

<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 114

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttcggggtcg ccctcccggc catgggccgc 3120

ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg 3180

gagagcgggc gcgcccagct gcccgactgg agccgcctgc tgaccgccat caccaccgtg 3240

ttcgtgaagt ccaagcgccc cgacatgcac gaccgcaagt ccaagcgccc cgacatgctg 3300

gtggacagct tcggcctgga gtccaccgtg caggacggcc tggtgttccg ccagtccttc 3360

tccatccgct cctacgagat cggcaccgac cgcaccgcca gcatcgagac cctgatgaac 3420

cacctgcagg agacctccct gaaccactgc aagagcaccg gcatcctgct ggacggcttc 3480

ggccgcaccc tggagatgtg caagcgcgac ctgatctggg tggtgatcaa gatgcagatc 3540

aaggtgaacc gctaccccgc ctggggcgac accgtggaga tcaacacccg cttcagccgc 3600

ctgggcaaga tcggcatggg ccgcgactgg ctgatctccg actgcaacac cggcgagatc 3660

ctggtgcgcg ccaccagcgc ctacgccatg atgaaccaga agacccgccg cctgtccaag 3720

ctgccctacg aggtgcacca ggagatcgtg cccctgttcg tggacagccc cgtgatcgag 3780

gactccgacc tgaaggtgca caagttcaag gtgaagaccg gcgacagcat ccagaagggc 3840

ctgacccccg gctggaacga cctggacgtg aaccagcacg tgtccaacgt gaagtacatc 3900

ggctggatcc tggagagcat gcccaccgag gtgctggaga cccaggagct gtgctccctg 3960

gccctggagt accgccgcga gtgcggccgc gactccgtgc tggagagcgt gaccgccatg 4020

gaccccagca aggtgggcgt gcgctcccag taccagcacc tgctgcgcct ggaggacggc 4080

accgccatcg tgaacggcgc caccgagtgg cgccccaaga acgccggcgc caacggcgcc 4140

atctccaccg gcaagaccag caacggcaac tccgtgtcca tggactacaa ggaccacgac 4200

ggcgactaca aggaccacga catcgactac aaggacgacg acgacaagtg actcgaggca 4260

gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 4320

ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt 4380

gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata 4440

ccacccccag catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct 4500

acgctgtcct gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc 4560

ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 4620

ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agctt 4665

<210> 115
<211> 4668
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 115

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180

gcgcttcgtg ggcgcgccca gctgcccgac tggagccgcc tgctgaccgc catcaccacc 3240

gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300

ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc 3360

ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420

aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480

ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540

atcaaggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc 3600

cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660

atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc 3720

aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgtgatc 3780

gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag 3840

ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900

atcggctgga tcctggagag catgcccacc gaggtgctgg agacccagga gctgtgctcc 3960

ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020

atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080

ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140

gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200

gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag 4260

gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 4320

ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380

tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga 4440

ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta 4500

tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca 4560

gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca 4620

atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

<210> 116
<211> 4668
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 116

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120

cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180

cccgtgcgcg ggcgcgccca gctgcccgac tggagccgcc tgctgaccgc catcaccacc 3240

gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300

ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc 3360

ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420

aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480

ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540

atcaaggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc 3600

cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660

atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc 3720

aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgtgatc 3780

gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag 3840

ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900

atcggctgga tcctggagag catgcccacc gaggtgctgg agacccagga gctgtgctcc 3960

ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020

atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080

ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140

gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200

gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag 4260

gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 4320

ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380

tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga 4440

ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta 4500

tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca 4560

gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca 4620

atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

<210> 117
<211> 4656
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 117

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcaccatgtc ggcgtgcccc 3120

gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc cacccgcctg 3180

agggggcgcg cccccgactg gagccgcctg ctgaccgcca tcaccaccgt gttcgtgaag 3240

tccaagcgcc ccgacatgca cgaccgcaag tccaagcgcc ccgacatgct ggtggacagc 3300

ttcggcctgg agtccaccgt gcaggacggc ctggtgttcc gccagtcctt ctccatccgc 3360

tcctacgaga tcggcaccga ccgcaccgcc agcatcgaga ccctgatgaa ccacctgcag 3420

gagacctccc tgaaccactg caagagcacc ggcatcctgc tggacggctt cggccgcacc 3480

ctggagatgt gcaagcgcga cctgatctgg gtggtgatca agatgcagat caaggtgaac 3540

cgctaccccg cctggggcga caccgtggag atcaacaccc gcttcagccg cctgggcaag 3600

atcggcatgg gccgcgactg gctgatctcc gactgcaaca ccggcgagat cctggtgcgc 3660

gccaccagcg cctacgccat gatgaaccag aagacccgcc gcctgtccaa gctgccctac 3720

gaggtgcacc aggagatcgt gcccctgttc gtggacagcc ccgtgatcga ggactccgac 3780

ctgaaggtgc acaagttcaa ggtgaagacc ggcgacagca tccagaaggg cctgaccccc 3840

ggctggaacg acctggacgt gaaccagcac gtgtccaacg tgaagtacat cggctggatc 3900

ctggagagca tgcccaccga ggtgctggag acccaggagc tgtgctccct ggccctggag 3960

taccgccgcg agtgcggccg cgactccgtg ctggagagcg tgaccgccat ggaccccagc 4020

aaggtgggcg tgcgctccca gtaccagcac ctgctgcgcc tggaggacgg caccgccatc 4080

gtgaacggcg ccaccgagtg gcgccccaag aacgccggcg ccaacggcgc catctccacc 4140

ggcaagacca gcaacggcaa ctccgtgtcc atggactaca aggaccacga cggcgactac 4200

aaggaccacg acatcgacta caaggacgac gacgacaagt gactcgaggc agcagcagct 4260

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 4320

ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 4380

tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accacccccа 4440

gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 4500

tgctatccct cagcgctgct cctgctcctg ctcactgccc ctcgcacagc cttggtttgg 4560

gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac 4620

gggaagtagt gggatgggaa cacaaatgga aagctt 4656

<210> 118
<211> 4721
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 118

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacgg cgcgccatgg ccaccaccag cctggcctcc gccttctgct 3120

ccatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc agctccgacc 3180

tgcagctgcg cgccggcaac gcccccacct ccctgaagat gatcaacggc accaagttca 3240

gctacaccga gagcctgaag cgcctgcccg actggtccat gctgttcgcc gtgatcacca 3300

ccatcttcag cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaagc 3360

tgccccagct gctggacgac cacttcggcc tgcacggcct ggtgttccgc cgcaccttcg 3420

ccatccgctc ctacgaggtg ggccccgacc gcagcacctc catcctggcc gtgatgaacc 3480

acatgcagga ggccaccctg aaccacgcca agagcgtggg catcctgggc gacggcttcg 3540

gcaccaccct ggagatgtcc aagcgcgacc tgatgtgggt ggtgcgccgc acccacgtgg 3600

ccgtggagcg ctaccccacc tggggcgaca ccgtggaggt ggagtgctgg atcggcgcca 3660

gcggcaacaa cggcatgcgc cgcgacttcc tggtgcgcga ctgcaagacc ggcgagatcc 3720

tgacccgctg cacctccctg agcgtgctga tgaacacccg cacccgccgc ctgagcacca 3780

tccccgacga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg 3840

acgacgagat caagaagctg cagaagctga acgactccac cgccgactac atccagggcg 3900

gcctgacccc ccgctggaac gacctggacg tgaaccagca cgtgaacaac ctgaagtacg 3960

tggcctgggt gttcgagacc gtgcccgaca gcatcttcga gtcccaccac atcagctcct 4020

tcaccctgga gtaccgccgc gagtgcaccc gcgactccgt gctgcgcagc ctgaccaccg 4080

tgagcggcgg cagctccgag gccggcctgg tgtgcgacca cctgctgcag ctggagggcg 4140

gcagcgaggt gctgcgcgcc cgcaccgagt ggcgccccaa gctgaccgac tccttccgcg 4200

gcatcagcgt gatccccgcc gagccccgcg tgatggacta caaggaccac gacggcgact 4260

acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag 4320

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 4380

acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 4440

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 4500

ccccagcatc cccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc 4560

tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 4620

tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 4680

tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 4721

<210> 119
<211> 4650
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 119

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttcggggtcg ccctcccggc catgggccgc 3120

ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg 3180

gagagcgggc gcgcccccga ctggtccatg ctgttcgccg tgatcaccac catcttcagc 3240

gccgccgaga agcagtggac caacctggag tggaagccca agcccaagct gccccagctg 3300

ctggacgacc acttcggcct gcacggcctg gtgttccgcc gcaccttcgc catccgctcc 3360

tacgaggtgg gccccgaccg cagcacctcc atcctggccg tgatgaacca catgcaggag 3420

gccaccctga accacgccaa gagcgtgggc atcctgggcg acggcttcgg caccaccctg 3480

gagatgtcca agcgcgacct gatgtgggtg gtgcgccgca cccacgtggc cgtggagcgc 3540

taccccacct ggggcgacac cgtggaggtg gagtgctgga tcggcgccag cggcaacaac 3600

ggcatgcgcc gcgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gacccgctgc 3660

acctccctga gcgtgctgat gaacacccgc acccgccgcc tgagcaccat ccccgacgag 3720

gtgcgcggcg agatcggccc cgccttcatc gacaacgtgg ccgtgaagga cgacgagatc 3780

aagaagctgc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc 3840

cgctggaacg acctggacgt gaaccagcac gtgaacaacc tgaagtacgt ggcctgggtg 3900

ttcgagaccg tgcccgacag catcttcgag tcccaccaca tcagctcctt caccctggag 3960

taccgccgcg agtgcacccg cgactccgtg ctgcgcagcc tgaccaccgt gagcggcggc 4020

agctccgagg ccggcctggt gtgcgaccac ctgctgcagc tggagggcgg cagcgaggtg 4080

ctgcgcgccc gcaccgagtg gcgccccaag ctgaccgact ccttccgcgg catcagcgtg 4140

atccccgccg agccccgcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200

gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260

gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct 4320

tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg 4380

tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440

ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat 4500

ccctcagcgc tgctcctgct cctgctcact gcccctcgca cagccttggt ttgggctccg 4560

cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620

tagtgggatg ggaacacaaa tggaaagctt 4650

<210> 120
<211> 4653
<212> ADN
<213> Secuencia artificial

<220>

<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 120

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180

gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240

agcgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag 3300

ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360

tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420

gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480

ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540

cgctacccca cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac 3600

aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660

tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720

gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag 3780

atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840

ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900

gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg 3960

gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc 4020

ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080

gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140

gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 4320

ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440

tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 121
<211> 4653
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 121

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120

cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180

cccgtgcgcg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240

agcgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag 3300

ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360

tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420

gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480

ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540

cgctacccca cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac 3600

aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660

tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720

gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag 3780

atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840

ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900

gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg 3960

gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc 4020

ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080

gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140

gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 4320

ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440

tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 122
<211> 4647
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 122

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcaccatgtc ggcgtgcccc 3120

gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc cacccgcctg 3180

aggggcgcg cccccgactg gtccatgctg ttcgccgtga tcaccaccat cttcagcgcc 3240

gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaagctgcc ccagctgctg 3300

gacgaccact tcggcctgca cggcctggtg ttccgccgca ccttcgccat ccgctcctac 3360

gaggtgggcc ccgaccgcag cacctccatc ctggccgtga tgaaccacat gcaggaggcc 3420

accctgaacc acgccaagag cgtgggcatc ctgggcgacg gcttcggcac caccctggag 3480

atgtccaagc gcgacctgat gtgggtggtg cgccgcaccc acgtggccgt ggagcgctac 3540

cccacctggg gcgacaccgt ggaggtggag tgctggatcg gcgccagcgg caacaacggc 3600

atgcgccgcg acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgctgcacc 3660

tccctgagcg tgctgatgaa cacccgcacc cgccgcctga gcaccatccc cgacgaggtg 3720

cgcggcgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga cgagatcaag 3780

aagctgcaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc 3840

tggaacgacc tggacgtgaa ccagcacgtg aacaacctga agtacgtggc ctgggtgttc 3900

gagaccgtgc ccgacagcat cttcgagtcc caccacatca gctccttcac cctggagtac 3960

cgccgcgagt gcacccgcga ctccgtgctg cgcagcctga ccaccgtgag cggcggcagc 4020

tccgaggccg gcctggtgtg cgaccacctg ctgcagctgg agggcggcag cgaggtgctg 4080

cgcgcccgca ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgatc 4140

cccgccgagc cccgcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac 4200

atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta 4260

tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga 4320

cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac 4380

gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatccgct 4440

tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc 4500

tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct 4560

gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag 4620

tgggatggga acacaaatgg aaagctt 4647

<210> 123
<211> 4721
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 123

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacgg cgcgccatgg ccaccacctc cctggcctcc gccttctgca 3120

gcatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc tccagcgacc 3180

tgcagctgcg cgccggcaac gcccagacct ccctgaagat gatcaacggc accaagttct 3240

cctacaccga gagcctgaag aagctgcccg actggtccat gctgttcgcc gtgatcacca 3300

ccatcttctc cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaacc 3360

ccccccagct gctggacgac cacttcggcc cccacggcct ggtgttccgc cgcaccttcg 3420

ccatccgcag ctacgaggtg ggccccgacc gctccaccag catcgtggcc gtgatgaacc 3480

acctgcagga ggccgccctg aaccacgcca agtccgtggg catcctgggc gacggcttcg 3540

gcaccaccct ggagatgtcc aagcgcgacc tgatctgggt ggtgaagcgc acccacgtgg 3600

ccgtggagcg ctaccccgcc tggggcgaca ccgtggaggt ggagtgctgg gtgggcgcct 3660

ccggcaacaa cggccgccgc cacgacttcc tggtgcgcga ctgcaagacc ggcgagatcc 3720

tgacccgctg cacctccctg agcgtgatga tgaacacccg cacccgccgc ctgagcaaga 3780

tccccgagga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg 3840

acgaggagat caagaagccc cagaagctga acgactccac cgccgactac atccagggcg 3900

gcctgacccc ccgctggaac gacctggaca tcaaccagca cgtgaacaac atcaagtacg 3960

tggactggat cctggagacc gtgcccgaca gcatcttcga gagccaccac atctcctcct 4020

tcaccatcga gtaccgccgc gagtgcacca tggacagcgt gctgcagtcc ctgaccaccg 4080

tgagcggcgg ctcctccgag gccggcctgg tgtgcgagca cctgctgcag ctggagggcg 4140

gcagcgaggt gctgcgcgcc aagaccgagt ggcgccccaa gctgaccgac tccttccgcg 4200

gcatcagcgt gatccccgcc gagtccagcg tgatggacta caaggaccac gacggcgact 4260

acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag 4320

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 4380

acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 4440

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 4500

ccccagcatc cccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc 4560

tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 4620

tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 4680

tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 4721

<210> 124
<211> 4650
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 124

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttcggggtcg ccctcccggc catgggccgc 3120

ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg 3180

gagagcgggc gcgcccccga ctggtccatg ctgttcgccg tgatcaccac catcttctcc 3240

gccgccgaga agcagtggac caacctggag tggaagccca agcccaaccc cccccagctg 3300

ctggacgacc acttcggccc ccacggcctg gtgttccgcc gcaccttcgc catccgcagc 3360

tacgaggtgg gccccgaccg ctccaccagc atcgtggccg tgatgaacca cctgcaggag 3420

gccgccctga accacgccaa gtccgtgggc atcctgggcg acggcttcgg caccaccctg 3480

gagatgtcca agcgcgacct gatctgggtg gtgaagcgca cccacgtggc cgtggagcgc 3540

taccccgcct ggggcgacac cgtggaggtg gagtgctggg tgggcgcctc cggcaacaac 3600

ggccgccgcc acgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gacccgctgc 3660

acctccctga gcgtgatgat gaacacccgc acccgccgcc tgagcaagat ccccgaggag 3720

gtgcgcggcg agatcggccc cgccttcatc gacaacgtgg ccgtgaagga cgaggagatc 3780

aagaagcccc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc 3840

cgctggaacg acctggacat caaccagcac gtgaacaaca tcaagtacgt ggactggatc 3900

ctggagaccg tgcccgacag catcttcgag agccaccaca tctcctcctt caccatcgag 3960

taccgccgcg agtgcaccat ggacagcgtg ctgcagtccc tgaccaccgt gagcggcggc 4020

tcctccgagg ccggcctggt gtgcgagcac ctgctgcagc tggagggcgg cagcgaggtg 4080

ctgcgcgcca agaccgagtg gcgccccaag ctgaccgact ccttccgcgg catcagcgtg 4140

atccccgccg agtccagcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200

gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260

gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct 4320

tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg 4380

tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440

ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat 4500

ccctcagcgc tgctcctgct cctgctcact gcccctcgca cagccttggt ttgggctccg 4560

cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620

tagtgggatg ggaacacaaa tggaaagctt 4650

<210> 125
<211> 4653
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 125

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460

tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180

gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240

tccgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa ccccccccag 3300

ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc 3360

agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag 3420

gaggccgccc tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc 3480

ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag 3540

cgctaccccg cctggggcga caccgtggag gtggagtgct gggtgggcgc ctccggcaac 3600

aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660

tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag 3720

gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgaggag 3780

atcaagaagc cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840

ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg 3900

atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc 3960

gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc 4020

ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag 4080

gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140

gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 4320

ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440

tcccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 126
<211> 3669
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Polinucleótido sintético

<400> 126

cccgtgatca cacaggtgcc ttgcgagcgt gatcacacta ttttgggggt cctacagtac 60

tgaaatggtg agaagtcgta ctgaaatcaa ggatgaacaa tgaaaatggt gctgtggtgg 120

cttctcaaag gtcaagaatc agtcgctcgc gtcaggaaat cgcggcgtca accagcgtgg 180

gcgcggtcag tggccccgca ctggtcacca tagcctctcc tgccacagta gcgatcccct 240

gggcgttcac tctcagcagc ggctgtactg cctcccagat tttcttcttc tggacctgcg 300

ggcgtgagag gatgagcagg gtggggccaa gggctcaatc ctgaacggcc ctcattcggt 360

ttccaatccc acaacacata cccacagcag gtcagaccac gcattcgcac catgcgcacc 420

aaataacgtg tccttacctg attgggtgtg gcaggctccg tggacaggag tgcctcgtcc 480

cccgcccaga cccgctcccc cgtcacggcg gcgtccggga cccgcagcgg ctccaccgcg 540

gtgtgatccg cgttggcggc gcagagcagc atcccagccg atttgacccc gcgcatgctc 600

cgaggcttga ggttggccag caccaccacc cgccggccga caaggtcctc cagggtcacg 660

tgccggacca ggccactcac gatggtgcga gggcccccct cctcgccgag gtcgatctgc 720

tcgacgtaca gactgcgaca tgcgtggcga gtggtcatca gaaggaagca ggtgtgcaga 780

aggggcacgt ggttggtatt gagagtagcc aaagctttgt gccaatcaga aagtcaacgc 840

agctgcctgc ctggctcgcg tacaattcct ttcttgcgct atgacacttc cagcaaaagg 900

tagggcgggc tgcgagacgg cttcccggcg ctgcatgcaa caccgatgat gcttcgaccc 960

cccgaagctc cttcggggct gcatgggcgc tccgatgccg ctccagggcg agcgctgttt 1020

aaatagccag gcccccgatt gcaaagacat tatagcgagc taccaaagcc atattcaaac 1080

acctagatca ctaccacttc tacacaggcc actcgagctt gtgatcgcac tccgctaagg 1140

gggcgcctct tcctcttcgt ttcagtcaca acccgcaaac ggcgcgccat gctgctgcag 1200

gccttcctgt tcctgctggc cggcttcgcc gccaagatca gcgcctccat gacgaacgag 1260

acgtccgacc gccccctggt gcacttcacc cccaacaagg gctggatgaa cgaccccaac 1320

ggcctgtggt acgacgagaa ggacgccaag tggcacctgt acttccagta caacccgaac 1380

gacaccgtct gggggacgcc cttgttctgg ggccacgcca cgtccgacga cctgaccaac 1440

tgggaggacc agcccatcgc catcgccccg aagcgcaacg actccggcgc cttctccggc 1500

tccatggtgg tggactacaa caacacctcc ggcttcttca acgacaccat cgacccgcgc 1560

cagcgctgcg tggccatctg gacctacaac accccggagt ccgaggagca gtacatctcc 1620

tacagcctgg acggcggcta caccttcacc gagtaccaga agaaccccgt gctggccgcc 1680

aactccaccc agttccgcga cccgaaggtc ttctggtacg agccctccca gaagtggatc 1740

atgaccgcgg ccaagtccca ggactacaag atcgagatct actcctccga cgacctgaag 1800

tcctggaagc tggagtccgc gttcgccaac gagggcttcc tcggctacca gtacgagtgc 1860

cccggcctga tcgaggtccc caccgagcag gaccccagca agtcctactg ggtgatgttc 1920

atctccatca accccggcgc cccggccggc ggctccttca accagtactt cgtcggcagc 1980

ttcaacggca cccacttcga ggccttcgac aaccagtccc gcgtggtgga cttcggcaag 2040

gactactacg ccctgcagac cttcttcaac accgacccga cctacgggag cgccctgggc 2100

atcgcgtggg cctccaactg ggagtactcc gccttcgtgc ccaccaaccc ctggcgctcc 2160

tccatgtccc tcgtgcgcaa gttctccctc aacaccgagt accaggccaa cccggagacg 2220

gagctgatca acctgaaggc cgagccgatc ctgaacatca gcaacgccgg cccctggagc 2280

cggttcgcca ccaacaccac gttgacgaag gccaacagct acaacgtcga cctgtccaac 2340

agcaccggca ccctggagtt cgagctggtg tacgccgtca acaccaccca gacgatctcc 2400

aagtccgtgt tcgcggacct ctccctctgg ttcaagggcc tggaggaccc cgaggagtac 2460

ctccgcatgg gcttcgaggt gtccgcgtcc tccttcttcc tggaccgcgg gaacagcaag 2520

gtgaagttcg tgaaggagaa cccctacttc accaaccgca tgagcgtgaa caaccagccc 2580

ttcaagagcg agaacgacct gtcctactac aaggtgtacg gcttgctgga ccagaacatc 2640

ctggagctgt acttcaacga cggcgacgtc gtgtccacca acacctactt catgaccacc 2700

gggaacgccc tgggctccgt gaacatgacg acgggggtgg acaacctgtt ctacatcgac 2760

aagttccagg tgcgcgaggt caagtgatta attaactcga ggcagcagca gctcggatag 2820

tatcgacaca ctctggacgc tggtcgtgtg atggactgtt gccgccacac ttgctgcctt 2880

gacctgtgaa tatccctgcc gcttttatca aacagcctca gtgtgtttga tcttgtgtgt 2940

acgcgctttt gcgagttgct agctgcttgt gctatttgcg aataccaccc ccagcatccc 3000

cttccctcgt ttcatatcgc ttgcatccca accgcaactt atctacgctg tcctgctatc 3060

cctcagcgct gctcctgctc ctgctcactg cccctcgcac agccttggtt tgggctccgc 3120

ctgtattctc ctggtactgc aacctgtaaa ccagcactgc aatgctgatg cacgggaagt 3180

agtgggatgg gaacacaaat ggaaagcttg agctcggtac ccgtacccat cagcatccgg 3240

gtgaatcttg gcctccaaga tatggccaat cctcacatcc agcttggcaa aatcgactag 3300

actgtctgca agtgggaatg tggagcacaa ggttgcttgt agcgatcgac agactggtgg 3360

ggtacattga caggtgggca gcgccgcatc catcgtgcct gacgcgagcg ccgccggttg 3420

ctcgcccgtg cctgccgtca aagagcggca gagaaatcgg gaaccgaaaa cgtcacattg 3480

cctgatgttg ttacatgctg gactagactt tcttggcgtg ggtctgctcc tcgccaggtg 3540

cgcgacgcct cggggctggg tgcgagggag ccgtgcggcc acgcatttga caagacccaa 3600

agctcgcatc tcagacggtc aaccgttcgt attatacatt caacatatgg tacatacgca 3660

aaaagcatg 3669

<210> 127
<211> 39
<212> PRT
<213> Prototheca moriformis

<400> 127

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

Pro Arg Pro Arg Val Ala Val Arg Ala Gln Ser Ala Ser Gln Val Leu
20 25 30

Glu Ser Gly Arg Ala Gln Leu
35

<210> 128
<211> 40
<212> PRT
<213> Prototheca moriformis

<400> 128

Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15

Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
20 25 30

Ala Leu Arg Gly Arg Ala Gln Leu
35 40

<210> 129
<211> 36
<212> PRT
<213> Prototheca moriformis

<400> 129

Met Ala Ser Ala Ala Phe Thr Met Ser Ala Cys Pro Ala Met Thr Gly
1 5 10 15

Arg Ala Pro Gly Ala Arg Arg Ser Gly Arg Pro Val Ala Thr Arg Leu
20 25 30

Arg Gly Arg Ala
35

<210> 130
<211> 40
<212> PRT
<213> Chlorella protothecoides

<400> 130

Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15

Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30

Pro Val Arg Gly Arg Ala Gln Leu
35 40

<210> 131
<211> 87
<212> PRT
<213> Cuphea hookeriana

<400> 131

Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65 70 75 80

Pro Pro Arg Thr Phe Leu His
85

<210> 132
<211> 60
<212> PRT
<213> Umbellularia californica

<400> 132

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu
50 55 60

<210> 133
<211> 60
<212> PRT
<213> Cinnamomum camphora

<400> 133

```
Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1               5                  10                  15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
            20                  25                  30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly
        35                  40                  45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu
    50                  55                  60
```

<210> 134
<211> 1104
<212> ADN
<213> Prototheca moriformis

<400> 134

```
atggcaccga ccagcctgct tgccagtact ggcgtctctt ccgcttctct gtggtcctct      60
gcgcgctcca gcgcgtgcgc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg     120
ccgctgccca tgcagcgccg ctgcttccga acagtggcgg tcagggccgc acccgcggta     180
gccgtccgtc cggaacccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg     240
gaggacaagc gcatcttcct ggaggagcac cgcattcggg gcaacgaggt gggcccctcg     300
cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg     360
gtggccatgt ggggccggag ctcggagggt ttcgcgacgg acccggagct gcaggaggcg     420
ggtctcatct ttgtgatgac gcgcatgcag atccaaatgt accgctaccc gcgctggggc     480
gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag     540
tgggtgctgc gcgacaagct gaccggcgag gcgctgggcg cggccacctc cagctgggtc     600
atgatcaaca tccgcacgcg ccggccgtgc cgcatgcccg agctcgtccg cgtcaagtcg     660
gccttcttcg cgcgcgagcc gccgcgcctg gcgctgccgc ccacggtcac gcgcgccaag     720
ctgcccaaca tcgcgacgcc ggcgccgctg cgcgggcacc gccaggtcgc gcgccgcacc     780
gacatggaca tgaacgggca cgtgaacaac gtggcctacc tggcctggtg cctggaggcc     840
gtgcccgagc acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc     900
gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg     960
ctcacgcaca acggcgccgg ccgcaacccc tcctgcttcg tccatagcat tctgcgcgcc    1020
gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc    1080
aagccgccca aggcgagcca ctga                                           1104
```

<210> 135
<211> 1104
<212> ADN
<213> Prototheca moriformis

<400> 135

```
atggcaccga ccagcctgct tgcccgtact ggcgtctctt ccgcttctct gtgctcctct      60
acgcgctccg gcgcgtgcgc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg     120
ccgctgccca tgcagcgccg ctgcttccga acagtggctg tcagggccgc acccgcagta     180
gccgtccgtc cggaacccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg     240
gaggacaagc gcatcttcct ggaggagcac cgcattcgtg gcaacgaggt gggcccctcg     300
cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg     360
gtggccatgt ggggtcggag ctcggagggt ttcgcgacgg acccggagct gcaggaggcg     420
ggcctcatct ttgtgatgac gcgcatgcag atccaaatgt accgctaccc gcgctggggc     480
gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag     540
tgggtgctgc gcgacaagct gaccggcgag gcgctgggcg cggccacctc cagctgggtc     600
atgatcaaca tccgcacgcg ccggccgtgc cgcatgcccg agctcgtccg cgtcaagtcg     660
gccttcttcg cgcgcgagcc gccgcgcctg gcgctgccgc ccgcggtcac gcgtgccaag     720
ctgcccaaca tcgcgacgcc ggcgccgctg cgcgggcacc gccaggtcgc gcgccgcacc     780
gacatggaca tgaacggcca cgtgaacaac gttgcctacc tggcctggtg cctggaggcc     840
gtgcccgagc acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc     900
gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg     960
ctcacgcaca acggcgccgg ccgcaacccc tcctgcttcg tccatagcat tctgcgcgcc    1020
gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc    1080
aagccgccca aggcgagcca ctga                                           1104
```

<210> 136
<211> 367
<212> PRT
<213> Prototheca moriformis

<400> 136

```
Met Ala Pro Thr Ser Leu Leu Ala Ser Thr Gly Val Ser Ser Ala Ser
1               5                  10                  15

Leu Trp Ser Ser Ala Arg Ser Ser Ala Cys Ala Phe Pro Val Asp His
            20                  25                  30

Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
        35                  40                  45
```

```
Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
    50                  55                  60

Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65                  70                  75                  80

Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
                85                  90                  95

Val Gly Pro Ser Gln Arg Leu Thr Ile Thr Ala Val Ala Asn Ile Leu
            100                 105                 110

Gln Glu Ala Ala Gly Asn His Ala Val Ala Met Trp Gly Arg Ser Ser
        115                 120                 125

Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
    130                 135                 140

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145                 150                 155                 160

Asp Leu Met Gln Val Glu Thr Trp Phe Gln Thr Ala Gly Lys Leu Gly
                165                 170                 175

Ala Gln Arg Glu Trp Val Leu Arg Asp Lys Leu Thr Gly Glu Ala Leu
            180                 185                 190

Gly Ala Ala Thr Ser Ser Trp Val Met Ile Asn Ile Arg Thr Arg Arg
        195                 200                 205

Pro Cys Arg Met Pro Glu Leu Val Arg Val Lys Ser Ala Phe Phe Ala
    210                 215                 220

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Thr Val Thr Arg Ala Lys
225                 230                 235                 240

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
                245                 250                 255

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
            260                 265                 270

Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
        275                 280                 285

Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
```

```
            290                 295                 300

Gly Asp Val Ile Ser Ser Gln Ala Glu Gln Ile Pro Pro Gln Glu Ala
305                 310                 315                 320

Leu Thr His Asn Gly Ala Gly Arg Asn Pro Ser Cys Phe Val His Ser
                325                 330                 335

Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
            340                 345                 350

Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
        355                 360                 365
```

<210> 137
<211> 367
<212> PRT
<213> Prototheca moriformis

<400> 137

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

Leu Cys Ser Ser Thr Arg Ser Gly Ala Cys Ala Phe Pro Val Asp His
            20                  25                  30

Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
        35                  40                  45

Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
    50                  55                  60

Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65                  70                  75                  80

Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
                85                  90                  95

Val Gly Pro Ser Gln Arg Leu Thr Ile Thr Ala Val Ala Asn Ile Leu
            100                 105                 110

Gln Glu Ala Ala Gly Asn His Ala Val Ala Met Trp Gly Arg Ser Ser
        115                 120                 125

Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
    130                 135                 140

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145                 150                 155                 160
```

```
Asp Leu Met Gln Val Glu Thr Trp Phe Gln Thr Ala Gly Lys Leu Gly
                165                 170                 175

Ala Gln Arg Glu Trp Val Leu Arg Asp Lys Leu Thr Gly Glu Ala Leu
            180                 185                 190

Gly Ala Ala Thr Ser Ser Trp Val Met Ile Asn Ile Arg Thr Arg Arg
        195                 200                 205

Pro Cys Arg Met Pro Glu Leu Val Arg Val Lys Ser Ala Phe Phe Ala
    210                 215                 220

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Ala Val Thr Arg Ala Lys
225                 230                 235                 240

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
                245                 250                 255

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
            260                 265                 270

Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
        275                 280                 285

Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
    290                 295                 300

Gly Asp Val Ile Ser Ser Gln Ala Glu Gln Ile Pro Pro Gln Glu Ala
305                 310                 315                 320

Leu Thr His Asn Gly Ala Gly Arg Asn Pro Ser Cys Phe Val His Ser
                325                 330                 335

Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
            340                 345                 350

Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
        355                 360                 365
```

<210> 138
<211> 328
<212> PRT
<213> Cuphea hookeriana

<400> 138

```
Gln Leu Pro Asp Trp Ser Arg Leu Leu Thr Ala Ile Thr Thr Val Phe
1               5                   10                  15
```

```
Val Lys Ser Lys Arg Pro Asp Met His Asp Arg Lys Ser Lys Arg Pro
            20                  25                  30

Asp Met Leu Val Asp Ser Phe Gly Leu Glu Ser Thr Val Gln Asp Gly
        35                  40                  45

Leu Val Phe Arg Gln Ser Phe Ser Ile Arg Ser Tyr Glu Ile Gly Thr
    50                  55                  60

Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu Thr
65                  70                  75                  80

Ser Leu Asn His Cys Lys Ser Thr Gly Ile Leu Leu Asp Gly Phe Gly
                85                  90                  95

Arg Thr Leu Glu Met Cys Lys Arg Asp Leu Ile Trp Val Val Ile Lys
            100                 105                 110

Met Gln Ile Lys Val Asn Arg Tyr Pro Ala Trp Gly Asp Thr Val Glu
        115                 120                 125

Ile Asn Thr Arg Phe Ser Arg Leu Gly Lys Ile Gly Met Gly Arg Asp
    130                 135                 140

Trp Leu Ile Ser Asp Cys Asn Thr Gly Glu Ile Leu Val Arg Ala Thr
145                 150                 155                 160

Ser Ala Tyr Ala Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys Leu
                165                 170                 175

Pro Tyr Glu Val His Gln Glu Ile Val Pro Leu Phe Val Asp Ser Pro
            180                 185                 190

Val Ile Glu Asp Ser Asp Leu Lys Val His Lys Phe Lys Val Lys Thr
        195                 200                 205

Gly Asp Ser Ile Gln Lys Gly Leu Thr Pro Gly Trp Asn Asp Leu Asp
    210                 215                 220

Val Asn Gln His Val Ser Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu
225                 230                 235                 240

Ser Met Pro Thr Glu Val Leu Glu Thr Gln Glu Leu Cys Ser Leu Ala
                245                 250                 255

Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Glu Ser Val
            260                 265                 270
```

Thr Ala Met Asp Pro Ser Lys Val Gly Val Arg Ser Gln Tyr Gln His
275 280 285

Leu Leu Arg Leu Glu Asp Gly Thr Ala Ile Val Asn Gly Ala Thr Glu
290 295 300

Trp Arg Pro Lys Asn Ala Gly Ala Asn Gly Ala Ile Ser Thr Gly Lys
305 310 315 320

Thr Ser Asn Gly Asn Ser Val Ser
325

<210> 139
<211> 322
<212> PRT
<213> Umbellularia californica

<400> 139

Pro Asp Trp Ser Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala
1 5 10 15

Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Lys Leu
20 25 30

Pro Gln Leu Leu Asp Asp His Phe Gly Leu His Gly Leu Val Phe Arg
35 40 45

Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr
50 55 60

Ser Ile Leu Ala Val Met Asn His Met Gln Glu Ala Thr Leu Asn His
65 70 75 80

Ala Lys Ser Val Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu
85 90 95

Met Ser Lys Arg Asp Leu Met Trp Val Val Arg Arg Thr His Val Ala
100 105 110

Val Glu Arg Tyr Pro Thr Trp Gly Asp Thr Val Glu Val Glu Cys Trp
115 120 125

Ile Gly Ala Ser Gly Asn Asn Gly Met Arg Arg Asp Phe Leu Val Arg
130 135 140

Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val
145 150 155 160

```
Leu Met Asn Thr Arg Thr Arg Arg Leu Ser Thr Ile Pro Asp Glu Val
                165                 170                 175

Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp
            180                 185                 190

Asp Glu Ile Lys Lys Leu Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr
        195                 200                 205

Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln
    210                 215                 220

His Val Asn Asn Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro
225                 230                 235                 240

Asp Ser Ile Phe Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr
                245                 250                 255

Arg Arg Glu Cys Thr Arg Asp Ser Val Leu Arg Ser Leu Thr Thr Val
            260                 265                 270

Ser Gly Gly Ser Ser Glu Ala Gly Leu Val Cys Asp His Leu Leu Gln
        275                 280                 285

Leu Glu Gly Gly Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro
    290                 295                 300

Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro
305                 310                 315                 320

Arg Val
```

<210> 140
<211> 345
<212> PRT
<213> Cinnamomum camphora

<400> 140

```
Pro Asp Trp Ser Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala
1               5                   10                  15

Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Asn Pro
            20                  25                  30

Pro Gln Leu Leu Asp Asp His Phe Gly Pro His Gly Leu Val Phe Arg
        35                  40                  45
```

```
Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr
    50                  55                  60

Ser Ile Val Ala Val Met Asn His Leu Gln Glu Ala Ala Leu Asn His
65                  70                  75                  80

Ala Lys Ser Val Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu
                85                  90                  95

Met Ser Lys Arg Asp Leu Ile Trp Val Val Lys Arg Thr His Val Ala
            100                 105                 110

Val Glu Arg Tyr Pro Ala Trp Gly Asp Thr Val Glu Val Glu Cys Trp
        115                 120                 125

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

Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val
145                 150                 155                 160

Met Met Asn Thr Arg Thr Arg Arg Leu Ser Lys Ile Pro Glu Glu Val
                165                 170                 175

Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp
            180                 185                 190

Glu Glu Ile Lys Lys Pro Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr
        195                 200                 205

Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Ile Asn Gln
    210                 215                 220

His Val Asn Asn Ile Lys Tyr Val Asp Trp Ile Leu Glu Thr Val Pro
225                 230                 235                 240

Asp Ser Ile Phe Glu Ser His His Ile Ser Ser Phe Thr Ile Glu Tyr
                245                 250                 255

Arg Arg Glu Cys Thr Met Asp Ser Val Leu Gln Ser Leu Thr Thr Val
            260                 265                 270

Ser Gly Gly Ser Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln
        275                 280                 285

Leu Glu Gly Gly Ser Glu Val Leu Arg Ala Lys Thr Glu Trp Arg Pro
    290                 295                 300
```

```
Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Ser
305                 310                 315                 320

Ser Val Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp
                325                 330                 335

Ile Asp Tyr Lys Asp Asp Asp Asp Lys
            340                 345
```

<210> 141
<211> 4
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de Secuencia artificial: Péptido sintético

<400> 141

```
Lys Asp Glu Leu
1
```

<210> 142
<211> 6
<212> PRT
<213> Desconocido

<220>
<223> Descripción de Desconocido: Secuencia de acil-ACP tioesterasa grasa vegetal más alta

<400> 142

```
Leu Asp Met Asn Gln His
1               5
```

<210> 143
<211> 6
<212> PRT
<213> Desconocido

<220>
<223> Descripción de Desconocido: Secuencia de acil-ACP tioesterasa grasa de alga

<400> 143

```
Met Asp Met Asn Gly His
1               5
```

REIVINDICACIONES

1.Una célula del género *Prototheca que* comprende un gen exógeno, en donde el gen exógeno está en un ligamiento operable con un promotor, y codifica una acil-ACP tioesterasa grasa.

2. La célula de la reivindicación 1, en la que el promotor es de un gen endógeno a una especie del género *Prototheca.*

3. La célula de la reivindicación 1 o 2, en la que la acil-ACP tioesterasa grasa tiene actividad de hidrólisis hacia uno o más sustratos de acil-ACP grasos de longitud de cadena C8, C10, C12 o C14, opcionalmente en la que la acil-ACP tioesterasa tiene al menos el 50% de identidad de aminoácidos con una o más secuencias seleccionadas del grupo que consiste de las SEQ ID NO: 59, 61, 63 y 138-140.

4. La célula de la reivindicación 3, en la que la secuencia codificante del gen exógeno comprende una secuencia de direccionamiento de plástido de microalgas, preferiblemente de una especie del género *Prototheca,* y más preferiblemente en la que la secuencia de direccionamiento de plástido tiene al menos un 25% de identidad de secuencia de aminoácidos con una o más de SEQ ID NO: 127-133 y es capaz de dirigir una proteína codificada por el gen exógeno al plastidio.

5. La célula de la reivindicación 1, en la que el promotor está regulado por incremento al menos 3 veces en una célula del género *Prototheca* cuando el entorno extracelular cambia de contener al menos 5 mM de nitrógeno a no contener nitrógeno, por ejemplo, en donde el promotor comprende un segmento de 50 o más nucleótidos de una de las SEQ ID NO: 91-102.

6. La célula de la reivindicación 1, en la que:

(i) la célula tiene una secuencia de ARNr 23S con al menos un 75% de identidad de nucleótidos con una o más de las SEQ ID NO: 11-19, y / o
(ii) la célula *es Prototheca moriformis, Prototheca krugani, Prototheca stagnora* o *Prototheca zopfii.*

7. Una célula del género *Prototheca que* comprende un gen exógeno que codifica una acil-ACP tioesterasa grasa, y un perfil lipídico de al menos el 4% de C8-C14.

8. La célula de la reivindicación 7, en la que:

(i) la cantidad de C8 es de al menos el 0,3%;
(ii) la cantidad de C10 es de al menos el 2%;
(iii) la cantidad de C12 es de al menos el 2%;
(iv) la cantidad de C14 es de al menos el 4%; o
(v) la cantidad de C8-C14 es de al menos el 30%.

9. La célula de la reivindicación 7, que comprende además un gen exógeno de invertasa de sacarosa.

10. La célula de la reivindicación 7, en la que la célula es como se define en la reivindicación 6, y opcionalmente en la que el gen exógeno que codifica la acil-ACP tioesterasa grasa está integrado en un cromosoma de la célula.

11. Un método para elaborar una composición de triglicéridos de al menos un 4% de C8-C14 que comprende cultivar la célula de la reivindicación 7.

12. El método de la reivindicación 11, en el que

(i) la cantidad de C8 es de al menos el 0,3%;
(ii) la cantidad de C10 es de al menos el 2%;
(iii) la cantidad de C12 es de al menos el 2%;
(iv) la cantidad de C14 es de al menos el 4%; o
(v) la cantidad de C8-C14 es de al menos el 30%.

13. Un método para elaborar una composición de triglicéridos que comprende cultivar la célula de la reivindicación 7, en donde la célula comprende además un gen exógeno que codifica una invertasa de sacarosa y se proporciona sacarosa como fuente de carbono, opcionalmente en donde la invertasa de sacarosa tiene al menos un 50% de identidad de aminoácidos con una o más de las SEQ ID NO: 3, 20-29 y 90.

14.. El método de la reivindicación 11 o 13 que comprende además lisar la célula y opcionalmente separar los lípidos del lisado celular.

15. Una composición de aceite de triglicérido que comprende un perfil lipídico de:

a. al menos un 4% de C8-C14; y
b. uno o más de los siguientes atributos:

i. menos de 0,4 microgramos/ml de carotenoides totales;
ii. menos de 0,001 microgramos/ml de licopeno;
iii. menos de 0,02 microgramos/ml de betacaroteno;
iv. menos de 0,02 miligramos de clorofila por kilogramo de aceite;
v. 0,40-0,60 miligramos de gamma tocoferol por 100 gramos de aceite;
vi. 3-9 mg de campesterol por 100 gramos de aceite; y
vii menos de 0,5 miligramos de tocotrienoles totales por gramo de aceite,

y opcionalmente en donde

c. la cantidad de C8 es de al menos el 0,3%,
d. la cantidad de C10 es de al menos el 2%,
e. la cantidad de C12 es de al menos el 2%,
f. la cantidad de C12 es de al menos el 4%, o
g. la cantidad de C10-C14 es de al menos el 30%,

en donde dicho aceite se puede obtener mediante el método de cualquiera de las reivindicaciones 11 a 13 y además extrayendo la composición de aceite de triglicéridos mediante extracción con solvente o usando una prensa expulsora.

Figura 1

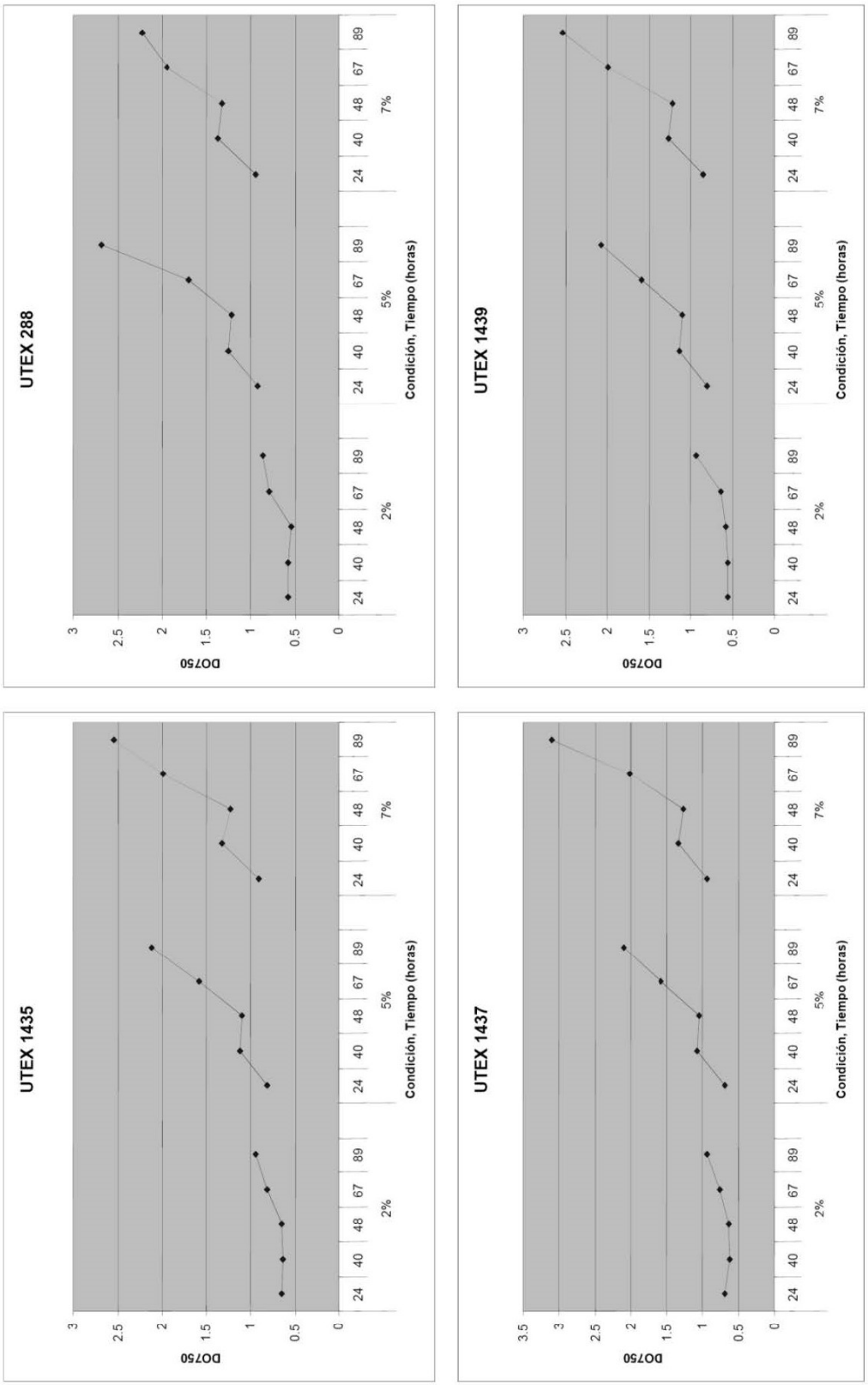

UTEX 288
UTEX 1439
UTEX 1435
UTEX 1437
DO750
Condición, Tiempo (horas)
2%
5%
7%
24
40
48
67
89

Figura 2

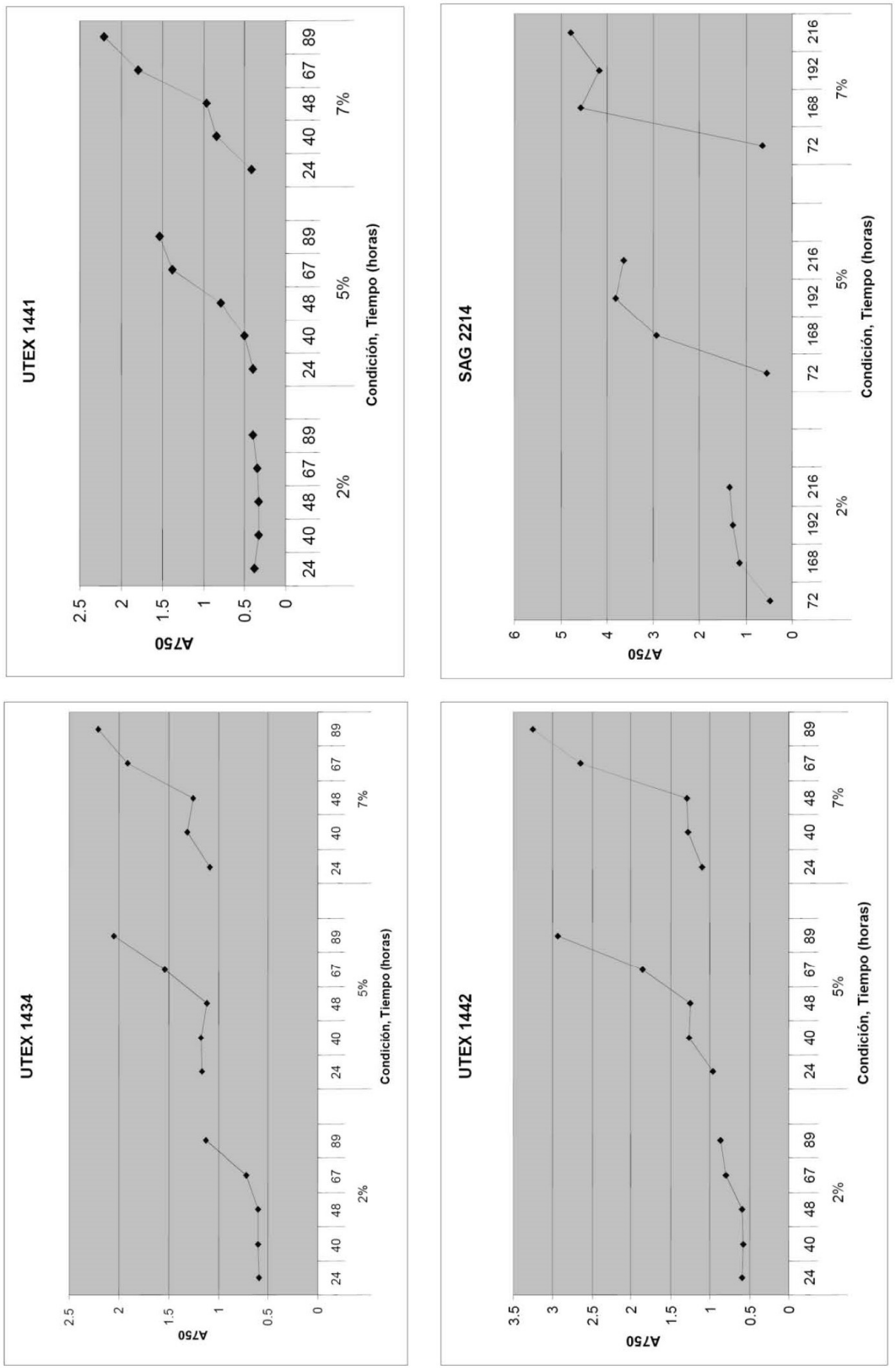
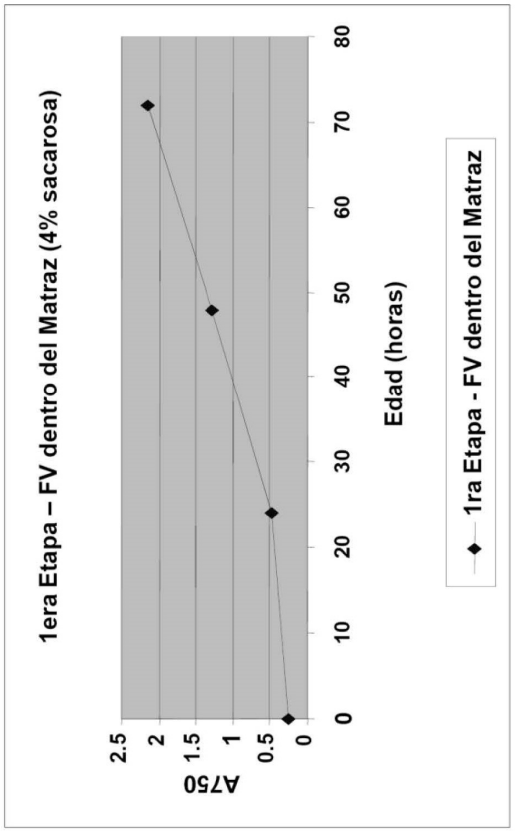

UTEX 1441
A750
Condición, Tiempo (horas)
2%
5%
7%
SAG 2214
UTEX 1434
UTEX 1442

Figura 3


1era Etapa – FV dentro del Matraz (4% sacarosa)
A750
2.5
2
1.5
1
0.5
0
0
10
20
30
40
50
60
70
80
Edad (horas)
1ra Etapa - FV dentro del Matraz

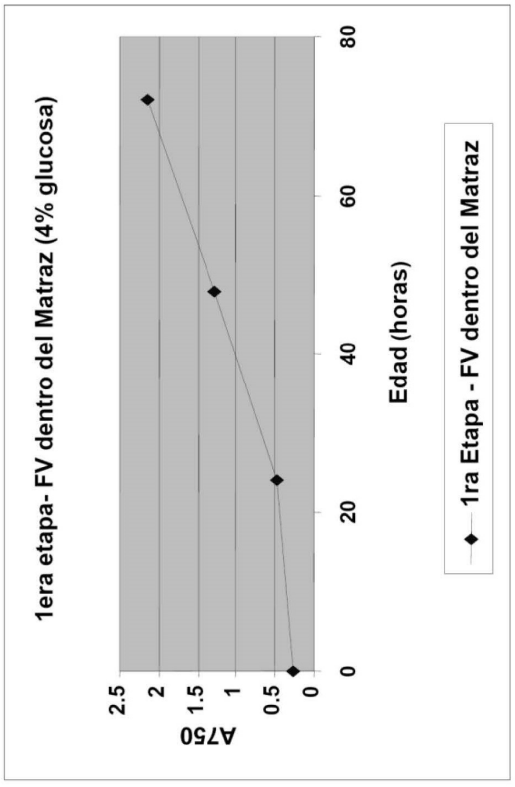

1era etapa- FV dentro del Matraz (4% glucosa)
A750
2.5
2
1.5
1
0.5
0
0
20
40
60
80
Edad (horas)
1ra Etapa - FV dentro del Matraz

Figura 4

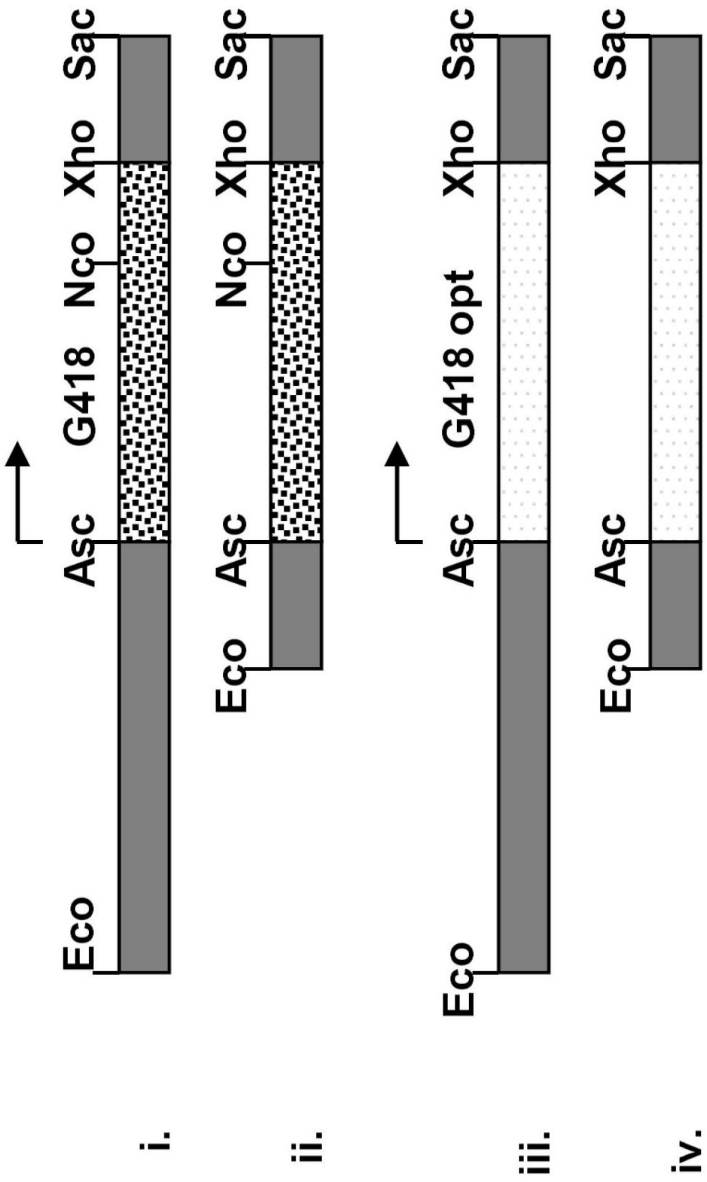

Sac
Xho
Nco
G418
Asc
Eco
i.
Sac
Xho
Nco
Asc
Eco
ii.
Sac
Xho
G418 opt
Asc
Eco
iii.
Sac
Xho
Asc
Eco
iv.

Figura 5

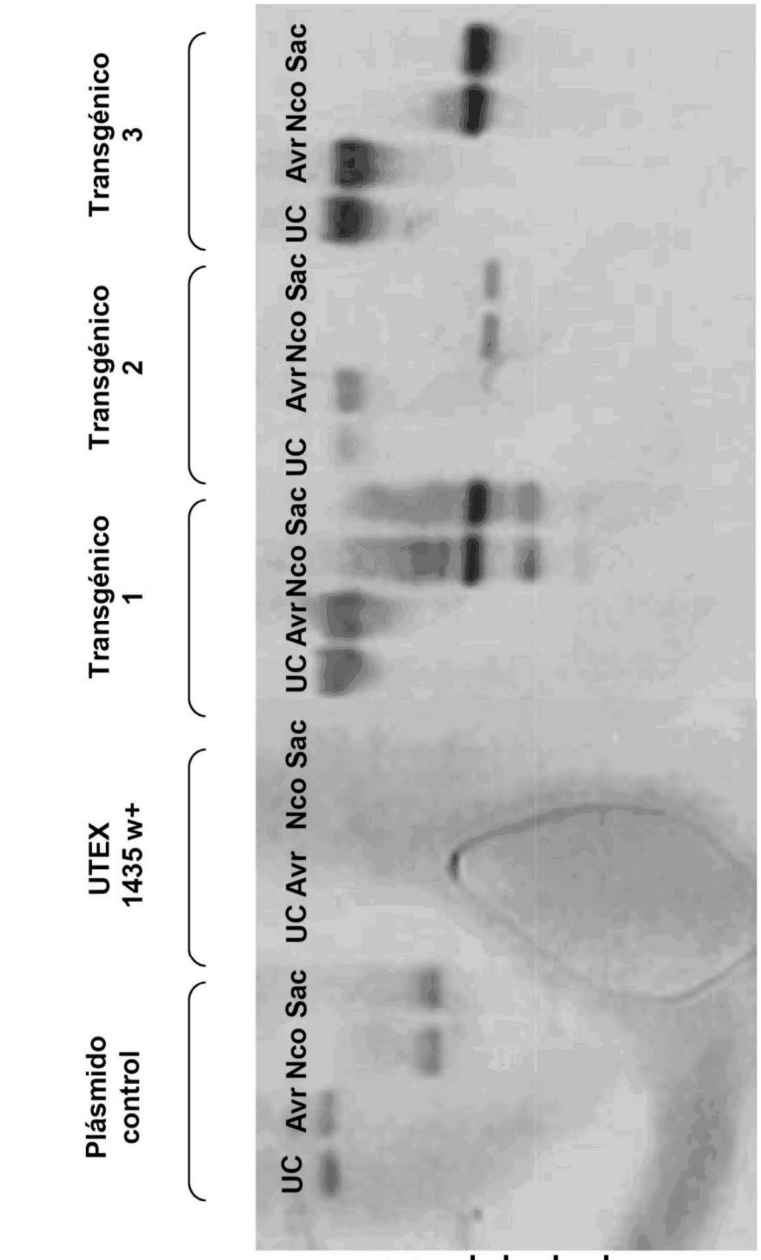

Plásmido control
UTEX 1435 w+
Transgénico 1
Transgénico 2
Transgénico 3
UC Avr Nco Sac
UC Avr Nco Sac
UC Avr Nco Sac UC
Avr Nco Sac UC
Avr Nco Sac
Tamaño (pb)
8,000 -
6,000 -
4,000 -
3,000 -
2,000 -
1,500 -

Figura 6

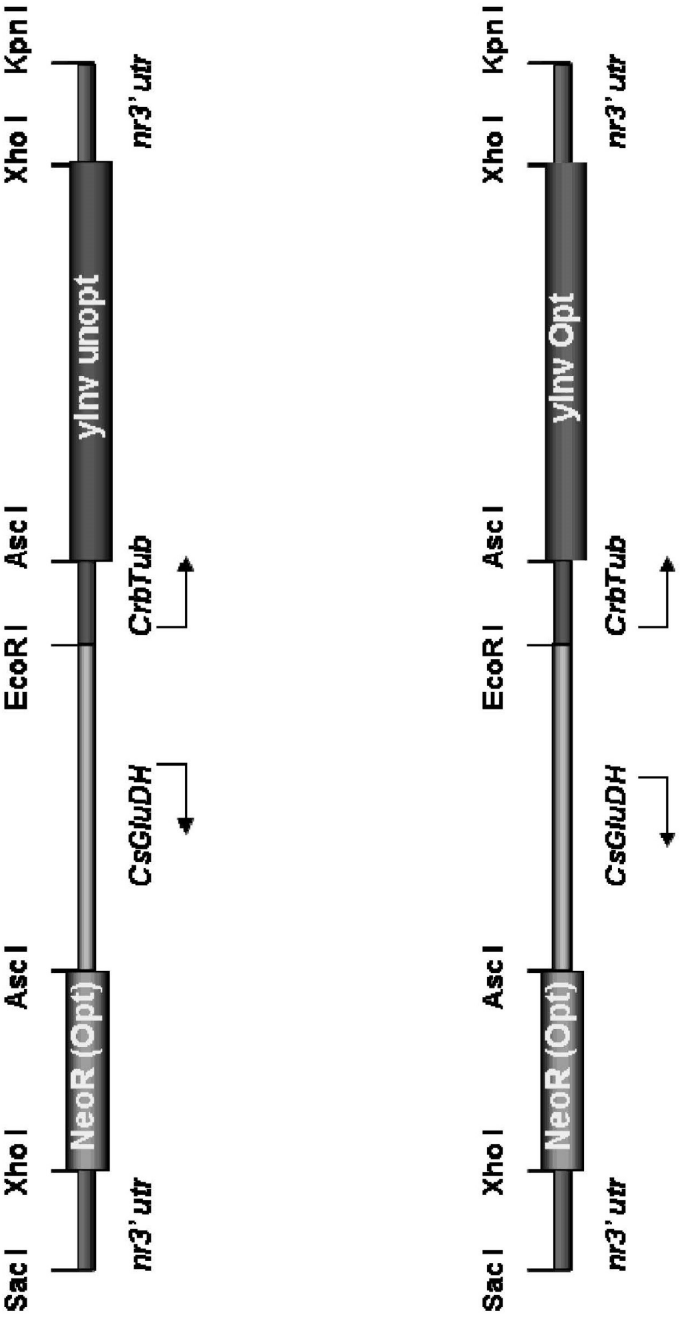

Sac I
Xho I
NeoR (Opt)
Asc I
EcoR I
CsGluDH
nr3' utr
CrbTub
Asc I
ylnv unopt
Xho I
Kpn I
nr3' utr
Sac I
Xho I
NeoR (Opt)
Asc I
EcoR I
CsGluDH
CrbTub
Asc I
ylnv Opt
Xho I
Kpn I
nr3' utr

Figura 7a

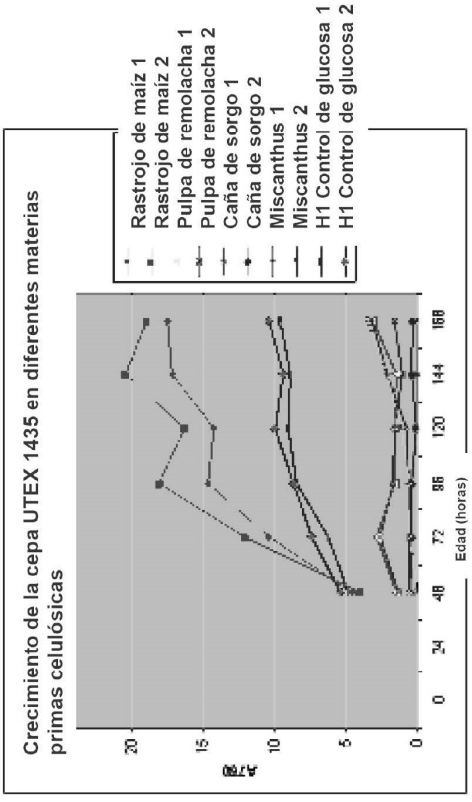


Figura 7c

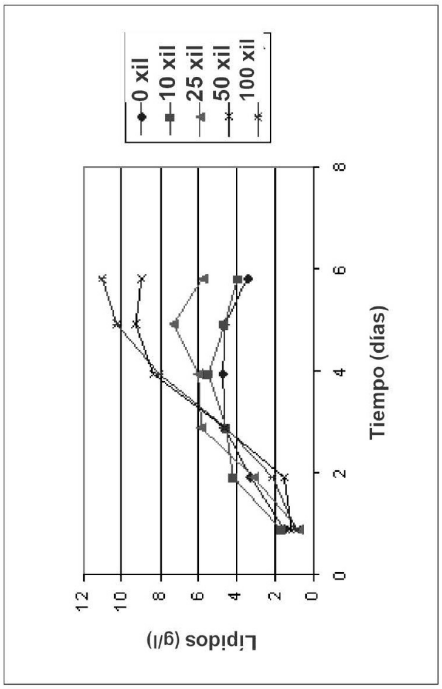


Figura 7b

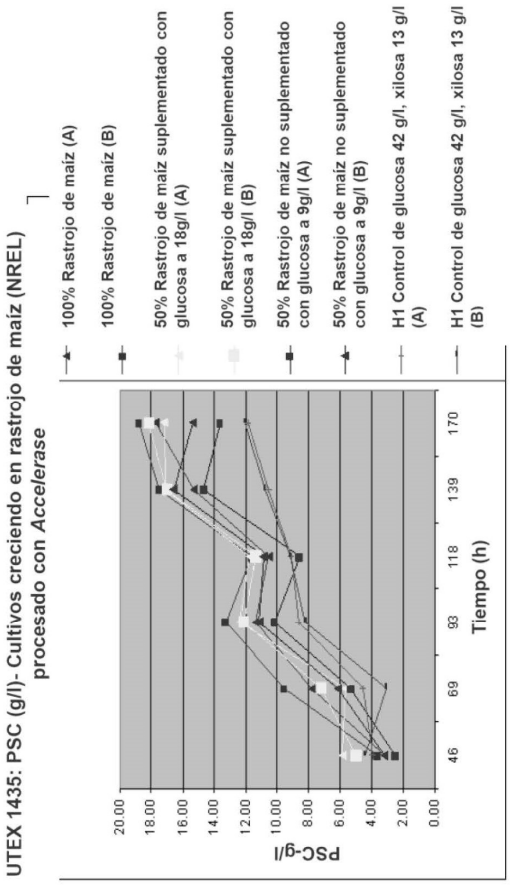


Figura 7d

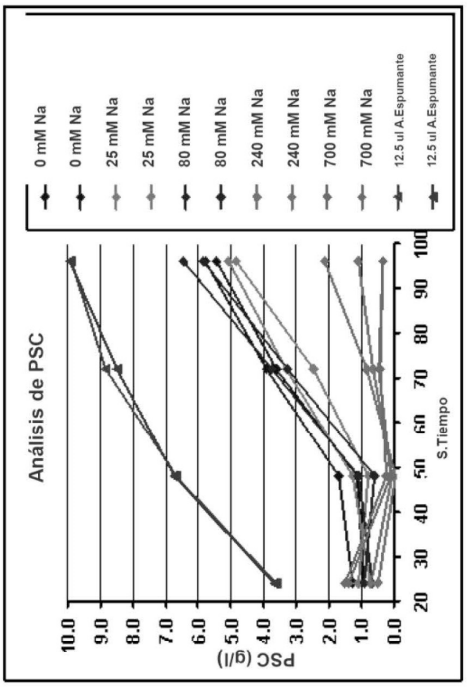


Figura 8

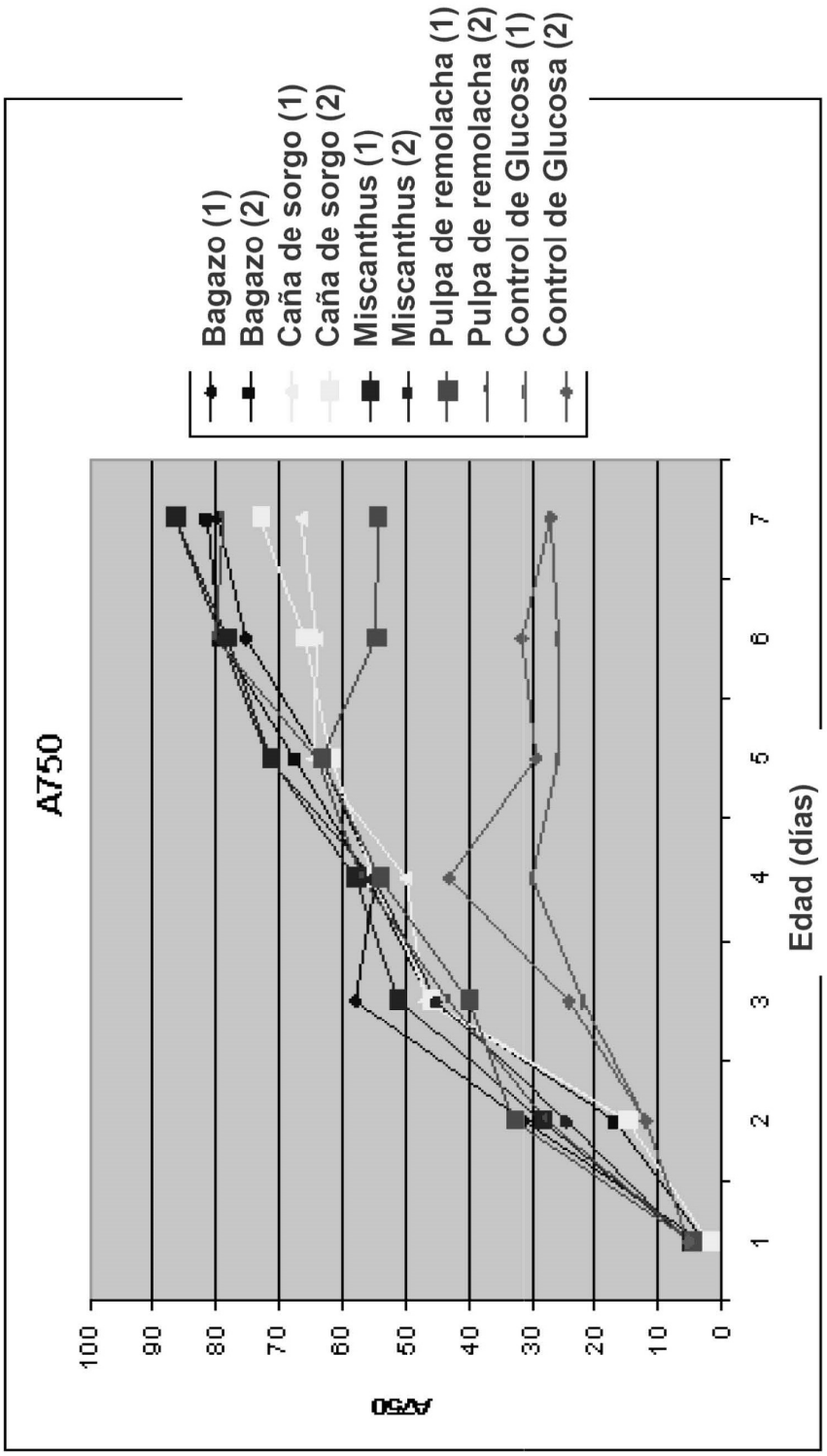

A750
Edad (días)
Bagazo (1)
Bagazo (2)
Caña de sorgo (1)
Caña de sorgo (2)
Miscanthus (1)
Miscanthus (2)
Pulpa de remolacha (1)
Pulpa de remolacha (2)
Control de Glucosa (1)
Control de Glucosa (2)

Figura 9

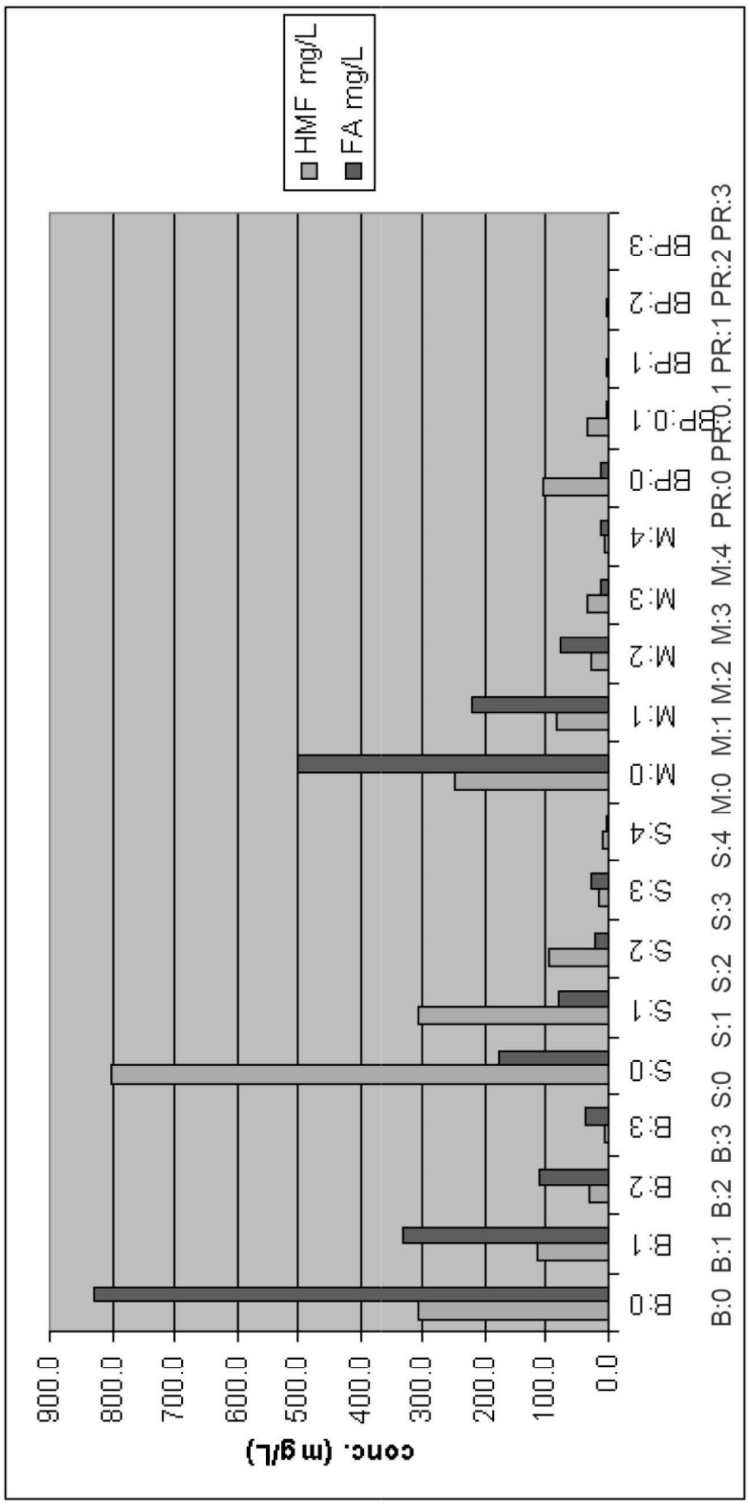

HMF mg/L
FA mg/L
conc. (m g/L)
900.0
800.0
700.0
600.0
500.0
400.0
300.0
200.0
100.0
0.0
B:0 B:1 B:2 B:3 S:0 S:1 S:2 S:3 S:4 M:0 M:1 M:2 M:3 M:4 PR:0 PR:0.1 PR:1 PR:2 PR:3

Figura 10

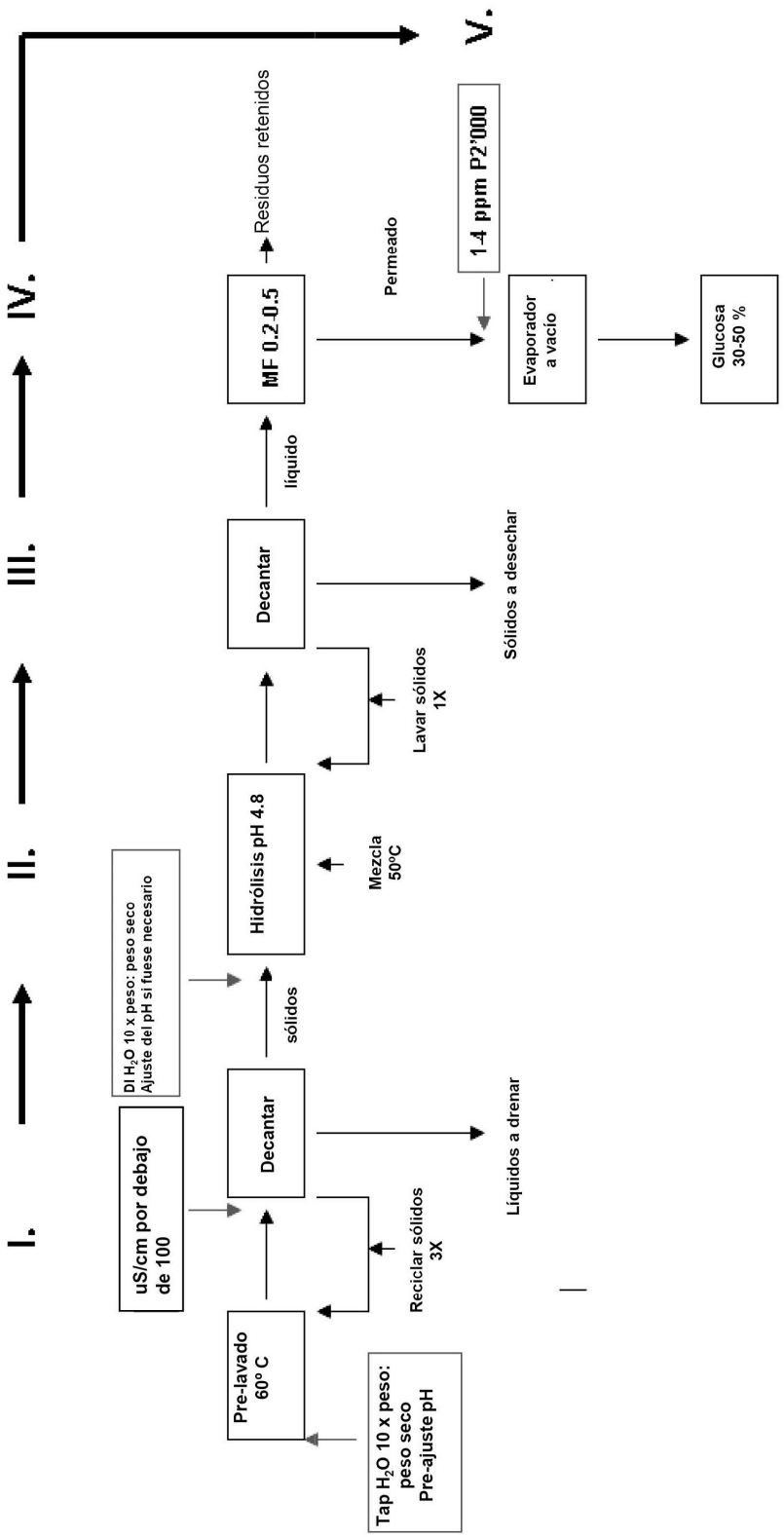

I.
II.
III.
IV.
V.
Tap H_2O 10 x peso: peso seco Pre-ajuste pH
Pre-lavado 60° C
uS/cm por debajo de 100
Reciclar sólidos 3X
Decantar
Líquidos a drenar
DI H_2O 10 x peso: peso seco Ajuste del pH si fuese necesario
sólidos
Hidrólisis pH 4.8
Mezcla 50°C
Lavar sólidos 1X
Decantar
Sólidos a desechar
líquido
MF 0.2 0.5
Residuos retenidos
Permeado
1-4 ppm P2'000
Evaporador a vacío
Glucosa 30-50 %

Figura 11

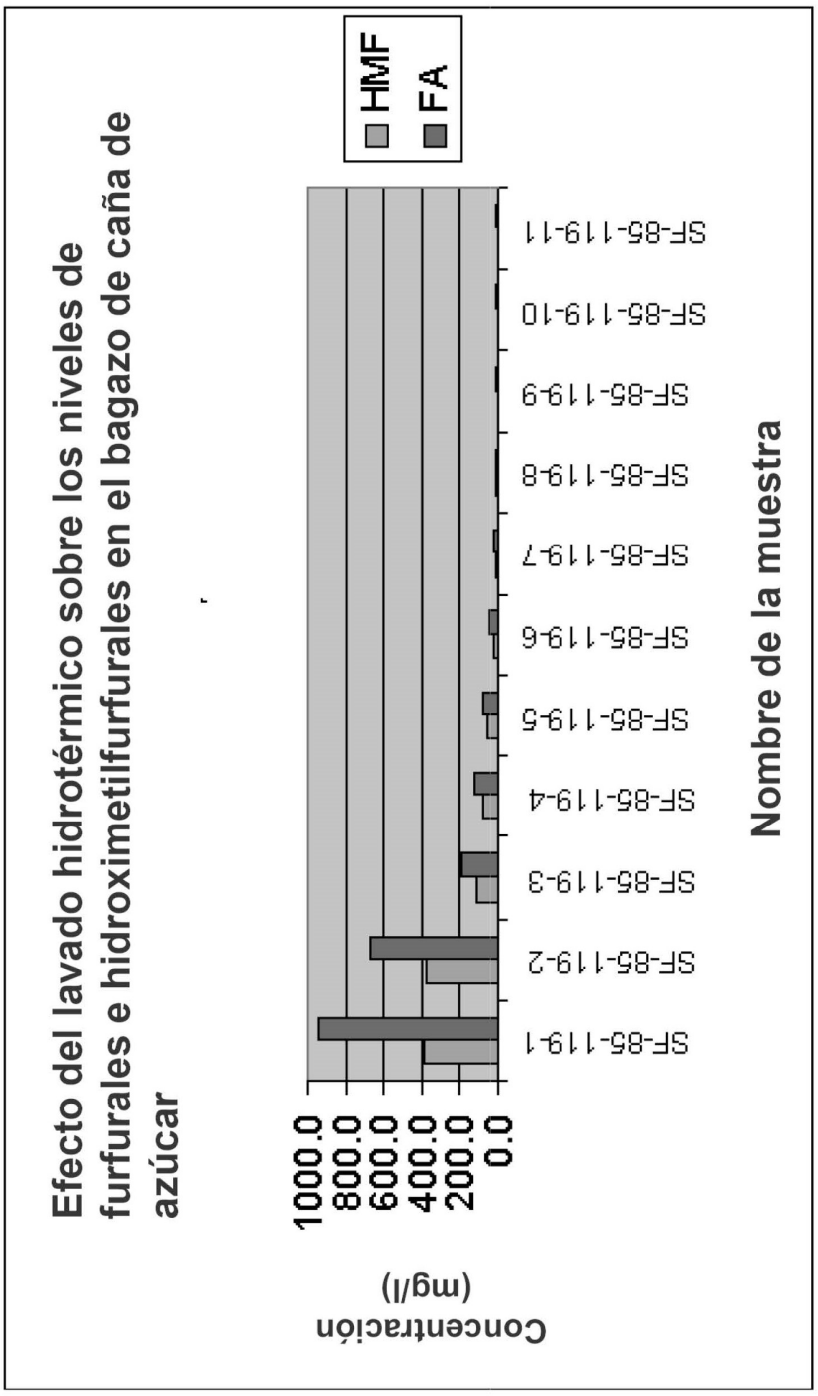

Efecto del lavado hidrotérmico sobre los niveles de furfurales e hidroximetilfurfurales en el bagazo de caña de azúcar
HMF
FA
Concentración (mg/l)
1000.0
800.0
600.0
400.0
200.0
0.0
SF-85-119-1
SF-85-119-2
SF-85-119-3
SF-85-119-4
SF-85-119-5
SF-85-119-6
SF-85-119-7
SF-85-119-8
SF-85-119-9
SF-85-119-10
SF-85-119-11
Nombre de la muestra

Figura 12

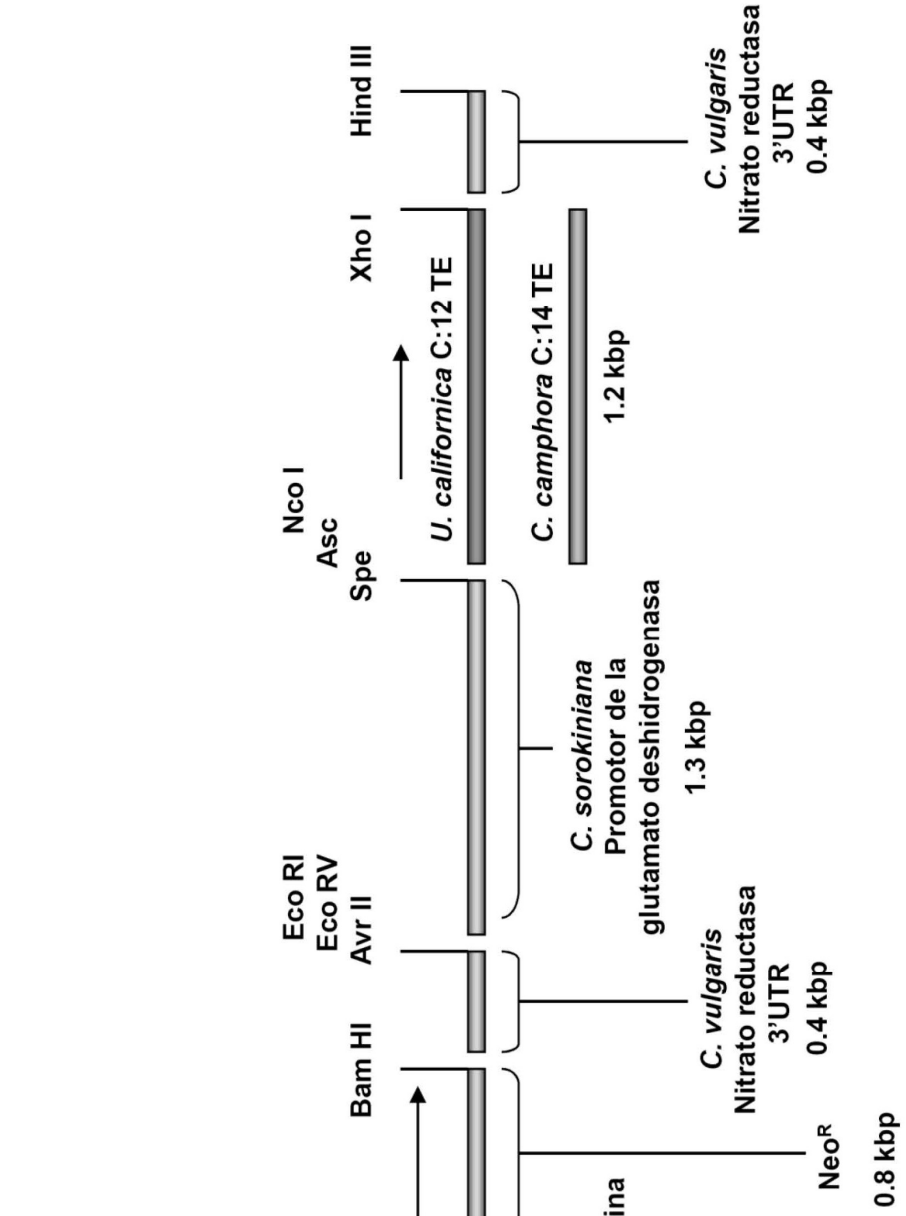
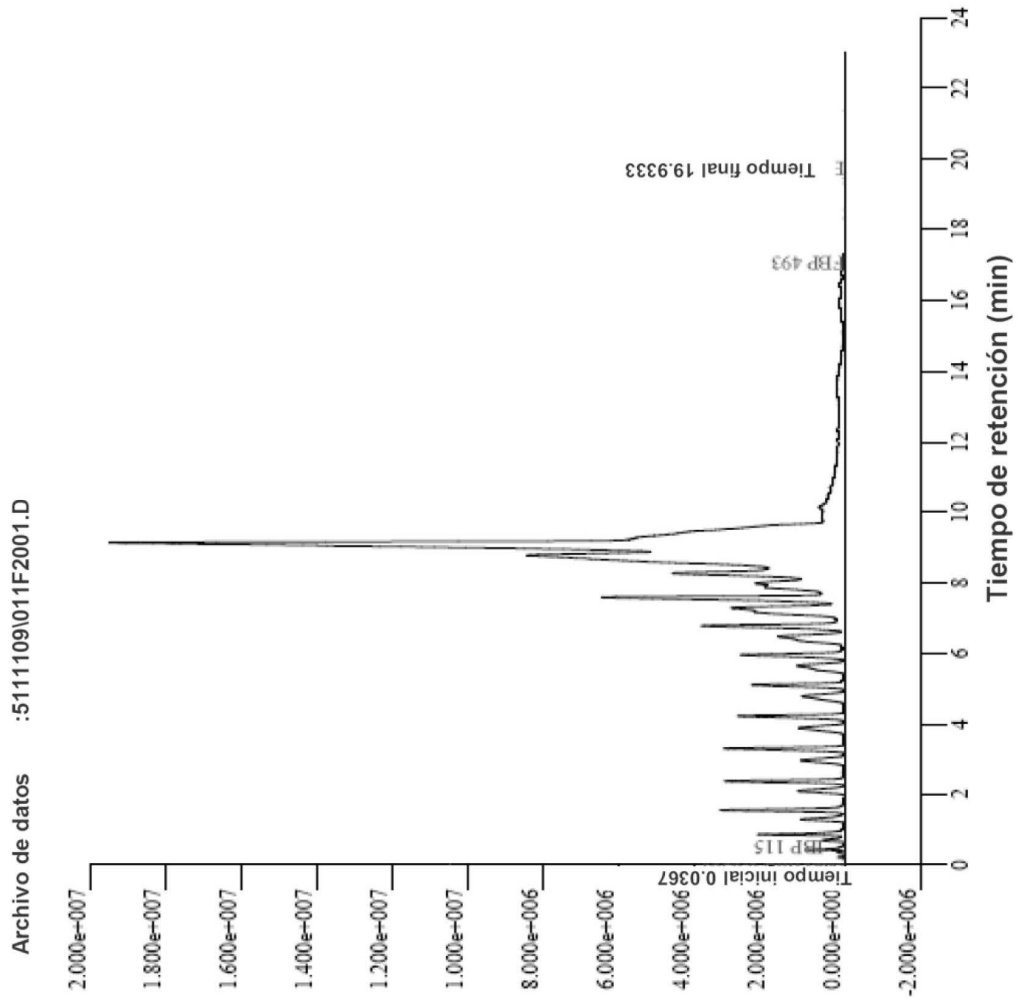

Kpn I
Xba I
Mfe
Bam HI
Eco RI
Eco RV
Avr II
Nco I
Asc
Spe
Xho I
Hind III
C. reinhardtii
Promotor de β tubulina
0.3 kbp
C. vulgaris
Nitrato reductasa
3'UTR
0.4 kbp
NeoR
0.8 kbp
C. sorokiniana
Promotor de la
glutamato deshidrogenasa
1.3 kbp
U. californica C:12 TE
C. camphora C:14 TE
1.2 kbp
C. vulgaris
Nitrato reductasa
3'UTR
0.4 kbp

Figura 13


Archivo de datos :5111109\011F2001.D
Tiempo final 19.9333
Tiempo inicial 0.0367
Tiempo de retención (min)