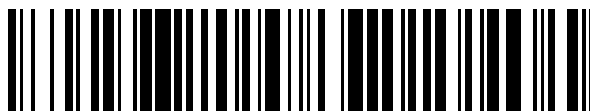


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 583 639**

51 Int. Cl.:

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 09829851 (6)**

97 Fecha y número de publicación de la concesión europea: **20.04.2016 EP 2370555**

54 Título: **Producción de aceites específicos en microorganismos heterótrofos**

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:
21.09.2016

73 Titular/es:

TERRAVIA HOLDINGS, INC. (100.0%)
225 Gateway Boulevard
South San Francisco, CA 94080, US

72 Inventor/es:

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

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 583 639 T3

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

DESCRIPCIÓN

Producción de aceites específicos en microorganismos heterótrofos

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

5 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 SECUENCIAS

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

CAMPO DE LA INVENCIÓN

15 La presente invención se relaciona con 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

20 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.

25 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.

30 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

35 La invención, según se reivindica, proporciona células del género *Prototheca* que contienen un gen exógeno y en donde dicho gen exógeno está en unión operativa con un promotor y codifica una sacarosa invertasa. 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 sec. con núm. de ident.: 11-19. En la invención, según se reivindica, 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 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 sec. con núm. de ident.: 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 sec. con núm. de ident.: 91-102. En otras realizaciones, la célula tiene una secuencia de ARNr 23S con al menos 70, 75, 80, 85 o 95% de identidad de nucleótidos con una o más de las sec. con núm. de ident.: 11-19. En otras realizaciones, el gen exógeno se integra a un cromosoma de la célula. La invención también proporciona el uso de células del género *Prototheca* que comprenden un gen exógeno, donde dicho gen exógeno está en unión operativa con un promotor y codifica una sacarosa invertasa para producir un lípido, opcionalmente en donde el lípido se utiliza además para la producción de un aceite, combustible o producto oleoquímico.

55 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 y 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 sec. con núm. de ident.: 11-19. En otras realizaciones, el gen exógeno acil graso-ACP tioesterasa se integra al cromosoma de la célula. También se divulgan 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 y 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 de esta divulgación la sacarosa invertasa tiene al menos 50, 60, 70, 80, o 90% de identidad de aminoácidos con una o más de las sec. con núm. de ident.: 3, 20-29 y 90.

Se divulgan 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 miligramos 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 de la divulgación, 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 divulgados también incluyen procesar los aceites antes mencionados realizando una o más de las reacciones químicas de la lista consistente de: transesterificación, hidrogenación, hidrocrqueo, desoxigenación, isomerización, interesterificación, hidroxilación, hidrólisis para producir ácidos grasos libres, y saponificación. Se divulgan 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 de la divulgación, 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 25°C. 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.

Se divulga una composición que comprende (a) un polisacárido que comprende uno o más monosacáridos de la lista que consiste en 20-30 por ciento molar de galactosa; 55-65 por ciento molar de glucosa; y 5-15 por ciento 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 sec. con núm. de ident.: 11-19; y (d) un gen exógeno. En algunas realizaciones de la divulgación 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.

Se divulgan 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 transcrito 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 de la divulgación, el ácido nucleico recombinante contiene un segmento de 50 nucleótidos o más de una de las sec. con núm. de ident.: 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, u 80% de los codones de la secuencia codificante. En algunos vectores, la secuencia codificante comprende una secuencia dirigida al plastidio, dentro del 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 sec. con núm. de ident.: 127-133 y es capaz de dirigir una proteína al plastidio de una célula del género *Prototheca*. Otros vectores adicionales de la divulgación comprenden secuencias endógenas de ácidos nucleicos al genoma nuclear de una célula del género
- 5 *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.
- 10 Se divulga un ácido nucleico recombinante que tiene al menos 80, 90, 95 o 98% de identidad de nucleótidos con una o ambas de las sec. con núm. de ident.: 134-135 y un ácido nucleico recombinante que codifica una proteína que tiene al menos 80, 90, 95 o 98% de identidad de aminoácidos con una o ambas de las sec. con núm. de ident.: 136-137.
- Se divulgan 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
- 15 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 de la divulgación, 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.
- 20
- 25
- 30 Otros métodos divulgados 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 de la divulgación, 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
- 35
- 40 *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 divulgado 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 sec. con núm. de ident.: 30 en presencia de sacarosa como fuente de carbono.
- Se divulgan 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 selecciona 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 realizar 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
- 45
- 50
- 55
- 60

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 mencionados anteriormente, opcionalmente una reacción de interesterificación o una reacción de transesterificación. Otros métodos incluyen realizar una reacción de hidroxilación sobre los aceites mencionados anteriormente, en donde, opcionalmente, una reacción de hidroxilación es seguida por la reacción de condensación.

BREVE DESCRIPCIÓN DE LAS FIGURAS

10 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.

15 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 (*ylnv*)). Se indican los sitios de restricción relevantes del clonaje y las flechas indican la dirección de la transcripción.

20 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.

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

La Figura 7d 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*.

30 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.

35 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.

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

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

45 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 alimentaria 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. Se divulgan también 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.

5 Esta descripción detallada de la invención se divide en secciones para conveniencia del lector. La Sección 1 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 divulgados. 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

15 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 invención. Las siguientes referencias proporcionan al experto una definición general de muchos de los términos usados en esta invención: 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.

20 "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.

25 "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 la porción de 4-fosfopantoteína de la coenzima A.

30 "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 y de la divulgación, "al menos 4% de C8-C14" significa que al menos el 4% de los ácidos grasos totales en la célula o en la composición de los glicerolípidos extraídos, tiene una longitud de cadena que incluye 8, 10, 12 o 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.

40 "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.

45 "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.

50 "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.

- 5 “Co-cultivo”, y variantes de este 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.
- 10 “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”).
- 15 “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 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 específicas 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 crudo geológico.
- 20 “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.
- 25 “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 expulsora) 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.
- 30 “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.
- 35 “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.
- 40 “Proporcionada exógenamente”, se refiere a una molécula proporcionada al medio de cultivo de un *cultivo* celular.
- 45 “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.
- 50 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 en un medio de cultivo, que se puede utilizar por un microorganismo cultivado allí.

“Homogenato”, es la biomasa que se ha roto físicamente.

5 “Hidrocarburo”, es una molécula que contiene sólo átomos de hidrógeno y carbono, donde los átomos de carbono están 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 (CH₄), que es un constituyente del gas natural, a una muy fuerte y muy 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.

15 “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 CH₄ (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.

20 “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.

25 “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”.

30 “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.

35 “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.

40 “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.

45 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”.

50 “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.
- 5 “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.
- 10 “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*.
- 15 “Microorganismo” y “microbio”, son organismos microscópicos unicelulares.
- 20 “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.
- 25 “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.
- 30 “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. Un 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.
- 35 “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 ambos recombinantes para los propósitos de esta 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 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.
- 40 45
- 50 “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.
- 55 “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 derivado 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.

5 “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

10 La presente invención se relaciona, en general, con el cultivo de cepas de *Prototheca*, en particular con cepas 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.

15 Especies y cepas de *Prototheca*

La *Prototheca* es un microorganismo notable para su uso en la producción de lípidos, debido a que 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. 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 cepa 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 e 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.

40 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 una 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 especies 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 sec. con núm. de ident.: 11-19.

55 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. Appl. 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 de valor positivo del 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 iniciales 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 incrementa la puntuación de alineamiento acumulativa. Para las secuencias nucleotídicas, las puntuaciones 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 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 TBLASTN (usando secuencias de proteínas para secuencias 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* 90:5873-5877 (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. Biorreactor

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 divulgación 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). 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, también se pueden manipular más fácilmente usando un biorreactor.

5 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.

10 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.

15 Los puertos de los biorreactores se pueden usar para introducir, o extraer, gases, sólidos, semisólidos y líquidos, de la cámara del biorreactor 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 biorreactor 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 comprometedoramente 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.

20 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/CO₂, 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.

25 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. Medio

30 El medio de cultivo de microalgas contiene, típicamente, componentes tales como 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, amoníaco (puro o en forma de sal, como (NH₄)₂SO₄ y NH₄OH), 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 ZnCl₂, H₃BO₃, CoCl₂·6H₂O, CuCl₂·2H₂O, MnCl₂·4H₂O y (NH₄)₆Mo₇O₂₄·4H₂O.

35 De acuerdo con los métodos de la presente divulgación 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.

40 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.

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 NaNO₃, 0.17 mM de CaCl₂·2H₂O, 0.3 mM de MgSO₄·7H₂O, 0.43 mM, 1.29 mM de KH₂PO₄, 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 hidrocloreuro 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, ramosa, sacarosa, y/o xilosa, la selección de materias primas que contengan dichos compuestos es un aspecto importante de los métodos divulgados. Las materias primas adecuadas, útiles de acuerdo con los métodos divulgados 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, tal 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 µM, al menos aproximadamente 100 µM, al menos aproximadamente 500 µM, al menos aproximadamente 5 mM, 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 divulgación incluyen, celulosa (en forma despolimerizada), glicerol, sacarosa, y sorgo, cada uno de estos se discute abajo con más detalle.

De acuerdo con la presente divulgación, los microorganismos se pueden cultivar usando biomasa celulósica despolimerizada como una materia prima. La biomasa celulósica (por ejemplo, rastrojo como rastrojo de maíz) es económica 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 municipales, 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. Se divulgan 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 manera predominante, 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 divulgación 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 a lo largo del tiempo mientras crecen y acumulan lípidos. En la producción de etanol celulósico, las corrientes de azúcares celulósicos no se usan a estos intervalos de concentración ni cerca de estos. 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, 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 divulgadas, 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 divulgación, el proceso de eliminación de los subproductos de carbohidratos se realiza mediante tratamiento hidrotérmico de los materiales celulósicos explotados. Además, se divulgan métodos en los cuales las cepas capaces de tolerar compuestos tales como los furfurales o 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 divulgación, 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 también con la neutralización de los materiales explotados para el proceso de sacarificación enzimática posterior. Se divulgan 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 desionización 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 resuspensión repetida en agua desionizada y centrifugación. Esta etapa de lavado resulta en una corriente celulósica cuya conductividad se encuentra entre 100-300 $\mu\text{S}/\text{cm}$ 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 agentes 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.

La fuente de carbono puede ser, como alternativa, 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 y al menos una de las otras fuentes de carbono. En algunos casos, todo el glicerol y al menos una de las otras fuentes fijas de 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, se proporcionan métodos para alimentar las microalgas con azúcares alternativos, tales como la sacarosa en varias combinaciones con glicerol.

En otra realización de la divulgación 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 pueden proporcionar 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 divulgados, 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. Producción de aceite

Para la producción de aceite de acuerdo con los métodos de la divulgación, 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 de crecimiento, 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 divulgada 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 divulgación 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 divulgación, 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 divulgados, 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 divulgados, 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 divulgada 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, cacahuete, 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 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 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 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 tioesterasa 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 divulgados en la presente. Se divulgan las células recombinantes que contienen genes exógenos integrados a los cromosomas nucleares.

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.

- 5 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.

10 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

15 La presente invención divulga métodos y materiales para modificar genéticamente células de *Prototheca* y células huésped recombinantes útiles en la divulgación 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

20 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 filamentos finos de carburo de silicio. Otro método que se puede usar comprende la formación de protoplastos y el uso de CaCl₂ 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 cotransformación 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).

30 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

35 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 mecánicas 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.

40 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.

50 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 de orientación 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 región 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 K_m , 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. Vectores y componentes de vectores

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 divulgación 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 de control que regulan la expresión de genes o dirigen el producto del gen a una localización particular en la célula recombinante. Para ayudar al lector, esta subsección se divide en subsecciones. La Subsección A describe las secuencias de 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. Como alternativa, 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), p.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* (sec. con núm. de ident.: 69). Opcionalmente, se usan al menos 10, 20, 30, 40, 50, o 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 HUP1 de la *Chlorella* (sec. con núm. de ident.: 1) y el promotor de la nitrato reductasa *Chlorella ellipsoidea* (sec. con núm. de ident.: 2). Los promotores del virus de la *Chlorella* también se pueden usar para expresar genes en *Prototheca*, tales como sec. con núm. de ident.: 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 divulgados. 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 sec. con núm. de ident.: 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 inactivo 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 principalmente 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 divulgación 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 divulgación 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 divulgados en la presente 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 compartimento correcto en la célula huésped. Las bibliotecas de ADNc se realizaron por medio del uso de células de *Prototheca moriformis* y *Chlorella protothecoides* 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 sec. con núm. de ident.: 127-133.

En otra realización de la presente divulgación, 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 señal 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-traducciona 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 divulgación, 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), p. 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. Como alternativa, 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 bleomicina (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, un gen exógeno, que codifica una sacarosa invertasa y opcionalmente un 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).

También se divulga una modalidad en la que los dos o más genes exógenos (además de cualquier marcador de selección) son: una acil graso ACP tioesterasa 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 tioesterasa y una acil graso-CoA reductasa para generar aldehídos. En una realización, el vector proporciona la combinación de una acil graso ACP tioesterasa, una acil graso-CoA reductasa y una aldehído graso decarboxilasa 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 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 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 divulgación 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 ADNc 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)		Ser	AGT	16 (0.03)
	GGA	56 (0.07)			AGC	123 (0.22)
	GGT	76 (0.10)			TCG	152 (0.28)
	GGC	559 (0.71)			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)		Val	GTG	308 (0.50)
	AAA	7 (0.02)			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)		Trp	TGG	107 (1.00)
	CTA	20 (0.03)				
	CTT	45 (0.06)		Tyr	TAT	10 (0.05)
	CTC	190 (0.26)			TAC	180 (0.95)
Met	ATG	191 (1.00)		Terminación	TGA/TAG/TAA	

5 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

10 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.

20 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 sec. con núm. de ident: 3) con una señal de secreción (tal como sec. con núm. de ident: 4 (de levadura), sec. con núm. de ident: 5 (de plantas superiores), sec. con núm. de ident: 6 (señal consenso de secreción de eucariotas), y sec. con

núm. de ident: 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 divulgada 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.

5 *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 divulgados 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 en Genbank	sec. con núm. de ident.:
Invertasa	<i>Chicorium intybus</i>	Y11124	sec. con núm. de ident.:20
Invertasa	<i>Schizosaccharomyces pombe</i>	AB011433	sec. con núm. de ident.:21
beta-fructofuranosidasa (invertasa)	<i>Pichia anomala</i>	X80640	sec. con núm. de ident.:22
Invertasa	<i>Debaryomyces occidentalis</i>	X17604	sec. con núm. de ident.:23
Invertasa	<i>Oryza sativa</i>	AF019113	sec. con núm. de ident.:24

Invertasa	<i>cepa Allium</i>	AJ006067	sec. con núm. de ident.:25
Invertasa	<i>Beta vulgaris subsp. Vulgaris</i>	AJ278531	sec. con núm. de ident.:26
beta-fructofuranosidasa (invertasa)	<i>Bifidobacterium breve UCC2003</i>	AAT28190	sec. con núm. de ident.:27
Invertasa	<i>Saccharomyces cerevisiae</i>	NP_012104	sec. con núm. de ident.:8 (nucleótido)
			sec. con núm. de ident.:28 (aminoácido)
Invertasa A	<i>Zymomonas mobilis</i>	AAO38865	sec. con núm. de ident.: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 microbiano 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 divulgados en la presente 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 relevantes se pueden aplicar de manera que las materias primas, tales como las celulósicas son utilizables por un microbio huésped modificado 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 divulgación también proporciona *Prototheca* recombinante que se modificó para alterar las propiedades y/o proporciones de los lípidos producidos. La vía puede además, o como alternativa, 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 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 clave que controlan puntos de ramificación en el metabolismo de la síntesis de ácidos grasos se regularon ascendente o descendientemente 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 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 ascendientemente 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 esteroil, (SREBP), tales como SREBP-1a y SREBP-1c (para ejemplos ver los números de acceso en Genbank NP_035610 y Q9WTN3).

Se divulgan también 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 descarboxilasa (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 coexpresa 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, sin tomar en cuenta si ellas se coexpresan naturalmente o no en un tejido u organismo particular. Por lo tanto, la presente divulgación contempla tanto pares de estas enzimas que se coexpresan 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 aún otras realizaciones de la divulgación, 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 de la divulgación, los microbios que se divulgan 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 coexpresada. 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 entonces 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 producir 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 ACP y la tioesterasa se pueden coexpresar naturalmente en un determinado tejido u organismo del cual se derivan). Por lo tanto, en varias realizaciones, las células de *Prototheca* recombinantes divulgadas 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 sustrato 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 sustrato que tiene menos, es decir, 12, en lugar de más, es decir, 18 átomos de carbono.

En ejemplos no limitantes sino ilustrativos, la presente divulgación 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 (número de acceso al GenBank Q39473, la secuencia de aminoácidos está en la sec. con número de identificación: 59, la secuencia de aminoácidos sin secuencia dirigida a plastidios (STP) está en la sec. con número de identificación: 139, y la secuencia de ADNc con el codón optimizado que se basa en la Tabla 1 está en la sec. con número de identificación: 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 (número de acceso al GenBank AAC49269, la secuencia de aminoácidos está en la sec. con número de identificación: 61, la secuencia de aminoácidos sin STP está en la sec. con número de identificación: 138, y la secuencia de ADNc con el codón optimizado que se basa en la Tabla 1 se encuentra en la sec. con número de identificación: 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 (número de acceso al GenBank Q41635, la secuencia de aminoácidos se incluye en la sec. con número de identificación: 63, la secuencia de aminoácidos sin STP está en la sec. con número de identificación: 139, y la secuencia de ADNc con el codón optimizado que se basa en la Tabla 1 se encuentra en la sec. con número de identificación: 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 divulgación incluyen, sin limitarse a, las enumeradas en la Tabla 3.

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

<i>Umbellularia californica</i> acil graso-ACP tioesterasa (GenBank #AAC49001)
<i>Cinnamomum camphora</i> acil graso-ACP tioesterasa (GenBank #Q39473)
<i>Umbellularia californica</i> acil graso-ACP tioesterasa (GenBank #Q41635)
<i>Myristica fragrans</i> acil graso-ACP tioesterasa (GenBank #AAB71729)
<i>Myristica fragrans</i> acil graso-ACP tioesterasa (GenBank #AAB71730)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #ABD83939)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #AAD42220)
<i>Populus tomentosa</i> acil graso-ACP tioesterasa (GenBank #ABC47311)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #NP_172327)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #CAA85387)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #CAA85388)
<i>Gossypium hirsutum</i> acil graso-ACP tioesterasa (GenBank #Q9SQI3)
<i>Cuphea lanceolata</i> acil graso-ACP tioesterasa (GenBank #CAA54060)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #AAC72882)
<i>Cuphea calophylla</i> subsp. <i>mesostemon</i> acil graso-ACP tioesterasa (GenBank #ABB71581)
<i>Cuphea lanceolata</i> acil graso-ACP tioesterasa (GenBank #CAC19933)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #AAL15645)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #Q39513)
<i>Gossypium hirsutum</i> acil graso-ACP tioesterasa (GenBank #AAD01982)
<i>Vitis vinifera</i> acil graso-ACP tioesterasa (GenBank #CAN81819)
<i>Garcinia mangostana</i> acil graso-ACP tioesterasa (GenBank #AAB51525)
<i>Brassica juncea</i> acil graso-ACP tioesterasa (GenBank #AB118986)
<i>Madhuca longifolia</i> acil graso-ACP tioesterasa (GenBank #AAX51637)
<i>Brassica napus</i> acil graso-ACP tioesterasa (GenBank #ABH11710)
<i>Oryza sativa</i> (cultivares grupo indica) acil graso-ACP tioesterasa (GenBank #EAY86877)
<i>Oryza sativa</i> (cultivares grupo japónica) acil graso-ACP tioesterasa (GenBank #NP_001068400)
<i>Oryza sativa</i> (cultivares grupo indica) acil graso-ACP tioesterasa (GenBank #EAY99617)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #AAC49269)
<i>Ulmus Americana</i> acil graso-ACP tioesterasa (GenBank #AAB71731)
<i>Cuphea lanceolata</i> acil graso-ACP tioesterasa (GenBank #CAB60830)
<i>Cuphea palustris</i> acil graso-ACP tioesterasa (GenBank #AAC49180)
<i>Iris germanica</i> acil graso-ACP tioesterasa (GenBank #AAG43858)
<i>Cuphea palustris</i> acil graso-ACP tioesterasa (GenBank #AAC49179)
<i>Myristica fragrans</i> acil graso-ACP tioesterasa (GenBank #AAB71729)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #U39834)
<i>Umbellularia californica</i> acil graso-ACP tioesterasa (GenBank #M94159)
<i>Cinnamomum camphora</i> acil graso-ACP tioesterasa (GenBank #U31813)

Los ejemplos debajo describen el éxito de la orientación y la expresión heteróloga de la acil graso ACP tioesterasas de *Cuphea hookeriana*, *Umbellularia californica*, *Cinnamomum 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

5 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 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 aminoácidos 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 sec. con núm. de ident.: 134-137.

10 Cuando las secuencias de aminoácidos de estos dos ADNc se analizaron por BLAST contra la base de datos NCBI, las dos secuencias más homólogas 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 (NXHX₃₆C) 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 divulgación incluyen, sin limitarse a, los enumerados en la Tabla 4.

25 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 divulgación 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
--

30 Las aldehído graso decarbonilasas adecuadas para su uso con los microbios y métodos de la divulgación 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

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

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 adicionalmente 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 divulgación 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.

Al seleccionar la combinación deseada de genes exógenos a expresarse, se puede adaptar el producto generado por el microbio, el cual se puede extraer entonces 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 coexpresada de manera natural o una proteína portadora de acilo que tiene, de otra manera, 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 decarboxilasa. El microbio, bajo las condiciones de cultivo que se describen en la presente, sintetiza un ácido graso unido a una ACP 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 graso-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 producir 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 decarboxilasa 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 cuáles 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 divulgación 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 de la divulgación, 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 divulgación 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 graso-CoA reductasa, una aldehído graso reductasa, una acil graso-CoA/aldehído reductasa, una aldehído graso decarboxilasa, 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 de la divulgación, 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 divulgación también proporciona una célula de *Prototheca* recombinante que contiene dos genes exógenos, en donde el primer gen exógeno codifica una acil graso-ACP tioesterasa y el segundo gen exógeno codifica una proteína que se selecciona del grupo que consiste de una acil graso-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 divulgados 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 divulgan 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 divulgados. 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 divulgados. 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 divulgados 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 divulgados.

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 carbonos de una ACP, y la reductasa codificada por el segundo gen exógeno cataliza la reducción de un acil graso-CoA de 10 a 14 carbonos 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 coexpresa naturalmente con la acil graso- ACP tioesterasa.

En algunas realizaciones, el segundo gen exógeno codifica para una acil graso-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 a 18 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 graso-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 las células sintetizan un ácido graso vinculado a una proteína portadora de acilo (ACP), la acil graso-ACP tioesterasa cataliza la escisión del ácido graso de la ACP para producir, a través de la transformación posterior, una acil graso-CoA y la acil graso-CoA/aldehído reductasa cataliza la reducción de la acil-CoA a un alcohol.

La presente divulgación 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 que codifica una acil graso-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 graso-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 divulgación 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 esteroil 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 graso-ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, una aldehído graso decarboxilasa, 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 divulgación, 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 de la culminación 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. Como alternativa, 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 contienen 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 divulgada, 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. Como alternativa, 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 divulgación, 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 estos métodos 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 divulgación, 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 mM 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 ruptura celular.

En otra realización de la presente divulgación, 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 presente divulgación, 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, p. 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 divulgación, 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 divulgación, 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 célula que contiene la mezcla 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. Como alternativa, 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 divulgación, la etapa de lisis de un microorganismo se realiza mediante la aplicación de un choque osmótico.

5 En otra realización de la presente divulgación, 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, 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 microalga 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.

15 En otra realización de la presente divulgación, la etapa de lisis de un microorganismo comprende la autólisis. En esta realización, un microorganismo de acuerdo con 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 lítico codifica para una enzima o divulgación 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.

25 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 CO₂ 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 protothecoides*, en el 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).

40 De esta manera, los lípidos, derivados de lípidos e hidrocarburos generados por los microorganismos de la presente invención o divulgación 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.

45 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. Como alternativa, 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*.

50 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 de manera adicional 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 automultiplicativa. 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 hidrocarburos más pequeños.

5 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
10 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
15 Bioengineering, Vol. XXIV: 193-205 (1982)), una publicación sobre hidrocrqueo 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 divulgación, son útiles en una variedad de
20 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
25 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 los materiales de partida biológicos renovables pueden sustituir a los materiales de partida derivados 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
30 partir de materiales biológicos. La presente divulgación 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
35 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 combustible, 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
40 con la especificación apropiada de la ASTM, se puede definir como combustible diesel (ASTM D975), combustible de reactor (ASTM D1655), o como biodiesel 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.

45 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.
50 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
55 mezcla con el combustible diesel convencional (diesel), el biodiesel puede estar presente en aproximadamente de 0.1% 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.

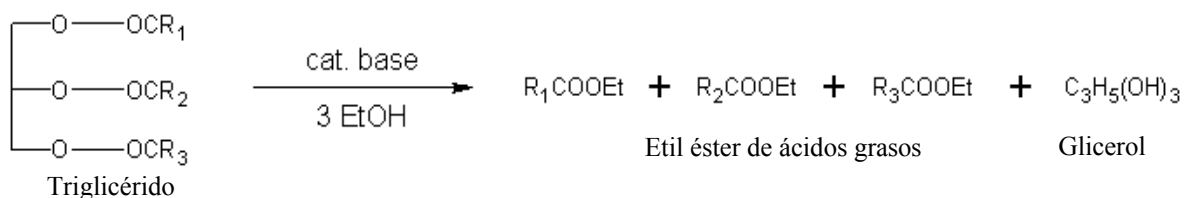
El biodiesel también puede usarse como combustible de calefacción en calderas domésticas y comerciales. Las calderas de aceite que existen pueden contener partes de caucho, y pueden requerir conversión para funcionar con biodiesel. El proceso de conversión es, por lo general, relativamente sencillo, 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, también 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 divulgadas en la presente 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:



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.

<p>Lipasa de <i>Aspergillus niger</i> ABG73614, lipasa B de <i>Candida antarctica</i> (novozym-435) CAA83122, lipase de <i>Candida cylindracea</i> AAR24090, lipasa de <i>Candida lipolytica</i> (Lipasa L; Amano Pharmaceutical Co., Ltd.), lipasa de <i>Candida rugosa</i> (por ejemplo, Lipasa-OF; Meito Sangyo Co., Ltd.), lipasa de <i>Mucor miehei</i> (Lipozyme IM 20), lipasa de <i>Pseudomonas fluorescens</i> AAA25882, lipasa de <i>Rhizopus japonicas</i> (Lilipase A-10FG) Q7M4U7_1, lipasa de <i>Rhizomucor miehei</i> B34959, lipasa de <i>Rhizopus oryzae</i> (Lipasa F) AAF32408,</p>
--

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 el 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 4 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 recombinantes 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 se 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 (DL₅₀) 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 de Estados Unidos 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 proveniente 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 como se divulgan en la presente pueden transesterificarse para generar biodiesel con un tiempo de maceración en frío de menos de 120 segundos, menos de 100 segundos 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 divulgación 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 divulgación pueden servir como materia prima para producir el diesel renovable. Por lo tanto, se proporciona un método para producir diesel renovable. El diesel renovable puede producirse mediante al menos tres procesos: procesamiento hidrotérmico (hidrotratamiento); hidroprocesamiento y licuefacción indirecta. Estos procesos generan destilados que no son ésteres. 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 hidrotatar 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, hidrocrqueo, 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 300° a 660°F, 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 divulgado, 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 300° a 700° 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 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 triglicéridos. 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 divulgadas en la presente, 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) divulgados 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 242.1°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) divulgados 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) divulgados 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) divulgados 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 divulgación 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 aún 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.

5 La presente divulgación también proporciona métodos para producir combustible de avión. 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 de 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 Kerosene (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.

15 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 principio activo, es un ejemplo. Los inhibidores de la corrosión, por ejemplo, el DCI-4A, se usa para combustibles civiles y militares y el DCI-6A se usa para combustibles militares. Los agentes FSII incluyen, por ejemplo, el Di-EGME.

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

25 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 divulgado 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 en 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.

30 En una realización, el método para producir combustible de reactor comprende (a) cultivar un microorganismo que contiene lípido usando los métodos divulgados 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 C₂-C₅.

35 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 AmberlystTM-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.

40 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 hidrocrqueo, 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 C₂-C₅ se describen en las patentes de los Estados Unidos núms. 6,538,169 y 7,288,685.

5 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 10 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.

15 En una realización, el craqueo de la composición lipídica de la presente divulgación, se lleva a cabo en la sección del ascensor o, como alternativa, 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 149°C a aproximadamente 316°C (300°F a 600°F). 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 aproximadamente 2 segundos o menos.

20 El catalizador mezclado y los vapores de la composición lipídica reaccionada se descargan entonces 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 25 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.

35 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.

40 La temperatura de la corriente de craqueo en la salida del ascensor será de aproximadamente 510°C a 621°C (950°F a 1150°F). Sin embargo, las temperaturas de salida del ascensor por encima de 566°C (1050°F) producen más gas seco y más olefinas. Mientras que las temperaturas de salida del ascensor por debajo de 566°C (1050°F) producen menos etileno y propileno. Por consiguiente, se prefiere para llevar a cabo el proceso de FCC a una temperatura preferida de aproximadamente 566°C 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.

50 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 (1292°F) a aproximadamente 760°C (1400°F), 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. 55

60 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 hidredesoxigenación (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.

5 Generalmente en la HDO/etapa de HDS, 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 cocorriente, ya sea como una sola fase o una materia prima de la fuente en dos fases. Después de la HDO/etapa de MDS, 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 cocorriente.

10 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 corrientes 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 C₂-C₅ se describen en la patente de los Estados Unidos núm. 7,232,935.

15 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.

20 En la primera etapa, 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 cocorriente 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 cocorriente. 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/Al₂O₃ y CoMo/Al₂O₃.

30 Antes de la etapa de HDO, la composición lipídica o los lípidos de 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.

35 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.

40 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.

45 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 cocorriente o contracorriente.

50 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 de 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 Al₂O₃ y SiO₂. Los catalizadores de isomerización típicos son, por ejemplo, Pt/SAPO-11/Al₂O₃, Pt/ZSM-22/Al₂O₃, Pt/ZSM-23/Al₂O₃ y Pt/SAPO-11/SiO₂. 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/M³ y 840K/M³. En aún 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, se divulgan una variedad de métodos en los cuales se lleva a cabo la modificación química de los lípidos de las microalgas para producir 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 divulgados 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 divulgados, 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 divulgada, 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 divulgados en la presente, mediante el fraccionamiento del aceite en glicerina y ácidos grasos. Como se mencionó anteriormente, los ácidos grasos pueden procesarse después adicionalmente 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 divulgadas, 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.

5 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 entonces 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.

10 Las reacciones de fraccionamiento enzimático generalmente tienen lugar en la delimitación de las fases entre la fase orgánica y la fase 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.

15 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 libres 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 nuevo o la grasa para someterse al fraccionamiento. De esta manera, los glicéridos reciclados se someten entonces 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.

20 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. Como alternativa, la densidad de la fase acuosa del glicerol se puede tomar como una medida del grado de fraccionamiento.

25 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.

30 En una realización, los aceites producidos de acuerdo con los métodos divulgados en la presente también pueden someterse a saponificación como un método de hidrólisis. Los aceites animales y vegetales se hacen típicamente de triacilglicerolos (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. Como alternativa, 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.

35 En algunos métodos, la primera etapa de modificación química puede ser el hidrogenamiento 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. 5,475,160 (Hidrogenación de triglicéridos); 5,091,116 (Desoxigenación, hidrogenación y eliminación de gas); 6,391,815 (Hidrogenación); y 5,888,947 (Isomerización).

Los aceites triglicéridos se pueden desoxigenar 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 otra realización, la reacción de hidrogenación en el método divulgado 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 o cromo) en cantidades tales que aproximadamente del 1 al 2% del peso se mantiene en el catalizador en 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 de este tipo incluyen, pero sin limitarse a, carbono, sílice, sílice-alúmina, zirconio, titanía, 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 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 llevar a cabo la reacción de hidrogenación varían en función del tipo de material de partida 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.

5 El hidrógeno usado en la reacción de hidrogenólisis divulgada en la presente 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.

10 En algunas realizaciones de la divulgación, 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 hidrogenólisis 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.

20 En algunas realizaciones, un alcohol de azúcar o azúcar de 5 y/o 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 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 descrito en la reacción de hidrogenólisis puede incluir un soporte de catalizador como el descrito anteriormente para la reacción de hidrogenación.

30 Las condiciones para llevar a cabo la reacción de hidrogenólisis varían en función del tipo de material de partida 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 225°C. 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 un 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 divulgación puede incluir hidrógeno externo, hidrógeno reciclado, hidrógeno generado in situ, y cualquier combinación de los mismos.

40 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.

50 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, heteropolíácidos, 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 autoportado. Tal como se usa en la presente, el término "autoportado" 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.

5 Las condiciones bajo las cuales la reacción de condensación se produce varían en función del tipo de material de partida 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 varía 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 125°C a 450°C, y con la máxima preferencia de 125°C 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.

10 Los alcanos superiores formados por la divulgación 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-dimetilhexano, 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.

15 En algunas realizaciones, los cicloalcanos y los cicloalquenos son no sustituidos. En otras realizaciones, los cicloalcanos y los cicloalquenos son monosustituidos. En aún otras realizaciones, los cicloalcanos y los cicloalquenos son multisustituidos. En las realizaciones que comprenden los cicloalcanos y cicloalquenos sustituidos, el grupo sustituido incluye, sin limitarse a, un alquilo de cadena recta o ramificada que tiene 1 a 12 átomos de carbono, un alqueno 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, ciclohexano, ciclohexeno, metil-ciclopentano, metil-ciclopenteno, etil-ciclopentano, etil-ciclopenteno, etil-ciclohexano, etil-ciclohexeno, isómeros y cualquier combinación de éstos.

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

25 Los alcoholes producidos de acuerdo con la divulgación 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 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.

30 Las cetonas producidas de acuerdo con la divulgación 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 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.

35 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 alimentarios); 5,434,278 (Grasas comestibles para productos alimentarios); 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 divulgación, 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: 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 aproximadamente 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 resto 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 entonces incrementarse hasta al menos aproximadamente 60°C y sagita 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 después 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 reacción puede secarse después 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 glicerolípidos 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, *JAACS* 71(2): 169-174 (1994). Los estolidos 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 oleaginosa 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 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 encuentran 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 K₂HPO₄, 3.1 g/l de NaH₂PO₄, 0.24 g/l de MgSO₄ · 7H₂O, 0.25 g/l de ácido cítrico monohidrato, 0.025 g/l de CaCl₂ 2H₂O, 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
<i>Prototheca stagnora</i>	UTEX 327	13.14
<i>Prototheca moriformis</i>	UTEX 1441	18.02
<i>Prototheca moriformis</i>	UTEX 1435	27.17

Se determinó el genotipo de las muestras de microalgas de las cepas que se enumeran en la Tabla 22 mostrada anteriormente. 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 molienda (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 después 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'-TGTTGAAGAATGAGCCGCGAC-3' (sec. con núm. de ident.: 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' (sec. con núm. de ident.: 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 dH₂O. 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 secuencias 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 (sec. con núm. de ident.: 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	Sec. con núm ident
<i>Prototheca kruegani</i>	UTEX 329	75.2	sec. con núm ident.: 11
<i>Prototheca wickerhamii</i>	UTEX 1440	99	sec. con núm ident.: 12
<i>Prototheca stagnora</i>	UTEX 1442	75.7	sec. con núm. ident.: 13
<i>Prototheca moriformis</i>	UTEX 288	75.4	sec. con núm ident.: 14
<i>Prototheca moriformis</i>	UTEX 1439; 1441; 1435; 1437	100	sec. con núm ident.: 15
<i>Prototheca wickerhamii</i>	UTEX 1533	99.8	sec. con núm ident.: 16
<i>Prototheca moriformis</i>	UTEX 1434	75.9	sec. con núm ident.: 17
<i>Prototheca zopfii</i>	UTEX 1438	75.7	sec. con núm ident.: 18
<i>Prototheca moriformis</i>	UTEX 1436	88.9	sec. con núm ident.: 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	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

1435									
------	--	--	--	--	--	--	--	--	--

Los péptidos de tránsito al plastidio de algas se identificaron a través del análisis de colecciones de ADNc de UTEX 1435 (*Prototheca moriformis*) o de UTEX 250 (*Chlorella protothecoides*) como se describen en el Ejemplo 12 y en el Ejemplo 11 a continuación. 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 plastidios 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	sec. con núm. ident.
	<i>P.moriformis</i> isopentenil difosfato sintasa PTS	sec. con núm. ident.: 127
5	<i>P.moriformis</i> delta 12-desaturasa de ácido graso PTS	sec. con núm. ident.: 128
	<i>P.moriformis</i> estearoil ACP desaturasa PTS	sec. con núm. ident.: 129
	<i>C.protothecoides</i> estearoil ACP desaturasa PTS	sec. con núm. ident.: 130
	<i>Cuphea hookeriana</i> acilo graso-ACP tioesterasa (C8-10)	sec. con núm. ident.: 131
	<i>Umbellularia californica</i> acil graso-ACP tioesterasa (C12)	sec. con núm. ident.: 132
20	<i>Cinnamomum camphora</i> acil graso-ACP tioesterasa (C14)	sec. con núm. ident.: 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 ° 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 posteriormente a 4.0 y se realizó un tratamiento con pectinasa a 50 ° 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 entonces 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.

5 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 un sistema de Ankom de hidrólisis ácida/extracción por solvente como se describió en el Ejemplo 1 anterior.

20 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/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 presente en la materia prima de azúcar 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 Na_2SO_4 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 cultivo de los métodos divulgados 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 divulgación 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*

5 *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.

10 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 se ve en la Figura 8).

20 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 a continuación. 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 celulósicas.

	glucosa 1 (ctrl)	glucosa 2 (ctrl)	bagazo 1	bagazo 2	sorgo 1	sorgo 2	<i>Miscan</i> 1	<i>Miscan</i> 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

30 n.d. denota ninguno detectado

35 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 divulgación 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, 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 μ S/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 IEX para reducir las sales.

Volumen del lecho de resina : celulósicos	pH post-desionización	Conductividad post-desionización (μ S/cm)	Conductividad post-desionización calculada y re-concentración 8x (μ S/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. Como alternativa, 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 μ S/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 μ S/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 centrifuga decantadora Sharples Modelo 660. El permeado líquido se mantuvo en lotes (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 centrifuga, 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 μ S/cm	total de xilosa eliminada (g)	total de glucosa eliminada (g)
1 (material inicial)	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 (NH₄)₂SO₄, 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, amoníaco, 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. Método general para la transformación biolística de *Prototheca*

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 2x10⁶ 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 10⁷ 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 (sec. con núm. de ident.: 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-Ascl. 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:

15 5'TGACCTAGGTGATTAATTAAGCTCGAGGCAGCAGCAGCTCGGATAGTATCG 3' (sec. con núm. de ident.:35)

Inverso:

5' CTACGAGCTCAAGCTTTCCATTTGTGTTT CCATCCCCTACTTCC 3' (sec. con núm. de ident.: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:

20 Directo: 5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (sec. con núm. de ident.:37)

Inverso: 5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (sec. con núm. de ident.: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 Ascl-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 heterofilicamente en medio líquido que contenía 2% de glucosa hasta alcanzar la densidad celular deseada (1×10^7 células/ml a 5×10^7 células/ml). Las células se cosecharon, se lavaron una vez con agua destilada estéril y se resuspendieron a 1×10^8 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 métodos de biología molecular estándar. 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 (sec. con núm. de ident.: 39) (los transformantes 1 y 2) o con la glutamato deshidrogenasa::neo::nit (sec. con num. de ident.: 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 (sec. con num. de ident.: 41) y la β -tubulina::neo::nit con codón optimizado (sec. con num. de ident.: 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 integración 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 nailon. 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' (sec. con núm. de ident.: 43)

Inverso: 5' AGGTCCTCGCCGTCGGGCATG 3' (sec. con núm. de ident.: 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. Transformación de *Prototheca* con sacarosa invertasa secretada heteróloga

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 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), sec. con núm. de ident.: 57, y la construcción con codón optimizado (Crβ-tub::CO-suc2::CvNitRed), sec. con núm. de ident.: 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 heterotróficamente 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 1.25×10^6 células/ml en medio fresco.

En una campana estéril se retiraron 0.4 ml de las células que se prepararon anteriormente 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 a continuación 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 a continuación 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 a continuación 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 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 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 1435 transformante) 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'GGTGGCGGGTCCAGGGTGT 3', sec. con núm. de ident.: 65 y 66, respectivamente) y los transcritos de *suc2* de invertasa (cebador directo 5'CGGCCGGCGGCTCCTTCAAC 3' y cebador inverso 5'GGCGCTCCCGTAGGTCGGGT 3', sec. con num. de ident.: 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 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, 2^{da} 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 las 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 K₂HPO₄ 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 µg 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 aproximadamente 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 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 1435, que creció en glucosa como única fuente de carbono y dos cepas colonial T98 (UTEX 1435 transformada con un transgén de sacarosa invertasa), que crecieron en sacarosa como única

fuelle de carbono, se describen en la Tabla 13 (tipo salvaje UTEX 1435 y clon 8 y el clon 11 de T98 a continuaci3n. El l3pido C:19:0 se us3n como un control de calibraci3n interna.

Tabla 13. Perfil lip3dico de UTEX 1435 tipo salvaje y los clones de UTEX 1435 con el transg3n *suc2*.

Nombre	tipo salvaje (% 3rea - ISTD)	T98 clon 11 (% 3rea - ISTD)	T98 clon 8 (% 3rea - 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			

- 5 El aceite extra3do de *Prototheca moriformis* UTEX 1435 de tipo salvaje (a trav3s de extracci3n con disolventes o mediante el uso de una prensa expulsora (ver m3todos en el Ejemplo 44 anterior) se analiz3 para los carotenoides, la clorofila, los tocoferoles, otros esterole y tocotrienoles. Los resultados se resumen a continuaci3n en la Tabla 14.

Tabla 14. An3lisis de carotenoide, clorofila, tocoferol/esterole y tocotrienol en aceite extra3do de *Prototheca moriformis* (UTEX 1435).

	Aceite prensado (mcg/ml)	Aceite extra3do del solvente (mcg/ml)
cis-Lute3na	0.041	0.042
trans-Lute3na	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 extra3do del solvente (mg/100g)
gamma Tocoferol	0.49	0.49
Campesterol	6.09	6.05
Stigmasterol	47.6	47.8
Beta-sitosterol	11.6	11.5
Otros esterole	445	446
Tocotrienoles		
	Aceite prensado (mg/g)	Aceite extra3do del solvente (mg/g)
alfa Tocotrienol	0.26	0.26
beta Tocotrienol	< 0.01	< 0.01

gamma Tocotrienol	0.10	0.10
delta 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.

15 Ejemplo 4: *Prototheca* recombinante 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 (sec. con núm. de ident.: 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* (sec. con num. de ident.: 70) o la tioesterasa C14 de *Cinnamomum camphora* (sec. con num. de ident.: 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 (sec. con núm. de ident.: 72), seguida de la 3'UTR de la nitrato reductasa de *Chlorella vulgaris* (sec. con num. de ident.: 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 en sec. con núm. de ident.: 70 y sec. con núm. de ident.: 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-PCR cuantitativa 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-PCR cuantitativa de C12 de transformación de la tioesterasa fueron:

U. californica cebadores PCR de tioesterasa C12:

50 Directo: 5' CTGGGCGACGGCTTCGGCAC 3' (sec. con núm. ident.: 74)

Inverso: 5' AAGTCGCGGCGCATGCCGTT 3' (sec. con núm. ident.: 75)

Los cebadores que se usaron en la detección RT-PCR cuantitativa de la transformación de tioesterasa C14 fueron:

Cinnamomum camphora cebadores de PCR tioesterasa C14:

Directo: 5' TACCCCGCCTGGGGCGACAC 3' (sec. con núm. ident.: 76)

Inverso: 5' CTTGCTCAGGCGGGGTGC 3' (sec. con núm. ident.: 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 H₂SO₄ 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 K₂CO₃ (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 Na₂SO₄ (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 C14 *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	11.68	10.71	12.00	10.15	11.03	12.08	9.98	10.47
alfa								
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	0.19	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 anteriormente 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, (sec. con num. de ident.: 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 plastidio, "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 Número	Promotor/5'UTR	sec. dirigida a plastidio	Gen	3'UTR	sec. con núm. ident...
Construcción 1	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 79
Construcción 2	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 80
Construcción 3	<i>C. sorokiniana</i>	<i>P. moriformis</i>	<i>Cuphea</i>	<i>C. vulgaris</i>	sec. con núm.

	glutamato deshidrogenasa	Isopentenil difosfato sintasa	<i>hookeriana</i> C8-10 TE	nitrato reductase	ident.: 81
Construcción 4	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> Isopentenil difosfato sintasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 82
Construcción 5	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 83
Construcción 6	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoide</i> s estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 84
Construcción 7	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 85
Construcción 8	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoide</i> s estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 86
Construcción 9	<i>Chlamydomonas reinhardtii</i> β -tubulina	Native	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 113
Construcción 10	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> Isopentenil difosfato sintasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 114
Construcción 11	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 115
Construcción 12	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C. protothecoide</i> s estearoil ACP desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 116
Construcción 13	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 117
Construcción 14	<i>Chlamydomonas reinhardtii</i> β -tubulina	Native	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 118
Construcción 15	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> Isopentenil difosfato sintasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 119
Construcción 16	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 120
Construcción 17	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C. protothecoide</i> s estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 121
Construcción 18	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 122
Construcción 19	<i>Chlamydomonas reinhardtii</i> β -tubulina	Native	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 123
Construcción 20	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> Isopentenil difosfato sintasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 124
Construcción	<i>Chlamydomonas</i>	<i>P. moriformis</i>	<i>Cinnamomum</i>	<i>C. vulgaris</i>	sec. con núm.

21	<i>reinhardtii</i> β -tubulina	delta 12 graso ácido desaturasa	<i>camphora</i> C14 TE	nitrato reductase	ident.:
Construcción 22	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C.protothecoide</i> s estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 87
Construcción 23	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductase	sec. con núm. ident.: 88

5 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.

10 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 más 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 el 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 aproximadamente 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 reinhardtii*

35 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 huésped *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 huésped 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 más 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.

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 18:3 α	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. <i>Cuphea hookeriana</i> 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 18:3 α	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. <i>Cuphea hookeriana</i> 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 18:3 α	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. <i>Cuphea hookeriana</i> 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 18:3 α	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	

Construcción 14. <i>Umbellularia californica</i> 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 18:3 α	0.93	0.78	0.74	
C 20:0	0.33	0.29	0.27	

Construcción 15. <i>Umbellularia californica</i> 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 18:3 α	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. <i>Umbellularia californica</i> 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 18:3 α	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. <i>Umbellularia californica</i> 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 18:3 α	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. <i>Umbellularia californica</i> 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 18:3 α	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. <i>Cinnamomum camphora</i> 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 18:3 α	0.87	0.74	0.75	
C 20:0	0.33	0.33	0.31	

Construcción 20. <i>Cinnamomum camphora</i> 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 18:3 α	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. <i>Cinnamomum camphora</i> 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 18:3 α	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. <i>Cinnamomum camphora</i> 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 18:3 α	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. <i>Cinnamomum camphora</i> 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 18:3 α	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	

5 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 huésped. 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.

20 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 observaron con la Construcción 15 (la secuencia dirigida a plastidio de la isopentenil difosfato sintasa de *P.moriformis*) y la 17 (la secuencia dirigida a plastidio de la esteroil ACP desaturasa de *C.protothecoides*). 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 esteroil ACP desaturasa de *C.protothecoides* y la isopentenil difosfato sintasa de *P.moriformis* produjeron el mayor cambio en los niveles de ácidos grasos C12 con ambos promotores.

30 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 esteroil ACP desaturasa de *C.protothecoides* y la esteroil 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*.

- 5 Dos clones positivos de la Construcción 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 18:3 α	0.92	0.81
C 20:0	0.16	0.17

- 10 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 más 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 americana* en *Prototheca*

- 15 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	sec. con núm. ident..
Construcción 27	<i>C.reinhardtii</i> β -tubulina	<i>C.protothecoides</i> esteroil ACP desaturasa	<i>Cuphea palustris</i> tioesterasa	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 107

- 20 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 plastidios nativa, para la tioesterasa de *Ulmus americana* se muestra como sec. con núm. de ident.: 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)

- 25 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 huésped. 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

- 40 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 huésped para crecer (y utilizar) la sacarosa como única fuente de carbono en condiciones heterótrofas (como se demostró en el Ejemplo 3 anterior). 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 huésped. 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. sorokiniana* 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	sec. con núm. ident..
Construcción 24	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 109
Construcción 25	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> isopentenil difosfato sintasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 110
Construcción 26	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 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 se 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 sacarasa 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

5 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 huésped.

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

15 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 heterotróficamente 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 *Prototheca moriformis*

20 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 *Arabidopsis* 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 epítipo 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 sorokiniana* y una región nitrato reductasa 3'UTR de *Chlorella vulgaris*. La secuencia de ADN de este casete del transgén se muestra como sec. con núm. de ident.: 89 y la secuencia de aminoácidos traducida se muestra como sec. con núm. de ident.: 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.

40 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*

45 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	<i>P. krugani</i> (UTEX 329)	<i>P. stagnora</i> (UTEX 1442)	<i>P. zopfii</i> (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) (sec. con núm. de ident.: 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* hidrolizaron 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. Secuencias de promotores y 5'UTR de *Chlorella protothecoides*

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, 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 (sec. con núm. de ident.: 31) y el promotor del factor de elongación-1a (EF1a) (sec. con núm. de ident.: 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 (sec. con núm. de ident.: 33) y la isoforma B (sec. con núm. de ident.: 34).

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

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 (sec. con núm. de ident.: 45 y 46) y dos geranil geranil difosfato sintasas (sec. con núm. de ident.: 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 (sec. con núm. de ident.: 51 y 52), una proteína compleja de desprendimiento de oxígeno OEE33 (sec. con núm. de ident.: 53 y 54) y una proteasa Clp (sec. con núm. de ident.: 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.

5 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 entonces 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.

15 Como un primer examen, uno de los promotores putativos, el 5'UTR/promotor aislado de Aat2 (transportador de amonio, sec. con núm. de ident.: 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 sec. con núm. de ident.: 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	sec. con núm. ident.	veces de aumento
5'UTR/promotor FatB/A	sec. con núm. ident.: 91	n/a
5'UTR/promotor NRAMP transportadora de metales	sec. con núm. ident.: 92	9.65
5'UTR/promotor de la proteína Flap Flagelar asociada	sec. con núm. ident.: 93	4.92
5'UTR/promotor de sulfito reductasa SulfRed	sec. con núm. ident.: 94	10.91
5'UTR/promotor del transportador de azúcar SugT	sec. con núm. ident.: 95	17.35
5'UTR/promotor del transportador de amonio 03 Amt03	sec. con núm. ident.: 96	10.1
5'UTR/promotor del transportador de amonio 02 Amt02	sec. con núm. ident.: 97	10.76
5'UTR/promotor del transportador de aminoácido 01 Aat01	sec. con núm. ident.: 98	6.21
5'UTR/promotor del transportador de aminoácido 02 Aat02	sec. con núm. ident.: 99	6.5
5'UTR/promotor del transportador de aminoácido 03 Aat03	sec. con núm. ident.: 100	7.87
5'UTR/promotor del transportador de aminoácido 04 Aat04	sec. con núm. ident.: 101	10.95
5'UTR/promotor del transportador de aminoácido 05 Aat05	sec. con núm. ident.: 102	6.71

25 Las veces de incremento se refiere a las 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*

30 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 epítipo etiquetas, cambiar promotores de genes endógenos y otro tipo de modificaciones de genes diana (por ejemplo, la introducción de mutaciones puntuales).

40 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 (sec. con núm. de ident.: 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 (sec. con núm. de ident.: 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 total de ADN que contiene las secuencias KE858 que flanquean el gen de 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 las 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 SZ726 que 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 SZ726 (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 evento de 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 epítipo 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 anterior, se identificó el ADNc de una esteroil ACP desaturasa endógena (SAPD). Las enzimas esteroil 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 esteroil 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 huésped. Una secuencia codificante de aproximadamente 1.5kb de un gen de esteroil ACP desaturasa de *Prototheca moriformis* (UTEX 1435) se identificó y se clonó (sec. con núm. de ident.: 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 -tubulina 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 sec. con núm. de ident.: 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 linealizó y luego se transformó en la célula huésped (*Prototheca moriformis*, UTEX 1435). Una integración exitosa eliminaría del genoma huésped 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 de 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 entonces 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-trimetilsilil (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). El análisis 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 divulgación proporciona comida deslipidada que tiene el contenido de los productos establecidos en la tabla anterior.

C. Producción de biodiesel a partir de aceite de *Prototheca*

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
5	C18:1	59.21
	C18:2	8.91
	C18:3	0.28
	C20:0	0.23
	C20:1	0.10
10	C20:1	0.08
	C21:0	0.02
	C22:0	0.06
	C24:0	0.10

15 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	Viscosidad cinemática / dinámica	viscosidad cinemática @ 104 °F/ 40 °C	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 50°C (122°F)/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 lubricantes por ICP-AES	Fósforo	0.000200	%p

Método	Prueba	Resultado	Unidades	
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	%
		% Pérdidas	1.5	%
		% Residuo	0.0	%
		Volumen de la trampa para frío	0.0	ml
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	

5 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 divulgadas se pueden someter a transesterificación para producir 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.

10 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 divulgados se pueden transesterificar para generar biodiesel con un tiempo de inmersión en frío de menos de 120 segundos, menos de 100 segundos y menos de 90 segundos.

D. La producción de diesel renovable

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.

15 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 entonces a craqueo e isomerización. El cromatograma del material se muestra en la Figura 13. El material se sometió entonces 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 muestran en la Tabla 20:

Tabla 20. Lecturas en °C:

	Volumen	Temperatura
	IBP	173
5	5	217.4
	10	242.1
	15	255.8
	20	265.6
	30	277.3
10	40	283.5
	50	286.6
	60	289.4
	70	290.9
	80	294.3
	90	300
15	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 divulgados en la presente, así como los métodos de destilación y fraccionamiento (tales como la filtración en frío) divulgados 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 divulgados 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 divulgados en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se divulgan 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 divulgados en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se divulgan 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 divulgan en la presente, así como los métodos de destilación y fraccionamiento (como la filtración en frío) que se divulgan 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 se proporcionan en la presente 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. Crecimiento de la SAG 2214 en glucosa y sacarosa

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. Productividad de lípidos y el perfil de ácidos grasos de la SAG 2214

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 anterior. 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.

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

5 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 (sec. con núm. de ident.: 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 (sec. con núm. de ident.: 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 simbiotes 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. 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. 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 de 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 oleaginosos."; 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 de los Estados Unidos núm. 61/118,994, presentada el 1 de diciembre de 2008, titulada "Producción de

aceite en microorganismos"; la solicitud de patente provisional de Estados Unidos núm. 61/174,357, presentada el 3 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 específicos en microorganismos heterótrofos recombinantes".

5

LISTADO DE SECUENCIAS

<110> SOLAZYME, INC.
5
<120> PRODUCCIÓN DE ACEITES ESPECÍFICOS EN MICROORGANISMOS HETERÓTROFOS
<130> 026172-005010PC
10 <140> PCT/US2009/066142
<141> 2009-11-30
<150> 61/219,525
<151> 2009-06-23
15 <150> 61/174,357
<151> 2009-04-30
<150> 61/118,994
20 <151> 2008-12-01
<150> 61/118,590
<151> 2008-11-28
25 <160> 143
<170> PatentIn versión 3.5
<210> 1
30 <211> 1187
<212> ADN
<213> Chlorella sp.

ES 2 583 639 T3

<400> 1

	gatcagacgg gcctgacctg cgagataatc aagtgctcgt aggcaaccaa ctcagcagct	60
	gcttgggtgtt gggctctgcag gatagtgttg cagggcccca aggacagcag gggaaacttac	120
5	accttgtccc cgacctcagtt ttatggagtg cattgcctca agagcctagc cggagcgcta	180
	ggctacatac ttgccgcacc ggtatgaggg gatatagtac tcgcactgcg ctgtctagtg	240
10	agatgggcag tgctgcccac aaacaactgg ctgctcagcc atttgttggc ggaccattct	300
	gggggggcca gcaatgcctg actttcgggt agggtgaaaa ctgaacaaag actaccaaaa	360
	cagaatttct tcctccttgg aggtaagcgc aggccggccc gcctgcgccc acatggcgct	420
15	ccgaacacct ccatagctgt aagggcgcaa acatggccgg actgtttgtca gactctttc	480
	atggccatac aaggatcatgt cgagattagt gctgagtaag aactatcac cccatgttcg	540
	attgaagccg tgacttcatg ccaacctgcc cctgggcgta gcagacgtat gccatcatga	600
20	ccactagccg acatgcgctg tcttttgcca ccaaaacaac tggtagaccg ctcgaagtcg	660
	tgccgcacac ctccgggagt gagtccggcg actcctcccc ggcgggccgc ggcctacct	720
25	gggtagggtc gccatacgcc cagacacaaa cgacgcagga ggggattggg gtagggaatc	780
	ccaaccagcc taaccaagac ggcacctata ataataggtg gggggactaa cagccctata	840
	tcgaagctt tgggtgccta tcttgagaag cacgagttgg agtggctgtg tacggctgac	900
30	cctaaggtgg gtgtgccgca gcctgaaaca aagcgtctag cagctgcttc tataatgtgt	960

ES 2 583 639 T3

	cagccgttgt gtttcagtta tattgtatgc tattgtttgt tcgtgctagg gtggcgcagg	1020
	cccacctact gtggcgggcc attggttggt gcttgaattg cctcaccatc taaggctctga	1080
5	acgctcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagtg	1140
	tgcaaaccag cacatagcac tcccttacat caccagcag tacaaca	1187
10	<210> 2	
	<211> 1414	
	<212> ADN	
	<213> Chlorella ellipsoidea	
15	<400> 2	
	cgctgcgcac cagggccgcc agctcgctga tgtcgctcca aatgcggtcc cccgattttt	60
	tgttcttcat cttctccacc ttggtggcct tcttggccag ggccttcagc tgcattgcga	120
20	cagaccgttg agctcctgat cagcatcctc aggaggcctt ttgacaagca agcccctgtg	180
	caagcccatt cacggggtac cagtggctgct gaggtagatg ggtttgaaa ggattgctcg	240
	gtcgattgct gctcatggaa ttggcatgtg catgcatgtt cacaatatgc caccaggctt	300
25	tggagcaaga gagcatgaat gccttcaggc aggttgaaag ttcttggggg tgaagaggca	360
	gggccgagga ttggaggagg aaagcatcaa gtcgctgctc atgctcatgt tttcagtcag	420
	agtttgccaa gctcacagga gcagagacaa gactggctgc tcagggtgtg catcgtgtgt	480
30	gtggtggggg ggggggggtt aatacggtac gaaatgcact tggaattccc acctcatgcc	540

ES 2 583 639 T3

agcggacca catgcttgaa ttcgaggcct gtggggtgag aaatgctcac tctgccctcg 600

ttgctgaggt acttcaggcc gctgagctca aagtcgatgc cctgctcgtc tatcagggcc 660

5 tgcacctctg ggctgaccgg ctcagcctcc ttcgcgggca tggagtaggc gccggcagcg 720

ttcatgtccg ggcccagggc agcgggtggtg ccataaatgt cggatgatgt ggggaggggg 780

gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc 840

10 ggcagcactg gtctccagcc agccagcaag tggctgttca ggaaagcggc catgtttgtg 900

gtccctgctc atgtaattcc ccagatcaaa ggaggaaca gcttggattt gatgtagtgc 960

15 ccaaccggac tgaatgtgcg atggcaggtc cttttgagtc tcccgaatta ctagcagggc 1020

actgtgacct aacgcagcat gcccaaccgca aaaaaatgat tgacagaaaa tgaagcggtg 1080

tgtcaatatt tgctgtattt attcgtttta atcagcaacc aagttcgaaa cgcaactatc 1140

20 gtggtgatca agtgaacctc atcagactta cctcgttcgg caaggaaacg gaggcaccaa 1200

attccaattt gatattatcg cttgccaagc tagagctgat ctttgggaaa ccaactgcca 1260

25 gacagtggac tgtgatggag tgccccgagt ggtggagcct cttcgattcg gttagtcatt 1320

actaacgtga accctcagtg aagggacat cagaccagaa agaccagatc tcctcctcga 1380

caccgagaga gtgttcggc agtaggacga caag 1414

30

<210> 3

ES 2 583 639 T3

<211> 512

<212> PRT

<213> Desconocido

5 <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

10 1 5 10 15

Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr Asp Glu Lys Asp

15 20 25 30

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

20 35 40 45

Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp Asp Leu Thr Asn

25 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

30 85 90 95

ES 2 583 639 T3

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

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

10 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

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

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

25 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

30 Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser Phe Asn Gln Tyr

ES 2 583 639 T3

	225		230		235		240
	Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala Phe Asp Asn Gln						
5		245		250		255	
	Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala Leu Gln Thr Phe						
10		260		265		270	
	Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly Ile Ala Trp Ala						
15		275		280		285	
	Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn Pro Trp Arg Ser						
20		290		295		300	
	Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr Glu Tyr Gln Ala						
25		305		310		315	
	Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu Pro Ile Leu Asn						
30		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						

ES 2 583 639 T3

355 360 365
Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr Gln Thr Ile Ser
370 375 380

5
Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys Gly Leu Glu Asp
385 390 395 400

10
Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser Ala Ser Ser Phe
405 410 415

15
Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val Lys Glu Asn Pro
420 425 430

20
Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro Phe Lys Ser Glu
435 440 445

25
Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu Asp Gln Asn Ile
450 455 460

30
Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn Met Thr Thr Gly
485 490 495

ES 2 583 639 T3

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

5

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

10

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

<400> 4

15

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

20

Ile Ser Ala Ser
20

25

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

30

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

<400> 5

Met Ala Asn Lys Ser Leu Leu Leu Leu Leu Leu Gly Ser Leu Ala

ES 2 583 639 T3

1 5 10 15

Ser Gly

5

<210> 6

<211> 10

10

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Secuencia de consenso
sintética

15

<400> 6

Met Ala Arg Leu Pro Leu Ala Ala Leu Gly

1 5 10

20

<210> 7

<211> 20

<212> PRT

25

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Péptido
sintético

30

<400> 7

Met Ala Asn Lys Leu Leu Leu Leu Leu Leu Leu Leu Leu Pro Leu

ES 2 583 639 T3

1 5 10 15

Ala Ala Ser Gly

5 20

<210> 8

<211> 2615

10 <212> ADN

<213> *Saccharomyces cerevisiae*

<400> 8

gaattcccca acatggtgga gcacgacact ctcgtctact ccaagaatat caaagataca 60

15 gtctcagaag accaaagggc tattgagact tttcaacaaa gggtaatatc gggaaacctc 120

ctcggattcc attgccagc tatctgtcac ttcatacaaaa ggacagtaga aaaggaaggt 180

20 ggcacctaca aatgccatca ttgcgataaa ggaaaggcta tcgttcaaga tgccctctgcc 240

gacagtggtc ccaaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt 300

ccaaccacgt cttcaaagca agtggattga tgtgaacatg gtggagcacg acaactctcgt 360

25 ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca 420

acaaagggta atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat 480

30 caaaaggaca gtagaaaagg aaggtggcac ctacaaatgc catcattgcg ataaaggaaa 540

ggctatcggt caagatgcct ctgccgacag tgggtcccaaa gatggacccc caccacagag 600

ES 2 583 639 T3

gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga 660

tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc 720

5 tatataagga agttcatttc atttggagag gacacgctga aatcaccagt ctctctctac 780

aaatctatct ctggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 840

10 gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcgataga ccacttgctg 900

actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 960

atgccaagtg gcatctgtac tttcaataca acccgaacga tactgtctgg gggacgcat 1020

15 tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 1080

tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 1140

20 atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 1200

cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca 1260

cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 1320

25 cgaaggcttt ttggtacgag ccctcgcaga agtggatcat gacagcggca aagtcacagg 1380

actacaagat cgaaatttac tcgtctgacg accttaaatc ctggaagctc gaatccgcgt 1440

30 tcgcaaacga gggctttctc ggctaccaat acgaatgcc aggctgata gaggtcccaa 1500

cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1560

ES 2 583 639 T3

	cggcaggagg ttcttttaat cagtacttcg tcggaagcct taacggaact catttcgagg	1620
	catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt	1680
5	tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg	1740
	agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgctc gtgaggaaat	1800
10	tctctctcaa cactgagtac caggccaacc cggaaccga actcataaac ctgaaagccg	1860
	aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt	1920
	tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccggtaca cttgaatttg	1980
15	aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct	2040
	ccctctgggt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt	2100
20	ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttggt aaggagaacc	2160
	catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt	2220
	cgactacaa agtgataggt ttgcttgatc aaaataticc tgaactctac ttcaacgatg	2280
25	gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga	2340
	acatgacgac ggggtgtgat aacctgttct acatcgaaa attccaggtg agggaagtca	2400
30	agtgagatct gtcgatcgac aagctcgagt ttctccataa taatgtgtga gtagttcca	2460
	gataagggaa ttagggttcc tatagggttt cgctcatgtg ttgagcatat aagaaaccct	2520

ES 2 583 639 T3

tagtatgtat ttgtatttgt aaaatacttc tatcaataaa atttctaatt cctaaaacca 2580

aaatccagta ctaaaatcca gatccccga attaa 2615

5

<210> 9

<211> 22

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

15

<400> 9

tgttgaagaa tgagccggcg ac 22

20 <210> 10

<211> 20

<212> ADN

<213> Secuencia Artificial

25 <220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 10

30 cagtgagcta ttacgcactc 20

ES 2 583 639 T3

<210> 11

<211> 541

<212> ADN

<213> Prototheca kruegani

5

<400> 11

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag 60

ccatagcgaa agcaagttaa gtaagcttag gtcattcttt ttagaccgga aaccgagtga 120

10

tctacccatg atcaggggtga agtgtagta aaataacatg gaggcccgaa ccgactaatg 180

ttgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

15

ctggttctcc ccgaaatgcg tttagggcga gcagtagcag tacaaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc 420

20

cagatcacca gttaaggccc caaatgaaa atgatagtga ctaaggatgt gggtagtca 480

aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540

25

g 541

<210> 12

<211> 573

30

<212> ADN

<213> Prototheca wickerhamii

ES 2 583 639 T3

<400> 12

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga 60

aacctaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat 120

5

aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240

10

tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360

ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

15

gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540

20

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 13

<211> 541

25

<212> ADN

<213> Prototheca stagnora

<400> 13

tgttgaagaa tgagccggcg agttaaaaaa aatggcatgg ttaaagatat ttctctgaag 60

30

ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagaccgga aaccgagtga 120

ES 2 583 639 T3

	tctacccatg atcagggatga agtggttggtc aaataacatg gaggcccgaa ccgactaatg	180
	gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
5	ctggttctcc ccgaaatgcg tttagggcgca gcagtagcaa cacaaataga ggggtaaagc	300
	actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta	360
	tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc	420
10	cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggacgt gagtatgtca	480
	aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
15	g	541
	<210> 14	
	<211> 541	
	<212> ADN	
	<213> Prototheca moriformis	
20		
	<400> 14	
	tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag	60
	ccatagcgaa agcaagttta acaagctaaa gtcacccttt ttagaccgga aaccgagtga	120
25		
	tctacccatg atcagggatga agtggttggtc aaataacatg gaggcccgaa ccgactaatg	180
	gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
30	ctggttctcc ccgaaatgcg tttagggcgca gcagtagcaa cacaaataga ggggtaaagc	300
	actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta	360

ES 2 583 639 T3

tttagatatic tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc 420
 cagatcacca gtttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggtagtatta 480
 5
 aaacctccag caggtagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
 g 541
 10
 <210> 15
 <211> 573
 <212> ADN
 <213> Prototheca moriformis
 15
 <400> 15
 tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga 60
 aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct 120
 20
 aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180
 gtgacaccaa gtggaagtcc gaaccgaccg atggtgaaaa atcggcggat gaactgtggt 240
 tagtggtgaa ataccagtcg aactcagagc tagctgggtc tccccgaaat gcgttgaggc 300
 25
 gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat 360
 ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420
 30
 gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaatg 480
 ataataaggt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca 540

ES 2 583 639 T3

	gccatccttt aaagagtgcg taatagctca ctg	573
5	<210> 16 <211> 573 <212> ADN <213> Prototheca wickerhamii	
10	<400> 16 tgttgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga	60
	aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct	120
15	aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
	gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
	tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc	300
20	gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat	360
	ggtagcaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
25	gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaatg	480
	ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca	540
	gccatccttt aaagagtgcg taatagctca ctg	573
30	<210> 17 <211> 541 <212> ADN	

ES 2 583 639 T3

<213> Prototheca moriformis

<400> 17

5	tgttgaagaa tgagccggcg agttaaaaag agtggcgtgg ttaaagaaaa ttctctggaa	60
	ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagaccgga aaccgagtga	120
	tctacccatg atcagggatg agtggttgta aaataacatg gaggcccgaa ccgactaatg	180
10	gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
	ctggttctcc ccgaaatgcy tttaggcgca gcagtagcaa cacaaataga ggggtaaagc	300
	actgtttctt ttgtgggctc cgaaagttgt acctcaaagt ggcaaactct gaatactcta	360
15	tttagatatc tactagttag accttggggg ataagctcct tggtcgaaag ggaaacagcc	420
	cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca	480
20	aaacctccag caggtagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
	g	541

25 <210> 18

<211> 541

<212> ADN

<213> Prototheca zopfii

30 <400> 18

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag	60
---	----

ES 2 583 639 T3

	ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga	120
	tctacccatg atcagggtga agtgttgta aaataacatg gaggcccgaa ccgactaatg	180
5	gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
	ctggttctcc ccgaaatgcg tttagggcgca gcagtagcaa cacaaataga ggggtaaagc	300
	actgtttctt tcgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta	360
10	tttagatatc tactagttag accttggggg ataagctcct tggcmetaag ggaacagcc	420
	cagatcacca gttaaggccc caaatgaaa atgatagtga ctaaggatgt gagtatgtca	480
	aaacctccag caggtagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
15	g	541
	<210> 19	
20	<211> 565	
	<212> ADN	
	<213> <i>Prototheca moriformis</i>	
	<400> 19	
25	tgttgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc	60
	cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatatttaga atctagtcac	120
	tttttctaga cccgaaccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc	180
30	aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcggtg	240

ES 2 583 639 T3

aaataccagt cgaacccgga gctagctggt tctccccgaa atgcggtgag gcgcagcagt 300
 acatctagtc tatctagggg taaagcactg tttcggtgcg ggctgtgaga acggtaccaa 360
 5 atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag 420
 ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
 gtgacaaagg aggtgaaaat gcaaatacaa ccaggagggt ggcttagaag cagccatcct 540
 10 ttaaagagtg cgtaatagct cactg 565
 <210> 20
 <211> 550
 <212> PRT
 15 <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
 20 Glu Gln Pro Tyr Arg Thr Ala Phe His Phe Gln Pro Pro Gln Asn Trp
 20 25 30
 25 Met Asn Asp Pro Asn Gly Pro Met Cys Tyr Asn Gly Val Tyr His Leu
 35 40 45
 30 Phe Tyr Gln Tyr Asn Pro Phe Gly Pro Leu Trp Asn Leu Arg Met Tyr
 50 55 60

ES 2 583 639 T3

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

5

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

10

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

15

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

20

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

25

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

30

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

ES 2 583 639 T3

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

5

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

10

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

15

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

20

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

25

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

30

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

ES 2 583 639 T3

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

5

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

10

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

15

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

20

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

25

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

30

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

ES 2 583 639 T3

	450	455	460	
	Glu Lys Arg Thr Tyr Gly Ala Phe Val Asp Ile Asp Pro Gln Gln Asp			
5	465	470	475	480
	Glu Ile Ser Leu Arg Thr Leu Ile Asp His Ser Ile Val Glu Ser Phe			
10		485	490	495
	Gly Gly Arg Gly Lys Thr Cys Ile Thr Thr Arg Val Tyr Pro Thr Leu			
15		500	505	510
	Ala Ile Gly Glu Gln Ala Arg Leu Phe Ala Phe Asn His Gly Thr Glu			
20		515	520	525
	Ser Val Glu Ile Ser Glu Leu Ser Ala Trp Ser Met Lys Lys Ala Gln			
25		530	535	540
	Met Lys Val Glu Glu Pro			
30		545	550	
	<210> 21			
	<211> 581			
	<212> PRT			
	<213> Schizosaccharomyces pombe			

ES 2 583 639 T3

<400> 21

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

5

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

10

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

15

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

20

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

25

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

30

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

ES 2 583 639 T3

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

5

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

10

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

15

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

20

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

25

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

30

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

ES 2 583 639 T3

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

5 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

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

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

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

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

30 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

ES 2 583 639 T3

	385		390			395		400
	Thr Ser Leu Ile Gln Thr Pro Leu Asn Val Ser Leu Leu Arg Asp Glu							
5			405			410		415
	Thr Leu Phe Thr Ala Pro Val Ile Asn Ser Ser Ser Ser Leu Ser Gly							
10			420			425		430
	Ser Pro Ile Thr Leu Pro Ser Asn Thr Ala Phe Glu Phe Asn Val Thr							
15			435			440		445
	Leu Ser Ile Asn Tyr Thr Glu Gly Cys Thr Thr Gly Tyr Cys Leu Gly							
			450			455		460
20	Arg Ile Ile Ile Asp Ser Asp Asp Pro Tyr Arg Leu Gln Ser Ile Ser							
			465			475		480
	Val Asp Val Asp Phe Ala Ala Ser Thr Leu Val Ile Asn Arg Ala Lys							
25			485			490		495
	Ala Gln Met Gly Trp Phe Asn Ser Leu Phe Thr Pro Ser Phe Ala Asn							
			500			505		510
30	Asp Ile Tyr Ile Tyr Gly Asn Val Thr Leu Tyr Gly Ile Val Asp Asn							
			515			520		525

ES 2 583 639 T3

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

5

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

10

Ala Phe Gln Gly Val Ser Phe Asn Asn Val Thr Val Thr Pro Leu Lys
565 570 575

Thr Ile Trp Asn Cys

15

580

<210> 22

<211> 550

20

<212> PRT

<213> Pichia anomala

<400> 22

Met Ile Gln Leu Ser Pro Leu Leu Leu Leu Pro Leu Phe Ser Val Phe

25

1

5

10

15

Asn Ser Ile Ala Asp Ala Ser Thr Glu Tyr Leu Arg Pro Gln Ile His

20

25

30

30

Leu Thr Pro Asp Gln Gly Trp Met Asn Asp Pro Asn Gly Met Phe Tyr

ES 2 583 639 T3

35

40

45

Asp Arg Lys Asp Lys Leu Trp His Val Tyr Phe Gln His Asn Pro Asp

5

50

55

60

Lys Lys Ser Ile Trp Ala Thr Pro Val Thr Trp Gly His Ser Thr Ser

10

65

70

75

80

Lys Asp Leu Leu Thr Trp Asp Tyr His Gly Asn Ala Leu Glu Pro Glu

15

85

90

95

Asn Asp Asp Glu Gly Ile Phe Ser Gly Ser Val Val Val Asp Arg Asn

20

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

25

130

135

140

Tyr Ser Leu Asp Lys Gly Tyr Ser Phe Ile Lys Tyr Asp Gln Asn Pro

30

145

150

155

160

Val Ile Asn Val Asn Ser Ser Gln Gln Arg Asp Pro Lys Val Leu Trp

ES 2 583 639 T3

165

170

175

His Asp Glu Ser Asn Gln Trp Ile Met Val Val Ala Lys Thr Gln Glu

5

180

185

190

Phe Lys Val Gln Ile Tyr Gly Ser Pro Asp Leu Lys Lys Trp Asp Leu

10

195

200

205

Lys Ser Asn Phe Thr Ser Asn Gly Tyr Leu Gly Phe Gln Tyr Glu Cys

15

210

215

220

Pro Gly Leu Phe Lys Leu Pro Ile Glu Asn Pro Leu Asn Asp Thr Val

225

230

235

240

20

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

25

260

265

270

Phe His Pro Asp Asp Gly Ala Thr Arg Phe Met Asp Ile Gly Lys Asp

30

275

280

285

Phe Tyr Ala Phe Gln Ser Phe Asp Asn Thr Glu Pro Glu Asp Gly Ala

ES 2 583 639 T3

290

295

300

Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asn Thr Val Pro

5 305 310 315 320

Thr Glu Asn Trp Arg Ser Ser Met Ser Leu Val Arg Asn Tyr Thr Leu

325 330 335

10 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

15 355 360 365

Thr Leu Glu Thr Ile Asn Glu Tyr Glu Val Asn Asp Leu Lys Leu Asp

370 375 380

20

Lys Ser Ser Phe Val Ala Thr Asp Phe Asn Thr Glu Arg Asn Ala Thr

385 390 395 400

25

Gly Val Phe Glu Phe Asp Leu Lys Phe Thr Gln Thr Asp Leu Lys Met

405 410 415

30 Gly Tyr Ser Asn Met Thr Thr Gln Phe Gly Leu Tyr Ile His Ser Gln

420 425 430

ES 2 583 639 T3

Thr Val Lys Gly Ser Gln Glu Thr Leu Gln Leu Val Phe Asp Thr Leu
435 440 445

5

Ser Thr Thr Trp Tyr Ile Asp Arg Thr Thr Gln His Ser Phe Gln Arg
450 455 460

10

Asn Ser Pro Val Phe Thr Glu Arg Ile Ser Thr Tyr Val Glu Lys Ile
465 470 475 480

15

Asp Thr Thr Asp Gln Gly Asn Val Tyr Thr Leu Tyr Gly Val Val Asp
485 490 495

20

Arg Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Ser Ile Ala Met Thr
500 505 510

25

Asn Thr Phe Phe Phe Arg Glu Gly Lys Ile Pro Thr Ser Phe Glu Val
515 520 525

30

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

ES 2 583 639 T3

<211> 533

<212> PRT

<213> *Debaryomyces occidentalis*

5 <400> 23

Met Val Gln Val Leu Ser Val Leu Val Ile Pro Leu Leu Thr Leu Phe
1 5 10 15

10 Phe Gly Tyr Val Ala Ser Ser Ser Ile Asp Leu Ser Val Asp Thr Ser
20 25 30

15 Glu Tyr Asn Arg Pro Leu Ile His Phe Thr Pro Glu Lys Gly Trp Met
35 40 45

20 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

25 Tyr Trp Gly His Ala Thr Ser Asn Asp Leu Val His Trp Asp Glu His
85 90 95

30 Glu Ile Ala Ile Gly Pro Glu His Asp Asn Glu Gly Ile Phe Ser Gly
100 105 110

ES 2 583 639 T3

Ser Ile Val Val Asp His Asn Asn Thr Ser Gly Phe Phe Asn Ser Ser
 115 120 125

5

Ile Asp Pro Asn Gln Arg Ile Val Ala Ile Tyr Thr Asn Asn Ile Pro
 130 135 140

10

Asp Leu Gln Thr Gln Asp Ile Ala Phe Ser Leu Asp Gly Gly Tyr Thr
 145 150 155 160

15

Phe Thr Lys Tyr Glu Asn Asn Pro Val Ile Asp Val Ser Ser Asn Gln
 165 170 175

20

Phe Arg Asp Pro Lys Val Phe Trp His Glu Arg Phe Lys Ser Met Asp
 180 185 190

25

His Gly Cys Ser Glu Ile Ala Arg Val Lys Ile Gln Ile Phe Gly Ser
 195 200 205

30

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

ES 2 583 639 T3

Asn Ser Asp Lys Ser Lys Trp Val Met Phe Leu Ala Ile Asn Pro Gly
245 250 255

5

Ser Pro Leu Gly Gly Ser Ile Asn Gln Tyr Phe Val Gly Asp Phe Asp
260 265 270

10

Gly Phe Gln Phe Val Pro Asp Asp Ser Gln Thr Arg Phe Val Asp Ile
275 280 285

15

Gly Lys Asp Phe Tyr Ala Phe Gln Thr Phe Ser Glu Val Glu His Gly
290 295 300

20

Val Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asp Gln Val
305 310 315 320

25

Pro Thr Asn Pro Trp Arg Ser Ser Thr Ser Leu Ala Arg Asn Tyr Thr
325 330 335

30

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

ES 2 583 639 T3

Asn Val Lys Leu Thr Asn Lys Lys Pro Ile Lys Thr Asn Phe Lys Gly
370 375 380

5

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

10

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

15

Gln Ser Leu Phe Tyr Ile Asp Arg His Ile Pro Asn Val Glu Phe Pro
435 440 445

20

Arg Lys Gln Phe Phe Thr Asp Lys Leu Ala Ala Tyr Leu Glu Pro Leu
450 455 460

25

Asp Tyr Asp Gln Asp Leu Arg Val Phe Ser Leu Tyr Gly Ile Val Asp
465 470 475 480

30

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

ES 2 583 639 T3

500

505

510

Val Thr Asp Thr Glu Glu Pro Leu Phe Glu Leu Glu Ser Val Ile Ile

5

515

520

525

Arg Glu Leu Asn Lys

530

10

<210> 24

<211> 654

<212> PRT

15

<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

20

Ala Ala Val Tyr Thr Pro Leu Pro Glu Gln Pro His Ser Ala Glu Val

20

25

30

25

Glu Ile Arg Asp Arg Lys Pro Phe Lys Ile Ile Ser Ala Ile Ile Leu

35

40

45

30

Ser Ser Leu Leu Leu Leu Ala Leu Ile Leu Val Ala Val Asn Tyr Gln

50

55

60

ES 2 583 639 T3

Ala Pro Pro Ser His Ser Ser Gly Asp Asn Ser Gln Pro Ala Ala Val
 65 70 75 80

5

Met Pro Pro Ser Arg Gly Val Ser Gln Gly Val Ser Glu Lys Ala Phe
 85 90 95

10

Arg Gly Ala Ser Gly Ala Gly Asn Gly Val Ser Phe Ala Trp Ser Asn
 100 105 110

15

Leu Met Leu Ser Trp Gln Arg Thr Ser Tyr His Phe Gln Pro Val Lys
 115 120 125

20

Asn Trp Met Asn Asp Pro Asn Gly Pro Leu Tyr Tyr Lys Gly Trp Tyr
 130 135 140

25

His Leu Phe Tyr Gln Tyr Asn Pro Asp Ser Ala Val Trp Gly Asn Ile
 145 150 155 160

30

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

ES 2 583 639 T3

Thr Gly Ser Ala Thr Ile Leu Pro Asp Gly Arg Ile Val Met Leu Tyr
 195 200 205

5

Thr Gly Asp Thr Asp Asp Tyr Val Gln Asp Gln Asn Leu Ala Phe Pro
 210 215 220

10

Ala Asn Leu Ser Asp Pro Leu Leu Val Asp Trp Val Lys Tyr Pro Asn
 225 230 235 240

15

Asn Pro Val Ile Tyr Pro Pro Pro Gly Ile Gly Val Lys Asp Phe Arg
 245 250 255

20

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

25

Glu Thr Thr Asn Phe Thr Thr Phe Lys Leu Leu Tyr Gly Val Leu His
 290 295 300

30

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

ES 2 583 639 T3

325 330 335

Val Lys His Val Leu Lys Thr Ser Leu Asp Asp Asp Lys His Asp Tyr
5 340 345 350

Tyr Ala Leu Gly Thr Tyr Asp Pro Val Lys Asn Lys Trp Thr Pro Asp
10 355 360 365

Asn Pro Asp Leu Asp Val Gly Ile Gly Leu Arg Leu Asp Tyr Gly Lys
15 370 375 380

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
20 405 410 415

Lys Gly Trp Ala Ser Leu Gln Ala Ile Pro Arg Thr Ile Val Phe Asp
25 420 425 430

Lys Lys Thr Gly Thr Asn Val Leu Gln Arg Pro Glu Glu Glu Val Glu
30 435 440 445

Ser Trp Ser Ser Gly Asp Pro Ile Thr Gln Arg Arg Ile Phe Glu Pro

ES 2 583 639 T3

	450		455		460											
	Gly	Ser	Val	Val	Pro	Ile	His	Val	Ser	Gly	Ala	Thr	Gln	Leu	Asp	Ile
5	465				470					475					480	
	Thr	Ala	Ser	Phe	Glu	Val	Asp	Glu	Thr	Leu	Leu	Glu	Thr	Thr	Ser	Glu
					485					490					495	
10	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
15		515						520						525		
	Leu	Ser	Glu	Leu	Thr	Pro	Val	Tyr	Leu	Tyr	Val	Ala	Lys	Gly	Gly	Asp
	530						535					540				
20																
	Gly	Lys	Ala	Lys	Ala	His	Leu	Cys	Ala	Tyr	Gln	Thr	Arg	Ser	Ser	Met
	545				550						555					560
25																
	Ala	Ser	Gly	Val	Glu	Lys	Glu	Val	Tyr	Gly	Ser	Ala	Val	Pro	Val	Leu
					565					570					575	
30	Asp	Gly	Glu	Asn	Tyr	Ser	Ala	Arg	Ile	Leu	Ile	Asp	His	Ser	Ile	Val
				580						585					590	

ES 2 583 639 T3

Glu Ser Phe Ala Gln Ala Gly Arg Thr Cys Val Arg Ser Arg Asp Tyr
595 600 605

5

Pro Thr Lys Asp Ile Tyr Gly Ala Ala Arg Cys Phe Phe Phe Asn Asn
610 615 620

10

Ala Thr Glu Ala Ser Val Arg Ala Ser Leu Lys Ala Trp Gln Met Lys
625 630 635 640

15

Ser Phe Ile Arg Pro Tyr Pro Phe Ile Pro Asp Gln Lys Ser
645 650

<210> 25

<211> 690

20

<212> PRT

<213> Allium cepa

<400> 25

25

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

30

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

ES 2 583 639 T3

35

40

45

Ala Thr Phe Leu Ser Asn Leu Glu Ser Pro Pro Ser Asp Ser Gly Leu

5

50

55

60

Val Ser Asp Pro Val Thr Phe Asp Val Asn Pro Ala Val Val Arg Arg

10

65

70

75

80

Gly Lys Asp Ala Gly Val Ser Asp Lys Thr Ser Gly Val Asp Ser Gly

15

85

90

95

Phe Val Leu Asp Pro Val Ala Val Asp Ala Asn Ser Val Val Val His

20

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

25

130

135

140

Met Leu Ser Trp Gln Arg Thr Gly Phe His Phe Gln Pro Val Lys Asn

30

145

150

155

160

Trp Met Asn Asp Pro Asn Gly Pro Leu Tyr Tyr Lys Gly Trp Tyr His

ES 2 583 639 T3

165

170

175

Phe Phe Tyr Gln Tyr Asn Pro Glu Gly Ala Val Trp Gly Asn Ile Ala

5

180

185

190

Trp Gly His Ala Val Ser Arg Asp Leu Val His Trp Thr His Leu Pro

10

195

200

205

Leu Ala Met Val Pro Asp Gln Trp Tyr Asp Ile Asn Gly Val Trp Thr

15

210

215

220

Gly Ser Ala Thr Ile Leu Pro Asp Gly Gln Ile Val Met Leu Tyr Thr

225

230

235

240

20

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

25

Pro Ile Leu Val Pro Pro Pro Gly Ile Gly Asp Lys Asp Phe Arg Asp

275

280

285

30

Pro Thr Thr Ala Trp Tyr Glu Pro Ser Asp Asp Thr Trp Arg Ile Val

290

295

300

ES 2 583 639 T3

Ile Gly Ser Lys Asp Ser Ser His Ser Gly Ile Ala Ile Val Tyr Ser
305 310 315 320

5

Thr Lys Asp Phe Ile Asn Tyr Lys Leu Ile Pro Gly Ile Leu His Ala
325 330 335

10

Val Glu Arg Val Gly Met Trp Glu Cys Val Asp Phe Tyr Pro Val Ala
340 345 350

15

Thr Ala Asp Ser Ser His Ala Asn His Gly Leu Asp Pro Ser Ala Arg
355 360 365

20

Pro Ser Pro Ala Val Lys His Val Leu Lys Ala Ser Met Asp Asp Asp
370 375 380

25

Arg His Asp Tyr Tyr Ala Ile Gly Thr Tyr Asp Pro Ala Gln Asn Thr
385 390 395 400

30

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

ES 2 583 639 T3

Lys Arg Arg Ile Leu Trp Ser Trp Ile Gly Glu Thr Asp Ser Glu Thr
435 440 445

5

Ala Asp Ile Ala Lys Gly Trp Ala Ser Leu Gln Gly Val Pro Arg Thr
450 455 460

10

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

15

Thr Val Asp Ala Gly Ser Thr Phe Lys Leu Asp Val Gly Gly Ala Ala
500 505 510

20

Gln Leu Asp Ile Glu Ala Glu Phe Lys Ile Ser Ser Glu Glu Leu Glu
515 520 525

25

Ala Val Lys Glu Ala Asp Val Ser Tyr Asn Cys Ser Ser Ser Gly Gly
530 535 540

30

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

ES 2 583 639 T3

565

570

575

Gly Met Asp Gly Gly Leu Asn Thr His Phe Cys Gln Asp Glu Lys Arg

5

580

585

590

Ser Ser Lys Ala Ser Asp Ile Val Lys Arg Ile Val Gly His Ser Val

595

600

605

10

Pro Val Leu Asp Gly Glu Ser Phe Ala Leu Arg Ile Leu Val Asp His

610

615

620

15

Ser Ile Val Glu Ser Phe Ala Gln Gly Gly Arg Ala Ser Ala Thr Ser

625

630

635

640

20

Arg Val Tyr Pro Thr Glu Ala Ile Tyr Asn Asn Ala Arg Val Phe Val

645

650

655

25

Phe Asn Asn Ala Thr Gly Ala Lys Val Thr Ala Gln Ser Leu Lys Val

660

665

670

30

Trp His Met Ser Thr Ala Ile Asn Glu Ile Tyr Asp Pro Ala Thr Ser

675

680

685

Val Met

ES 2 583 639 T3

690

<210> 26

<211> 501

<212> PRT

5 <213> Beta vulgaris

<220>

<221> MOD_RES

10 <222> (382)..(382)

<223> Qualquer aminoácido

<400> 26

Leu Phe Tyr Gln Tyr Asn Pro Asn Gly Val Ile Trp Gly Pro Pro Val

15 1 5 10 15

Trp Gly His Ser Thr Ser Lys Asp Leu Val Asn Trp Val Pro Gln Pro

20 20 25 30

Leu Thr Met Glu Pro Glu Met Ala Ala Asn Ile Asn Gly Ser Trp Ser

25 35 40 45

Gly Ser Ala Thr Ile Leu Pro Gly Asn Lys Pro Ala Ile Leu Phe Thr

50 55 60

30 Gly Leu Asp Pro Lys Tyr Glu Gln Val Gln Val Leu Ala Tyr Pro Lys

65 70 75 80

ES 2 583 639 T3

Asp Thr Ser Asp Pro Asn Leu Lys Glu Trp Phe Leu Ala Pro Gln Asn
 85 90 95

5

Pro Val Met Phe Pro Thr Pro Gln Asn Gln Ile Asn Ala Thr Ser Phe
 100 105 110

10

Arg Asp Pro Thr Thr Ala Trp Arg Leu Pro Asp Gly Val Trp Arg Leu
 115 120 125

15

Leu Ile Gly Ser Lys Arg Gly Gln Arg Gly Leu Ser Leu Leu Phe Arg
 130 135 140

20

Ser Arg Asp Phe Val His Trp Val Gln Ala Lys His Pro Leu Tyr Ser
 145 150 155 160

25

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

30

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

ES 2 583 639 T3

	210	215	220	
	Ile Gly Tyr Met Asn Asp Ser Ser Leu Arg Tyr Asp Tyr Gly Lys Tyr			
5	225	230	235	240
	Tyr Ala Ser Lys Thr Phe Phe Asp Asp Ala Lys Lys Glu Arg Ile Leu			
10		245	250	255
	Leu Gly Trp Ala Asn Glu Ser Ser Ser Val Glu Asp Asp Ile Lys Lys			
15		260	265	270
	Gly Trp Ser Gly Ile His Thr Ile Pro Arg Lys Ile Trp Leu Asp Lys			
20		275	280	285
	Leu Gly Lys Gln Leu Ile Gln Trp Pro Ile Ala Asn Ile Glu Lys Leu			
25		290	295	300
	Arg Gln Lys Pro Val Asn Ile Tyr Arg Lys Val Leu Lys Gly Gly Ser			
30		305	310	315
	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			

ES 2 583 639 T3

	340		345		350															
	Trp	Thr	Ser	Pro	Gln	Leu	Leu	Cys	Ser	Lys	Lys	Gly	Ala	Ser	Val	Lys				
5			355					360					365							
	Gly	Gly	Leu	Gly	Pro	Phe	Gly	Leu	Leu	Thr	Leu	Ala	Ser	Xaa	Gly	Leu				
10			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				
15				405						410						415				
	Pro	Thr	Asn	Asp	Lys	Thr	Thr	Tyr	Gly	Thr	Phe	Val	Asp	Val	Asn	Pro				
				420						425						430				
20																				
	Ile	Arg	Glu	Gly	Leu	Ser	Leu	Arg	Val	Leu	Ile	Asp	His	Ser	Val	Val				
			435						440						445					
25																				
	Glu	Ser	Phe	Gly	Ala	Lys	Gly	Lys	Asn	Val	Ile	Thr	Ala	Arg	Val	Tyr				
	450						455								460					
30																				
	Pro	Thr	Leu	Ala	Ile	Asn	Glu	Lys	Ala	His	Leu	Tyr	Val	Phe	Asn	Arg				
	465					470						475				480				

ES 2 583 639 T3

Gly Thr Ser Asn Val Glu Ile Thr Gly Leu Thr Ala Trp Ser Met Lys
485 490 495

5

Lys Ala Asn Ile Ala
500

10

<210> 27
<211> 518
<212> PRT
<213> Bifidobacterium breve

15

<400> 27
Met Thr Asp Phe Thr Pro Glu Thr Pro Val Leu Thr Pro Ile Arg Asp
1 5 10 15

20

His Ala Ala Glu Leu Ala Lys Ala Glu Ala Gly Val Ala Glu Met Ala
20 25 30

25

Ala Lys Arg Asn Asn Arg Trp Tyr Pro Lys Tyr His Ile Ala Ser Asn
35 40 45

30

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

ES 2 583 639 T3

	65		70		75		80
	Met His Trp Gly His Val Ser Ser Thr Asp Met Leu Asn Trp Lys Arg						
5		85		90		95	
	Glu Pro Ile Met Phe Ala Pro Ser Leu Glu Gln Glu Lys Asp Gly Val						
10		100		105		110	
	Phe Ser Gly Ser Ala Val Ile Asp Asp Asn Gly Asp Leu Arg Phe Tyr						
15		115		120		125	
	Tyr Thr Gly His Arg Trp Ala Asn Gly His Asp Asn Thr Gly Gly Asp						
20		130		135		140	
	Trp Gln Val Gln Met Thr Ala Leu Pro Asp Asn Asp Glu Leu Thr Ser						
		145		150		155	
	Ala Thr Lys Gln Gly Met Ile Ile Asp Cys Pro Thr Asp Lys Val Asp						
25		165		170		175	
	His His Tyr Arg Asp Pro Lys Val Trp Lys Thr Gly Asp Thr Trp Tyr						
30		180		185		190	
	Met Thr Phe Gly Val Ser Ser Glu Asp Lys Arg Gly Gln Met Trp Leu						

ES 2 583 639 T3

195 200 205

Phe Ser Ser Lys Asp Met Val Arg Trp Glu Tyr Glu Arg Val Leu Phe
5 210 215 220

Gln His Pro Asp Pro Asp Val Phe Met Leu Glu Cys Pro Asp Phe Phe
10 225 230 235 240

Pro Ile Lys Asp Lys Asp Gly Asn Glu Lys Trp Val Ile Gly Phe Ser
15 245 250 255

Ala Met Gly Ser Lys Pro Ser Gly Phe Met Asn Arg Asn Val Asn Asn
20 260 265 270

Ala Gly Tyr Met Ile Gly Thr Trp Glu Pro Gly Gly Glu Phe Lys Pro
25 275 280 285

Glu Thr Glu Phe Arg Leu Trp Asp Cys Gly His Asn Tyr Tyr Ala Pro
30 290 295 300

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

ES 2 583 639 T3

Leu Thr Leu Pro Arg Glu Ile Thr Leu Asp Asp Asp Gly Asp Val Val
340 345 350

5

Thr Ala Pro Val Ala Glu Met Glu Gly Leu Arg Glu Asp Thr Leu Asp
355 360 365

10

His Gly Ser Ile Thr Leu Asp Met Asp Gly Glu Gln Val Ile Ala Asp
370 375 380

15

Asp Ala Glu Ala Val Glu Ile Glu Met Thr Ile Asp Leu Ala Ala Ser
385 390 395 400

20

Thr Ala Asp Arg Ala Gly Leu Lys Ile His Ala Thr Glu Asp Gly Ala
405 410 415

25

Arg Gln Ala Met Ala Asn Gly Asp His Gly Tyr Arg Ala Ala Pro Leu
435 440 445

30

Thr Asp Ala Glu Leu Ala Ser Gly Lys Leu Asp Leu Arg Val Phe Val
450 455 460

ES 2 583 639 T3

Asp Arg Gly Ser Val Glu Val Tyr Val Asn Gly Gly His Gln Val Leu
465 470 475 480

5

Ser Ser Tyr Ser Tyr Ala Ser Glu Gly Pro Arg Ala Ile Lys Leu Val
485 490 495

10 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

15

515

<210> 28

<211> 532

<212> PRT

<213> *Saccharomyces cerevisiae*

20

<400> 28

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

25

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His
20 25 30

30

Phe Thr Pro Asn Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr
35 40 45

ES 2 583 639 T3

Asp Glu Lys Asp Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn
50 55 60

5

Asp Thr Val Trp Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp
65 70 75 80

10

Asp Leu Thr Asn Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg
85 90 95

15

Asn Asp Ser Gly Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn
100 105 110

20

Thr Ser Gly Phe Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val
115 120 125

25

Ala Ile Trp Thr Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser
130 135 140

30

Val Leu Ala Ala Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp
165 170 175

ES 2 583 639 T3

Tyr Glu Pro Ser Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp
180 185 190

5

Tyr Lys Ile Glu Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu
195 200 205

10

Glu Ser Ala Phe Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys
210 215 220

15

Pro Gly Leu Ile Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr
225 230 235 240

20

Trp Val Met Phe Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser
245 250 255

25

Phe Asn Gln Tyr Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala
260 265 270

30

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

ES 2 583 639 T3

Ile Ala Trp Ala Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn
 305 310 315 320

5

Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr
 325 330 335

10

Glu Tyr Gln Ala Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu
 340 345 350

15

Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr
 355 360 365

20

Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn
 370 375 380

25

Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr
 385 390 395 400

30

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

ES 2 583 639 T3

435

440

445

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

5

450

455

460

Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu

10

465

470

475

480

Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser

15

485

490

495

Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn

20

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

25

530

<210> 29

<211> 512

30

<212> PRT

<213> *Zymomonas mobilis*

ES 2 583 639 T3

<400> 29

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

5

Lys Lys Ala Gly Lys Arg Leu Leu Ser Ser Glu Trp Tyr Pro Gly Phe
 20 25 30

10

His Val Thr Pro Leu Thr Gly Trp Met Asn Asp Pro Asn Gly Leu Ile
 35 40 45

15

Phe Phe Lys Gly Glu Tyr His Leu Phe Tyr Gln Tyr Tyr Pro Phe Ala
 50 55 60

20

Pro Val Trp Gly Pro Met His Trp Gly His Ala Lys Ser Arg Asp Leu
65 70 75 80

25

Val His Trp Glu Thr Leu Pro Val Ala Leu Ala Pro Gly Asp Leu Phe
 85 90 95

30

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

ES 2 583 639 T3

Asp Ala Ile Arg Glu Val Gln Cys Met Ala Thr Ser Ile Asp Gly Ile
130 135 140

5

His Phe Gln Lys Glu Gly Ile Val Leu Glu Lys Ala Pro Met Pro Gln
145 150 155 160

10

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

15

Phe Met Val Val Gly Tyr Arg Thr Asp Asp Glu Lys His Gln Gly Ile
180 185 190

20

Gly His Val Ala Leu Tyr Arg Ser Glu Asn Leu Lys Asp Trp Ile Phe
195 200 205

25

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

30

Val Leu Met Phe Ser Pro Gln Gly Leu Lys Ala Ser Gly Tyr Lys Asn
245 250 255

ES 2 583 639 T3

Arg Asn Leu Phe Gln Asn Gly Tyr Ile Leu Gly Lys Trp Gln Ala Pro
260 265 270

5

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

10

Ile Ala Trp Phe Asp Met Trp Glu Asn Gln Lys Pro Ser Gln Arg Asp
305 310 315 320

15

Gly Trp Ala Gly Cys Met Thr Leu Pro Arg Lys Leu Asp Leu Ile Asp
325 330 335

20

Asn Lys Ile Val Met Thr Pro Val Arg Glu Met Glu Ile Leu Arg Gln
340 345 350

25

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

30

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

ES 2 583 639 T3

	385		390				395			400						
	Lys	Gly	Gln	Glu	Thr	Leu	Leu	Tyr	Ile	Asp	Arg	Ser	Gln	Asn	Arg	Ile
5						405				410						415
	Ile	Leu	Asp	Arg	Asn	Arg	Ser	Gly	Gln	Asn	Val	Lys	Gly	Ile	Arg	Ser
10						420				425						430
	Cys	Pro	Leu	Pro	Asn	Thr	Ser	Lys	Val	Arg	Leu	His	Ile	Phe	Leu	Asp
15						435				440						445
	Arg	Ser	Ser	Ile	Glu	Ile	Phe	Val	Gly	Asp	Asp	Gln	Thr	Gln	Gly	Leu
						450				455						460
20	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
25						485				490						495
	Trp	Thr	Leu	Gln	Asp	Ala	Asn	Leu	Ala	Ala	Phe	Ser	Ser	Asp	Ala	Cys
30						500				505						510
	<210>	30														
	<211>	987														
	<212>	ADN														

ES 2 583 639 T3

<213> Chlorella luteoviridis

<400> 30

	tgttgaagaa tgagccggcg acttatagga agtggcttgg ttaaggatac tttccgaagc	60
5		
	ctaagcgaag gcaagttgta acaatagcga tatacctctt tgtaggtcag tcacttctta	120
	tggacccgaa cccgggtgat ctaaccatga ccaggatgaa gcttgggtaa caccaagtga	180
10		
	aggtccgaac tcttcgatct ttaaaaatcg tgagatgagt tatggttagg ggtaaactcg	240
	gcagttttgc cccgaaaag ggtaaccttt tgtaattact gactcataac ggtgaagcct	300
	aaggcgtag ctatggtaat accgtgggaa gtttcaatac cttcttgcat attttttatt	360
15		
	tgcaccttta gtgcaaacag tgtaaagaaa gcgttttgaa acccctaac gactaatfff	420
	ttgcttttgc aagaacgtca gcactcacca atacactttc cgtttttttc ttttattaat	480
20		
	taaagcaaca taaaaatata ttttatagct ttaatcataa aactatgta gcacttcgtg	540
	ctaattgtgct aatgtgctaa tcaaatgaaa agtgttctta aaagtgagtt gaaggtagag	600
	tctaactcttg cctgaaaggg caagctgcac attttttttt gaatgtgcaa caatggaaat	660
25		
	gccaatcgaa ctcggagcta gctggttctc cccgaaatgt gttgaggcgc agcgattcat	720
	gattagtacg gtgtaggggt aaagcactgt ttcggtgcgg gctgtgaaaa cggtagcaaa	780
30		
	tcgtggcaaa ctaagaatac tacgcttgta taccatggat cagtgagact atgggggata	840
	agctccatag tcaagagggg aacagcccag atcaccagtt aaggcccaa aatgacagct	900

ES 2 583 639 T3

	aagtggcaaa ggaggtgaaa gtgcagaaac aaccaggagg tttgccaga agcagccatc	960
	ctttaaagag tgcgtaatag ctactg	987
5		
	<210> 31	
	<211> 1412	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
15		
	<400> 31	
	gaattcgagt ttaggtccag cgtccgtggg gggggacggg ctgggagctt gggccgggaa	60
	gggcaagacg atgcagtccc tctggggagt cacagccgac tgtgtgtgtt gcaactgtgcg	120
20		
	gcccgcagca ctcacacgca aaatgcctgg ccgacaggca ggcctgtcc agtgcaacat	180
	ccacggtccc tctcatcagg ctacacttgc tcattgacat aacggaatgc gtaccgctct	240
25		
	ttcagatctg tccatccaga gaggggagca ggctccccac cgacgctgtc aaacttgctt	300
	cctgccaac cgaaaacatt attgtttgag gggggggggg ggggggcaga ttgcatggcg	360
	ggatatctcg tgaggaacat cactgggaca ctgtggaaca cagtgagtgc agtatgcaga	420
30		
	gcatgtatgc taggggtcag cgcaggaagg gggcctttcc cagtctcca tgccactgca	480

ES 2 583 639 T3

	ccgatccac gactcaccag gaccagcttc ttgatcggct tccgctcccg tggacaccag	540
	tgtgtagcct ctggactcca ggtatgcgtg caccgcaaag gccagccgat cgtgccgatt	600
5	cctggggtgg aggatatgag tcagccaact tggggctcag agtgcacact ggggcacgat	660
	acgaacaac atctacaccg tgcctccat gctgacacac cacagcttcg ctccacctga	720
	atgtgggctc atgggcccga atcacagcca atgtcgtcgc tgccataatg tgatccagac	780
10	cctctccgcc cagatgccga gcggatcgtg ggcgctgaat agattcctgt ttcgatcact	840
	gtttgggtcc tttccttttc gtctcggatg cgcgtctcga aacaggctgc gtcgggcttt	900
15	cggatccctt ttgctccctc cgtcaccatc ctgcgcgcgg gcaagttgct tgaccctggg	960
	ctggtaccag ggttggaggg tattaccgcg tcaggccatt cccagcccgg attcaattca	1020
	aagtctgggc caccaccctc cgccgctctg tctgatcact ccacattcgt gcatacacta	1080
20	cgttcaagtc ctgatccagg cgtgtctcgg gacaaggtgt gcttgagttt gaatctcaag	1140
	gaccactcc agcacagctg ctggttgacc ccgccctcgc aactccctac catgtctgct	1200
25	ggtaggtcca gggatctttg ccatgcacac aggaccccgt ttgtgggggt ccccggtgca	1260
	tgctgtcgtc gtgcaggcgc cgggtgtgggg cctgggcccc gcgggagctc aactcctccc	1320
	catatgcctg ccgtccctcc caccaccgc gacctggccc cttttgcaga ggaaggcgaa	1380
30	gtcagcgcca tcgtgtgcga taatggatcc gg	1412

ES 2 583 639 T3

<210> 32

<211> 1627

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

10

<400> 32

gaattcgccc ttgagtttag gtccagcgtc cgtggggggg gcgtgagact cccccctgac 60

cttcgtatgg cagggactcc tacttgccaa gtaatcagtt gacaatgcca cttcaatgct 120

15

cgttgtggta cactgacgcg ggtctaacat actgggaagc atgaattgcc gacatggact 180

cagttggaga cagtaacagc tctttgtggt ctatcttcag gaacacattt ggcagcgcac 240

20

ccatacagtg gcgcacacgc agctgtacct gatgtggctc tattcccaca tgtttcaact 300

tgatccaaaa gtcactcaga ctctcagcag ctagacttga tcgcatcttt ggccatgaag 360

atgcttgctc aactctagga atgggacgag aaaagagcct gctctgatcg gatatttcca 420

25

ttctctggat gggactgaga tgattctgaa gaaatgctgc tcgacttatt tggaagaaca 480

gcacctgacg catgctttga ggctgctgtg gctgggatgt gctgtatttg tcagcattga 540

30

gcatctacgg gtagatggcc ataaccacgc gctgcctatc atgcggtggg ttgtgtggaa 600

aacgtacaat ggacagaaat caatcccatt gcgagcctag cgtgcagcca tgcgctcct 660

ES 2 583 639 T3

	ctgtagcccc gctccaagac aaagccagcc aatgccagaa cccacataga gagggatatct	720
	tcctaatagac ctcgcccatac atttcctcca aattaactat aatgccttga ttgtggagtt	780
5	ggctttggct tgcagctgct cgcgctggca cttttgtagg cagcacaggg tatgccagcg	840
	ccgaactttg tgccttgag caggccacaa gggcacaaga ctacaccatg cagctggtat	900
10	acttggaaact gataccattc ttaccaagca aggcacagca cagcctgcac cgactcactt	960
	tgcttgagcg gggcacagcg ccgcgactga tcctgcgagc tgtggggagt tccgactgtt	1020
	ctggacctcg gtctctgaaa gatgtgtacg atgggatcaa gtcattcaag tatgctcttc	1080
15	acatgagcaa tcgggggaga cacggtggcc ctaaaggtgt tcacttgatt caagtgtagt	1140
	gggggggtgc tgtttgtccc ggggcgcccc ccgctccccg accccggaga agggccccag	1200
20	aggactcggc cgcccacaga ggaataaccg ggcgtggctc ggccctgcgc ctccctcttt	1260
	caatatttca cctggtgttc agtgcacgga cacgtaaaga actagataca atggccgagg	1320
	gaaagacggt gagagcttgg cgttggtgga ccgggcagca tcagaaactc ctcttccccg	1380
25	cccgccttga aactcactgt aactccctcc tcttccccct cgcagcatct gtctatcgtt	1440
	atcgtgagtg aaagggactg ccatgtgtcg ggtcgttgac cacggtcggc tcgggvcgctg	1500
30	ctgcccgcgt cgcgaacgtt ccctgcaaac gccgcgcagc cgtccctttt tctgccgccg	1560
	ccccaccccc tcgctcccc cttcaatcac accgcagtgc ggacatgtcg attccggcaa	1620

ES 2 583 639 T3

	gtccacc	1627
5	<210> 33 <211> 570 <212> ADN <213> Chlorella protothecoides	
10	<400> 33 gaattccctg caggaagaag gccggcagca gctggtactt gtccttcacc tccttgatcg	60
	gctgggtgag cttggccggg tcgcagtcgt cgatgccggc atcgcccagc acgctgtgcg	120
15	gggagccggc atcgacaacc ttggcactgc tcaccttggc caccggcatg gggcatggc	180
	gctgcagacc agcggcctgt cagcatgctg caggcatctg tgttttgtag tagatacttt	240
	ctgatgcatc accacacggt tggaaggctc ccaagcccct tcaacagtct cgacatatga	300
20	cactcgcgcc ctcttcctcg tcccgtggcc tgatgagggt acgcaggtac cgcagctgcg	360
	ccccgtcccg ccagttgccc tggccccgcc gggcccaatc tgttcattgc cgctccctgg	420
25	cagccgtgaa cttcacacta ccgctctctg tgaccttcag cacagcagga atcgccattt	480
	caccggcggc cgttgctgcg gagcctcagc tgatctcgcc tgcgagacct cacagtttga	540
	atctgcggtc cccacacaac ctctgacgcc	570
30	<210> 34 <211> 568	

ES 2 583 639 T3

<212> ADN

<213> Chlorella protothecoides

5 <220>

<221> base modificada

<222> (133)..(133)

<223> a, c, t, g, desconocido u otro

10 <400> 34

gaattccctc aggaagaagg ccggcagcag ctggtacttg tccttcacct ccttgatcgg 60

ctgggtgagc ttcgcaggat cgcagtcgtc gatgccggca tcgcccagca cgctgtgcgg 120

15 ggagccggca tcnacaacct tggcactgct ccccttggtc accggcatgg ggtcatggcg 180

ctgcagccca gcggcctgtc agcatgctgc aggcattctgt gtattgtagt aggtacttcc 240

tgatgcatca acacacgttt ggaagctccc caagcccctt caacagtctc gacgtatgac 300

20 actcgcgcc tcttctcgc cccgtggcct gatgagggta cgcaggtacc acagctgcgc 360

cccgtcccgc cagttgccct ggcccggccg ggcccaatct gttcattgcc gtcacctggt 420

25 agccgtgaac tcacattacc gctctctgtg accttcagca cagcaggaat cgccatttca 480

ccggcggctc ttgctgcgga gcctcagctg atctcgcctg cgagacccca cagtttgaat 540

ttgcggtccc cacacaacct ctgacgcc 568

30 <210> 35

<211> 50

<212> ADN

<213> Secuencia Artificial

5

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 35

10 tgacctaggt gattaattaa ctcgaggcag cagcagctcg gatagtatcg

50

<210> 36

<211> 45

15

<212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

20

<400> 36

ctacgagctc aagctttcca tttgtgttcc catcccacta cttcc

45

25

<210> 37

<211> 32

<212> ADN

<213> Secuencia Artificial

30

<220>

<223> Descripción de Secuencia Artificial: Cebador

sintético

<400> 37

gatcagaatt cgcctgcaa cgcaaggga gc 32

5

<210> 38

<211> 30

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Cebador

sintético

15

<400> 38

gcatactagt ggcgggacgg agagaggcg 30

20

<210> 39

<211> 1568

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido

sintético

<400> 39

30 gaattccttt cttgcgctat gacacttcca gcaaaaggta ggcgggctg cgagacggct 60

tccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120

ES 2 583 639 T3

	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	180
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
5	cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	300
	cagtcacaac ccgcaaacgg cgcgccatat caatgattga acaagatgga ttgcacgcag	360
10	gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg	420
	gctgctctga tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca	480
	agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc	540
15	tggccacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg	600
	actggctgct attgggcgaa gtgccggggc aggatctcct gtcactcac cttgctcctg	660
20	ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta	720
	cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag	780
	ccggtcttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac	840
25	tgttcgccag gctcaaggcg cgcagccccg acggcgagga tctcgtcgtg acccatggcg	900
	atgcctgctt gccgaatc atggtgaaa atggccgctt ttctggattc atcgactgtg	960
30	gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctaccctg gatattgctg	1020
	aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg	1080

ES 2 583 639 T3

attcgagcg catgccttc tatgccttc ttgacgagtt cttctaagat ctgtcgatcg 1140

acaagtgact cgaggcagca gcagctcgga tagtatcgac aactctgga cgctggtcgt 1200

5 gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta 1260

tcaaacagcc tcagtgtgtt tgatccttggtg tgtacgcgct tttgagagtt gctagctgct 1320

10 tgtgctatatt gcgaatacca cccccagcat ccccttcctt cgtttcatat cgcttgcatc 1380

ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctccctgctca 1440

ctgccctcg cacagccttg gtttgggctc cgctgtatt ctccctgtac tgcaacctgt 1500

15 aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca aatggaaagc 1560

ttgagctc 1568

20

<210> 40

<211> 2571

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

30

<220>

<221> base modificada

ES 2 583 639 T3

<222> (997)..(999)

<223> a, c, t, g, desconocido u otro

<400> 40

5 gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgcccgc tgcctgccgc atgcttgtgc tggtaggct gggcagtgc 120

gccatgctga ttgaggcttg gttcatcggg tggagctta tgtgtgtgct gggcttgc 180

10 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcg acaggttgag 300

15 cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgcccgc tggctgtttc agctacaacg 480

20 gcaggagtca tccaactaac catagctgat caaactgca atcatcggcg gctgatgcaa 540

gcattctgca agacacatgc tgtgcgatgc tgcgctgctg cctgctgcgc acgccgttga 600

25 gttggcagca gctcagccat gcactggatc aggctgggct gccactgcaa tgtggtggat 660

aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc 720

accatcagca ggagcgcgat cgaagggact ggcccatgc acgcatgcc aaaccggagc 780

30 gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaacctg ctcattggacg 840

ES 2 583 639 T3

catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca 900

caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggccc acccttcttg 960

5 cggggtcgcg ccccagccag cggatgatgcg ctgatacnnnc caaacgagtt cacattcatt 1020

tgcagcctgg agaagcgagg ctggggcctt tgggctgggtg cagcccgcaa tggaatgcgg 1080

gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcatt 1140

10 tggactgcatt ttgggtgggg cggccggctg tttctttcgt gttgcaaac gcgccacgtc 1200

agcaacctgt cccgtgggtc ccccgtgccg atgaaatcgt gtgcacgccg atcagctgat 1260

15 tgcccggctc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag 1320

tggcgcgcca tatcaatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg 1380

gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc 1440

20 gtgttccggc tgtcagcga ggggcgccc gttctttttg tcaagaccga cctgtccggt 1500

gccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac gacgggcggt 1560

25 ccttgccgag ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc 1620

gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc 1680

atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgccc attcgaccac 1740

30 caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccgggtct tgtcgcagc 1800

ES 2 583 639 T3

gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag 1860

gcgcgcatgc ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat 1920

5 atcatggtgg aaaatggccg cttttctgga ttcacgact gtggccggct ggggtgtggc 1980

gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa 2040

tgggctgacc gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcatcgcc 2100

10 ttctatcgcc ttcttgacga gttcttctaa gatctgtcga tcgacaagtg actcgaggca 2160

gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 2220

15 ccacacttgc tgccttgacc tgtgaatata cctgccgctt ttatcaaaca gcctcagtgt 2280

gtttgatctt gtgtgtacgc gcttttgca gttgctagct gcttgtgcta tttgcaata 2340

ccacccccag catcccctc cctcgtttca tatcgcttgc atcccaaccg caacttatct 2400

20 acgctgtcct gctatccctc agcgctgctc ctgctcctgc tcaactgccc tcgcacagcc 2460

ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 2520

25 ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agcttgagct c 2571

<210> 41

<211> 2550

30 <212> ADN

<213> Secuencia Artificial

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5

<220>

<221> base modificada

<222> (997)..(999)

<223> a, c, t, g, desconocido u otro

10

<400> 41

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgct 120

15

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240

20

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcggt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

25

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaactaac catagctgat caaactgca atcatcggcg gctgatgcaa 540

30

gcatcctgca agacacatgc tgtgcatgct tgcgctgctg cctgctgctg acgccgttga 600

gttggcagca gctcagccat gcaactggatc aggctgggct gccactgcaa tgtggtggat 660

ES 2 583 639 T3

	aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc	720
	accatcagca ggagcgcgatg cgaagggact ggccccatgc acgcatgcc aaaccggagc	780
5	gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggacg	840
	catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca	900
10	caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggccc acccttcttg	960
	cggggtcgcg ccccagccag cggatgatgcg ctgacnnc caaacgagtt cacattcatt	1020
	tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg	1080
15	gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcac	1140
	tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaac gcgccacgtc	1200
20	agcaacctgt cccgtgggtc ccccgtgccg atgaaatcgt gtgcacgccg atcagctgat	1260
	tgcccggctc gcgaagtagg cgccctcttt ctgctcgcct tctctccgtc ccgccactag	1320
	tggcgcgcca tatcaatgat cgagcaggac ggcctccacg ccggctcccc cgccgcctgg	1380
25	gtggagcgcc tgttcggcta cgactgggcc cagcagacca tcggctgctc cgacgccgcc	1440
	gtgttccgcc tgtccgcca gggccgcccc gtgctgttcg tgaagaccga cctgtccggc	1500
30	gccctgaacg agctgcagga cgaggccgcc cgctgtcct ggctggccac caccggcgtg	1560
	ccctgcgccg ccgtgctgga cgtggtgacc gaggccggcc gcgactggct gctgctgggc	1620

ES 2 583 639 T3

gaggtgcccg gccaggacct gctgtcctcc cacctggccc ccgccgagaa ggtgtccatc 1680

atggccgacg ccatgcgccg cctgcacacc ctggaccccg ccacctgccc cttcgaccac 1740

5 caggccaagc accgcatcga gcgcgcccgc acccgcatgg aggccggcct ggtggaccag 1800

gacgacctgg acgaggagca ccagggcctg gccccgccg agctgttcgc ccgcctgaag 1860

10 gcccgcgatgc ccgacggcga ggacctggtg gtgaccacg gcgacgcctg cctgcccac 1920

atcatggtgg agaacggccg cttctccggc ttcacgact gcggccgcct gggcgtggcc 1980

gaccgctacc aggacatcgc cctggccacc cgcgacatcg ccgaggagct gggcggcgag 2040

15 tgggccgacc gcttcctggt gctgtacggc atcgcccgc ccgactcca gcgcatcgcc 2100

ttctaccgcc tgctggacga gttcttctga ctcgaggcag cagcagctcg gatagtatcg 2160

20 acacactctg gacgctggtc gtgtgatgga ctggtgccgc cacacttgct gccttgacct 2220

gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg 2280

cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cacccccagc atccccttcc 2340

25 ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca 2400

gcgctgctcc tgctcctgct cactgccct cgcacagcct tggtttgggc tccgcctgta 2460

30 ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg 2520

gatgggaaca caaatggaaa gcttgagctc 2550

ES 2 583 639 T3

<210> 42
 <211> 1547
 5 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Descripción de Secuencia Artificial: Polinucleótido
 10 sintético

<400> 42
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 15 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 20 cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacgg cgcgcatat caatgatcga gcaggacggc ctccacgccg 360
 25 gctccccgc cgcctgggtg gagcgcctgt tcggctacga ctgggcccag cagaccatcg 420
 gctgctccga cgccgccgtg ttccgcctgt ccgccaggg ccgccccgtg ctgttcgtga 480
 agaccgacct gtccggcgcc ctgaacgagc tgcaggacga ggccgcccgc ctgtcctggc 540
 30 tggccaccac cggcgtgccc tgcgccgccg tgctggacgt ggtgaccgag gccggccgcg 600

ES 2 583 639 T3

actggctgct gctgggagag gtgcccggcc aggacctgct gtcctccac ctggcccccg 660

ccgagaaggt gtccatcatg gccgacgcca tgcgccgcct gcacaccctg gaccccgcca 720

5 cctgcccctt cgaccaccag gccaagcacc gcatcgagcg cgcccgcacc cgcattggagg 780

ccggcctggt ggaccaggac gacctggacg aggagcacca gggcctggcc cccgccgagc 840

tgttcgcccg cctgaaggcc cgcattgccg acggcgagga cctggtggtg acccacggcg 900

10 acgcctgcct gcccaacatc atggtggaga acggccgctt ctccggcttc atcgactgcg 960

gccgcctggg cgtggccgac cgctaccagg acatcgccct ggccaccgc gacattgccg 1020

15 aggagctggg cggcgagtgg gccgaccgct tcctggtgct gtacggcatc gccgccccg 1080

actcccagcg catcgccttc taccgcctgc tggacgagtt cttctgactc gaggcagcag 1140

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 1200

20 acttgctgcc ttgacctgag aatatccctg ccgcttttat caaacagcct cagtgtgttt 1260

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 1320

25 ccccagcatc cccttcctc gtttcatatc gcttgcattc caaccgcaac ttattctacg 1380

tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1440

tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 1500

30 tgcacgggaa gtagtgggat gggaaacaaa atggaaagct tgagctc 1547

<210> 43
<211> 21
<212> ADN
5 <213> Secuencia Artificial

<220>
<223> Descripción de Secuencia Artificial: Cebador
sintético
10

<400> 43
gccgcgactg gctgctgctg g 21

15 <210> 44
<211> 21
<212> ADN
<213> Secuencia Artificial

20 <220>
<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 44
25 aggtcctcgc cgtcgggcat g 21

<210> 45
<211> 1292
30 <212> ADN
<213> Secuencia Artificial

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5 <400> 45
atcaaaggca tagattcaca tttgttggca ttgcagagca atcatcgcg aggacgaaca 60
tcgctcacca agcacgtact gggcatccgg aggcctccgc aaattcctgc aacaggactc 120
10 gctgatcagt tcgccaagg tctacgacgc tccctatcgg cgctagactt caacacatat 180
ttcactgtca cagcctcggc atgcatcagg cctcagtctc caccatgaag accatccagt 240
ctcggcacgc cgggccatc ggacatgtgc agtcgggctc cggatcggcg gggcgcgcg 300
15 gatcccgcat ggcgacccc gtggccgag ctaccgtcgc agcccctcgc tcggccctca 360
acctctccc caccatcatt cgacaggagg tgctccactc cgccagcgcc cagcaactag 420
20 actgcgtagc ctccctggcg cccgtcttcg agtcccagat cctccccctc ctgacgccc 480
tggacgagat gtggcagccc accgacttcc tccccgctc gaactcggag gcattcttcg 540
accagatcgg cgacctgagg ggcgatcgg cggccatccc cgacgacctg ctggtctgcc 600
25 tgggtggggga catgatcacg gaggaggccc tgcccaccta catggccatg ctgaacaccc 660
tggacgtcgt ggcgatgag acagggcaca gccagcacc ctacgccaag tggaccagg 720
30 cttggatcgc ggaggagaac cgccatggcg acctgctgaa caagtacatg tggctgacgg 780
ggcgggtggg acatgctggc ggtggagcgc accatccagc catgctggcg gtggagcgca 840

ES 2 583 639 T3

ccatccagcg cctcatctca tcgggcatgg acccgggcac ggagaaccac ccctaccag 900

cctttgtggt caccagcttc caggagcgcg ccaccaagct gagccacggc tccaccgccc 960

5 gcctggcggg cgccgccggg gacgaggccc tggccaagat ctgcccggacc attgcccggg 1020

acgagtcgcg ccacgaggcg gcgtacacgc ggaccatgga tgccatcttc cagcgcgacc 1080

10 ccagcggggc catggtggcg tttgcgaca tgatgatgcg caagatcacc atgcccgccc 1140

acctcatgga cgacggccag cacggcgcgc gcaacggggg ggcgcaactt gttcgacgac 1200

tttgcggcag tggcggagcg ggcaggggtg tacaccgccg gcgactacat cggcatcctg 1260

15 cgccacctca tccggcgctg ggacgtggag gg 1292

<210> 46

20 <211> 364

<212> PRT

<213> Secuencia Artificial

<220>

25 <223> Descripción de Secuencia Artificial: Polipéptido
sintético

<400> 46

Met His Gln Ala Ser Val Ser Thr Met Lys Thr Ile Gln Ser Arg His

30 1 5 10 15

ES 2 583 639 T3

Ala Gly Pro Ile Gly His Val Gln Ser Gly Arg Arg Ser Ala Gly Arg
20 25 30

5 Ala Gly Ser Arg Met Ala Thr Pro Val Ala Ala Ala Thr Val Ala Ala
35 40 45

10 Pro Arg Ser Ala Leu Asn Leu Ser Pro Thr Ile Ile Arg Gln Glu Val
50 55 60

15 Leu His Ser Ala Ser Ala Gln Gln Leu Asp Cys Val Ala Ser Leu Ala
65 70 75 80

20 Pro Val Phe Glu Ser Gln Ile Leu Pro Leu Leu Thr Pro Val Asp Glu
85 90 95

25 Met Trp Gln Pro Thr Asp Phe Leu Pro Ala Ser Asn Ser Glu Ala Phe
100 105 110

30 Phe Asp Gln Ile Gly Asp Leu Arg Ala Arg Ser Ala Ala Ile Pro Asp
115 120 125

35 Asp Leu Leu Val Cys Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu
130 135 140

40 Pro Thr Tyr Met Ala Met Leu Asn Thr Leu Asp Val Val Arg Asp Glu

ES 2 583 639 T3

	145		150		155		160									
	Thr	Gly	His	Ser	Gln	His	Pro	Tyr	Ala	Lys	Trp	Thr	Arg	Ala	Trp	Ile
5					165					170					175	
	Ala	Glu	Glu	Asn	Arg	His	Gly	Asp	Leu	Leu	Asn	Lys	Tyr	Met	Trp	Leu
10				180					185					190		
	Thr	Gly	Arg	Val	Gly	His	Ala	Gly	Gly	Gly	Ala	His	His	Pro	Ala	Met
15				195					200					205		
	Leu	Ala	Val	Glu	Arg	Thr	Ile	Gln	Arg	Leu	Ile	Ser	Ser	Gly	Met	Asp
20				210					215					220		
	Pro	Gly	Thr	Glu	Asn	His	Pro	Tyr	His	Ala	Phe	Val	Phe	Thr	Ser	Phe
25				225					230					235		240
	Gln	Glu	Arg	Ala	Thr	Lys	Leu	Ser	His	Gly	Ser	Thr	Ala	Arg	Leu	Ala
30				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		

ES 2 583 639 T3

Ile Phe Gln Arg Asp Pro Ser Gly Ala Met Val Ala Phe Ala His Met
290 295 300

5

Met Met Arg Lys Ile Thr Met Pro Ala His Leu Met Asp Asp Gly Gln
305 310 315 320

10

His Gly Ala Arg Asn Gly Gly Ala Gln Leu Val Arg Arg Leu Cys Gly
325 330 335

15

Ser Gly Gly Ala Gly Arg Gly Val His Arg Arg Arg Leu His Arg His
340 345 350

20

Pro Ala Pro Pro His Pro Ala Leu Gly Arg Gly Gly
355 360

25

<210> 47

<211> 1395

<212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

30

<400> 47

attatacatc ggcacgtct caggtttcac gatctgcatg ctatctatgg gactgtgact

60

ES 2 583 639 T3

ccgccggcca ggttggtgtg cgcgagaatc ctccccgctc ctgccttctc atttcctga 120
 cgggagtcgc cgctgagcac cgggcggatc atgggcgtcg gcacactcca aaccccatat 180
 5
 acatgtggtc gtgcattcac gcatagcgca cggtatgtcc cgcgacgcgc ggctcgaagc 240
 cgtggccatc cgacgcgctg cacggccgag gtgagggcac gccctccgc caatggcgcg 300
 10
 cagcccatga ccgccttca cttccggcag tacatgcagc agcgcgccgc gctggtggac 360
 gcagcgctgg acctggcagt gccgctgcag taccgccaga agatcaacga ggccatgcgg 420
 tacagcctgc tggccggggg caagcgcgtg cgccccgcgc tctgcctcgc tgcctgcgag 480
 15
 ctcgtgggcg gccccctgga ggcggccatg cccgccgcct gcgccatgga gatgatccac 540
 accatgagcc tcatccacga cgacctcccc gccatggaca acgacgactt ccggcgcggc 600
 20
 cagcccgcca accacaaggc ctatggcgag gagattgcga tcctggcggg cgacgcgctg 660
 ctgtcgctga gctttgagca catcgcgcgc gagacgcgag gcgtggacc ggtgcgcgtc 720
 ctggccgcca tctcggagtg gcgcgcggtg ggcagccgcg ggctggtggc ggggcaggtg 780
 25
 gtggacctgg gtttcgaggg cggcggcgtg gggctggccc cgctgcgcta catccacgag 840
 caaaaaccg cggcgtgct ggaggcggcg gtggtgtccg gcgcgctgct gggcggcgcg 900
 30
 gaggaggcgg acctggagcg cctgcgcacc tacaaccgcg ccatcggcct cgctttccag 960
 gtggtggggg acatcctgga catcccgggg accagcgagg agctgggcaa gaccgcgggc 1020

ES 2 583 639 T3

aaggacctga gctcccccaa aaccccctac ccgtccttgg tggggctggc caggtccaaa 1080

5
 aaaattgcgg acgaactgat tgaggacgcg aaaaccaac tcaccagta cgagccggcc 1140

cgagcggcgc ccctcgtaac cctggccgaa aacatttgaa accggaagaa ctgactgggg 1200

gcccccttg cccccagata cggcggggct cctccatcca gttttgggat gggaggagcg 1260

10
 acaaccgacc ccgtaaccct gtgacgcggt tgccttgcac acgtacgcat gccttgaaac 1320

ccatccatga ccctcaaca tacctgggtg tgtgtagctt ggtcctgaaa aaaaaaaaaa 1380

aaaaaaaaaa aaaaa 1395

15

<210> 48

<211> 342

<212> PRT

20

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polipéptido
 sintético

25

<400> 48

Met Gly Val Gly Thr Leu Gln Thr Pro Tyr Thr Cys Gly Arg Ala Phe

1

5

10

15

30

Thr His Ser Ala Arg Tyr Val Pro Arg Arg Ala Ala Arg Ser Arg Gly

20

25

30

ES 2 583 639 T3

His Pro Thr Arg Cys Thr Ala Glu Val Arg Ala Arg Pro Ser Ala Asn
35 40 45

5

Gly Ala Gln Pro Met Thr Ala Phe Asp Phe Arg Gln Tyr Met Gln Gln
50 55 60

10

Arg Ala Ala Leu Val Asp Ala Ala Leu Asp Leu Ala Val Pro Leu Gln
65 70 75 80

15

Tyr Pro Glu Lys Ile Asn Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly
85 90 95

20

Gly Lys Arg Val Arg Pro Ala Leu Cys Leu Ala Ala Cys Glu Leu Val
100 105 110

25

Gly Gly Pro Leu Glu Ala Ala Met Pro Ala Ala Cys Ala Met Glu Met
115 120 125

30

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

ES 2 583 639 T3

Glu Ile Ala Ile Leu Ala Gly Asp Ala Leu Leu Ser Leu Ser Phe Glu
165 170 175

5

His Ile Ala Arg Glu Thr Arg Gly Val Asp Pro Val Arg Val Leu Ala
180 185 190

10

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

15

Gln Val Val Asp Leu Gly Phe Glu Gly Gly Gly Val Gly Leu Ala Pro
210 215 220

20

Leu Arg Tyr Ile His Glu His Lys Thr Ala Ala Leu Leu Glu Ala Ala
225 230 235 240

25

Val Val Ser Gly Ala Leu Leu Gly Gly Ala Glu Glu Ala Asp Leu Glu
245 250 255

30

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

ES 2 583 639 T3

Ala Gly Lys Asp Leu Ser Ser Pro Lys Thr Pro Tyr Pro Ser Leu Val
290 295 300

5

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

10

Lys Thr Gln Leu Thr Gln Tyr Glu Pro Ala Arg Ala Ala Pro Leu Val
325 330 335

15

Thr Leu Ala Glu Asn Ile
340

<210> 49

20

<211> 833

<212> ADN

<213> Secuencia Artificial

<220>

25

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 49

cagatgccat ggcacctcgg gccgcgggcc tgagggtcca cgcagcgtcc tcggtggccc 60

30

agacgcacca ggccgcccc ccggcgga ga ggaggttcga cgactaccag ccccgaccg 120

ES 2 583 639 T3

ccatcctctt ccccggccag ggcgcgcaga gcgtgggcat ggcgggagag ctggcgaagg 180

ccgtccccgc cgccgcggcg ctgttcgacg ccgcctccga ccagctcggc tatgacctgc 240

5 tccgcgtgtg cgttgagggc cccaaggcgc gcctggacag caccgccgtc agccagcccc 300

ccatctacgt ggccagcctg gcggcggtgg agaagctgcg cgcggagggc ggggaggagg 360

cactggccgc catcgacgtc gctgccggtc tgtccttggg cgagtacacc gcgctggcct 420

10 ttgccggcgc cttctccttc gccgacgggc tgcgcctggt ggccctgcgc ggcgccagca 480

tgcaggccgc cgccgacgcc gcaccctcgg gcatggtctc cgatcatcggc ctgccctccg 540

15 acgcggtggc cgcgctgtgc gaggccgcca acgcgcaggt ggcccccgac caggccgtgc 600

gcatcgccaa ctacctctgc gacggcaact acgccgtcag cgggtgggctg gagggctgcg 660

cggcggtgga gggcctggcc aaggcccaca aggcgcgcat gacggtgcgc ctggcggtgg 720

20 cgggcgcctt ccacaccccc ttcattgcagc cggcggtgga ggcgctgagc gcgggcgctg 780

gcggacacgc cgctggtcgc gccgcgcatc cccgtggtca gcaacgggac gcc 833

25

<210> 50

<211> 275

<212> PRT

<213> Secuencia Artificial

30

<220>

<223> Descripción de Secuencia Artificial: Polipéptido

ES 2 583 639 T3

sintético

<400> 50

Met Arg Pro Arg Ala Ala Gly Leu Arg Val His Ala Ala Ser Ser Val
5 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

ES 2 583 639 T3

115

120

125

Ser Leu Gly Glu Tyr Thr Ala Leu Ala Phe Ala Gly Ala Phe Ser Phe
5 130 135 140

Ala Asp Gly Leu Arg Leu Val Ala Leu Arg Gly Ala Ser Met Gln Ala
10 145 150 155 160

Ala Ala Asp Ala Ala Pro Ser Gly Met Val Ser Val Ile Gly Leu Pro
15 165 170 175

Ser Asp Ala Val Ala Ala Leu Cys Glu Ala Ala Asn Ala Gln Val Ala
20 180 185 190

Pro Asp Gln Ala Val Arg Ile Ala Asn Tyr Leu Cys Asp Gly Asn Tyr
25 195 200 205

Ala Val Ser Gly Gly Leu Glu Gly Cys Ala Ala Val Glu Gly Leu Ala
30 210 215 220

Lys Ala His Lys Ala Arg Met Thr Val Arg Leu Ala Val Ala Gly Ala
35 225 230 235 240

Phe His Thr Pro Phe Met Gln Pro Ala Val Glu Ala Leu Ser Ala Gly
40 245 250 255

ES 2 583 639 T3

Ala Gly Gly His Ala Ala Gly Arg Ala Ala His Pro Arg Gly Gln Gln

260

265

270

5

Arg Asp Ala

275

10

<210> 51

<211> 787

<212> ADN

<213> Secuencia Artificial

15

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

20

<400> 51

tgtccatctc cccccaccct ccatccaacc atcgtcgacg gcatgcaggc gctgtgttct 60

caccccgct ccctcacggc gcgtgcggtta ccccatgggc gggccagccc agcacagcgg 120

25

gtgtccagcg ccggcccggc ctacaccggc ctgtcccggc acaccctggg ctgccccagc 180

acccccacc tccagtcccg cgccgcggtc cagaccgcg gtcctcctc cggtccacc 240

acgcgcatga ccaccaccgc ccagcgcaag atcaaggtag ccatcaacgg gttcggccgc 300

30

atcggccgcc agttcctgcg ctgcgtggag gggcgcgagg actcgtgct ggagatcgtg 360

ES 2 583 639 T3

gccgtgaacg actccggcgg cgtgaagcag gccagccacc tgctcaagta cgactccacc 420

atgggcacct tcaacgccga catcaagatc tcgggcgagg gcaccttctc cgtcaacggc 480

5 cgcgacatcc gcgtcgtctc ctcccgcgac cccctggccc tgcctgggg cgagctgggc 540

gtggacctgg tgatcgaggg gacgggagtg tttgtggacc gcaagggtgc cagcaagcac 600

ctgcaggcgg gggccaagaa ggtcatcatc accgcgccgg ccaagggtc cgacgtgccc 660

10 acctacgtca tgggcgtgaa cgcggaccag tactccaact ccgacgacat catctccaac 720

gcctcctgca ccaccaactg cctggcgccc tttgtcaagg tgctcaacga ccgcttcggc 780

15 atcgtga 787

<210> 52

<211> 248

20 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polipéptido
 sintético

25

<400> 52

Met Gln Ala Leu Cys Ser His Pro Ala Ser Leu Thr Ala Arg Ala Val

1

5

10

15

30

Pro His Gly Arg Ala Ser Pro Ala Gln Arg Val Ser Ser Ala Gly Pro

ES 2 583 639 T3

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

ES 2 583 639 T3

Asn Gly Arg Asp Ile Arg Val Val Ser Ser Arg Asp Pro Leu Ala Leu
145 150 155 160

5 Pro Trp Gly Glu Leu Gly Val Asp Leu Val Ile Glu Gly Thr Gly Val
165 170 175

10 Phe Val Asp Arg Lys Gly Ala Ser Lys His Leu Gln Ala Gly Ala Lys
180 185 190

15 Lys Val Ile Ile Thr Ala Pro Ala Lys Gly Ser Asp Val Pro Thr Tyr
195 200 205

20 Val Met Gly Val Asn Ala Asp Gln Tyr Ser Asn Ser Asp Asp Ile Ile
210 215 220

25 Ser Asn Ala Ser Cys Thr Thr Asn Cys Leu Ala Pro Phe Val Lys Val
225 230 235 240

30 Leu Asn Asp Arg Phe Gly Ile Val
245

<210> 53

30 <211> 860

<212> ADN

<213> Secuencia Artificial

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5

<400> 53

gatgttgaga atagtagctt gctgccttgt cgccatgcag agcgtgtgcg cgcagtcggt 60

ttcatgcaag ggggccttca cccagtcctt gcggaccccc cgatgcagca ggagccagct 120

10

cgtctgccgg gctgatggca aggccggagc cttcatcaag accgtaaaga gcggtgctgc 180

cgctctggct gcctccctcc tcctgtctgg ggggtgcgggc gactgacct ttgatgagct 240

15

gcagggcctg acctacctgc aggtgaaggc ctctggcatc gccaacacct gccccacct 300

gtctggcggc tcctccaaca tcaaggacct gaagagcggg acctactccg tcaacaagat 360

gtgcctggag cccacgtcct tcaaggtaa ggaggaggca cagttcaaga acggcgaggc 420

20

cgactttgtg cccaccaagc tcgtcacgcy tctgacctac accctggacg agatctctgg 480

ccagatgaag atcgacggca gcggcggcgt ggagttcaag gaggaggatg gcatcgacta 540

25

tgctgcagtc accgtgcagc ttccggggcgg ggagcgcgtg cccttcctct tcaccatcaa 600

ggagcttgac gccaagggga ctgccgacgg cttcaagggc gagttcaccg tgccctccta 660

ccgtgggtcc tccttcctgg accccaaggc ccgcgggcgc tccaccggct acgacaacgc 720

30

cgtggccctg cccgccgccg gcgattccga ggagttggag aaggagaaca acaagtccac 780

ES 2 583 639 T3

caaggctctg aagggggagg ccatcttctc catcgccaag gtggacgccg ggacagggga 840

ggtggcgggc atctttgagt 860

5

<210> 54

<211> 275

<212> PRT

<213> Secuencia Artificial

10

<220>

<223> Descripción de Secuencia Artificial: Polipéptido sintético

15

<400> 54

Met Gln Ser Val Cys Ala Gln Ser Val Ser Cys Lys Gly Ala Phe Thr
1 5 10 15

20

Gln Ser Leu Arg Thr Pro Arg Cys Ser Arg Ser Gln Leu Val Cys Arg
20 25 30

25

Ala Asp Gly Lys Ala Gly Ala Phe Ile Lys Thr Val Lys Ser Gly Ala
35 40 45

30

Ala Ala Leu Ala Ala Ser Leu Leu Leu Ser Gly Gly Ala Gly Ala Leu
50 55 60

ES 2 583 639 T3

	Thr	Phe	Asp	Glu	Leu	Gln	Gly	Leu	Thr	Tyr	Leu	Gln	Val	Lys	Gly	Ser
	65					70					75					80
5	Gly	Ile	Ala	Asn	Thr	Cys	Pro	Thr	Leu	Ser	Gly	Gly	Ser	Ser	Asn	Ile
						85				90						95
10	Lys	Asp	Leu	Lys	Ser	Gly	Thr	Tyr	Ser	Val	Asn	Lys	Met	Cys	Leu	Glu
				100					105							110
15	Pro	Thr	Ser	Phe	Lys	Val	Lys	Glu	Glu	Ala	Gln	Phe	Lys	Asn	Gly	Glu
				115					120							125
20	Ala	Asp	Phe	Val	Pro	Thr	Lys	Leu	Val	Thr	Arg	Leu	Thr	Tyr	Thr	Leu
				130					135							140
25	Phe	Lys	Glu	Glu	Asp	Gly	Ile	Asp	Tyr	Ala	Ala	Val	Thr	Val	Gln	Leu
							165								170	
30	Pro	Gly	Gly	Glu	Arg	Val	Pro	Phe	Leu	Phe	Thr	Ile	Lys	Glu	Leu	Asp
														180		185
																190

ES 2 583 639 T3

Ala Lys Gly Thr Ala Asp Gly Phe Lys Gly Glu Phe Thr Val Pro Ser
195 200 205

5 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

10 Leu Glu Lys Glu Asn Asn Lys Ser Thr Lys Ala Leu Lys Gly Glu Ala
245 250 255

15 Ile Phe Ser Ile Ala Lys Val Asp Ala Gly Thr Gly Glu Val Ala Gly
260 265 270

Ile Phe Glu
20 275

<210> 55

<211> 818

25 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido

30 sintético

<400> 55

ES 2 583 639 T3

ataatcggaa cccagctgca cgcacatca gtgcggcagc atgcagaccg tcgcagccag 60
 ctatggcgta ttggcgccct ccggctccag cgtgaccg ggtcgcacca gcagcaagca 120
 5 gcacttcacc accctcactc ccttttccgg cttcaggcgc ctgaatcatg tggatcgggc 180
 ggggcaggcg gggctctggga gccccagac cctgcagcag gccgtgggca aggccgtgcg 240
 ccggtcgcgg ggccgcacca ccagcgcctg gcgctgacc cgcctgatgt ttgagcgggt 300
 10 caccgagaag gccatcaagg tggctatgct cgcgcaggag gaggctcgcc gtctgggcca 360
 caacttcgtg gggacggagc aaatcctgct ggggttgatt ggggagtcca caggcatcgc 420
 15 cgccaaggtc ctcaagtcga tgggcgtcac gctgaaagat gcgctgtgag aggtcgagaa 480
 gatcatcggc cgggggagcg gctttgtggc cgtggagatc cccttcaccc cccgcgccaa 540
 gcgtgtgctg gagctgtccc tggaggaggc tcgccagctc ggccacaact acattggcac 600
 20 ggagcacatc ctgctgggcc tgctgcgca gggtagggc gtggcctccc gcgtgctgga 660
 gaccctgggc gccgacccc agaagatccg cactcaggtg gtacgcatgg tgggtgagtc 720
 25 gcaggagccc gtgggcacca cgggtggcgg agggctcacc ggctccaaca agatgccac 780
 cctggaggag tacggcacca acctgaccgc ccaggccg 818
 30 <210> 56
 <211> 259
 <212> PRT

ES 2 583 639 T3

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polipéptido
5 sintético

<400> 56

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

10

Ser Val Thr Arg Gly Ser Thr Ser Ser Lys Gln His Phe Thr Thr Leu
20 25 30

15

Thr Pro Phe Ser Gly Phe Arg Arg Leu Asn His Val Asp Arg Ala Gly
35 40 45

20

Gln Ala Gly Ser Gly Ser Pro Gln Thr Leu Gln Gln Ala Val Gly Lys
50 55 60

25

Ala Val Arg Arg Ser Arg Gly Arg Thr Thr Ser Ala Val Arg Val Thr
65 70 75 80

30

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

ES 2 583 639 T3

	100	105	110
	Glu Gln Ile Leu Leu Gly Leu Ile Gly Glu Ser Thr Gly Ile Ala Ala		
5	115	120	125
	Lys Val Leu Lys Ser Met Gly Val Thr Leu Lys Asp Ala Arg Val Glu		
10	130	135	140
	Val Glu Lys Ile Ile Gly Arg Gly Ser Gly Phe Val Ala Val Glu Ile		
15	145	150	155
	Pro Phe Thr Pro Arg Ala Lys Arg Val Leu Glu Leu Ser Leu Glu Glu		
20	165	170	175
	Ala Arg Gln Leu Gly His Asn Tyr Ile Gly Thr Glu His Ile Leu Leu		
25	180	185	190
	Gly Leu Leu Arg Glu Gly Glu Gly Val Ala Ser Arg Val Leu Glu Thr		
30	195	200	205
	Leu Gly Ala Asp Pro Gln Lys Ile Arg Thr Gln Val Val Arg Met Val		
35	210	215	220
	Gly Glu Ser Gln Glu Pro Val Gly Thr Thr Val Gly Gly Gly Ser Thr		

ES 2 583 639 T3

225

230

235

240

Gly Ser Asn Lys Met Pro Thr Leu Glu Glu Tyr Gly Thr Asn Leu Thr

5

245

250

255

Ala Gln Ala

10

<210> 57

<211> 2357

<212> ADN

15

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

20

<400> 57

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatggc 120

25

gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

30

ccactcgagc ttgtgatgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300

caaccgcaa acggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 360

ES 2 583 639 T3

	gTTTTgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgtgc	420
	actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag	480
5	atgccaagtg gcatctgtac tttcaataca acccgaacga tactgtctgg gggacgccat	540
	tgTTTTgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta	600
10	tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca	660
	atacttccgg cTTTTtcaac gataccattg acccgagaca acgctgctg gccatatgga	720
	cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca	780
15	cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc	840
	cgaaggtctt ttggtacgag ccctcgaga agtggatcat gacagcggca aagtcacagg	900
20	actacaagat cgaaatttac tcgtctgacg accttaaadc ctggaagctc gaatccgct	960
	tcgcaaacga gggctttctc ggctaccaat acgaatgcc aggctgata gaggccccaa	1020
	cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac	1080
25	cggcaggagg ttcttttaac cagtacttcg tcggaagctt taacggaact catttcgagg	1140
	catttgataa ccaatcaaga gtagttgatt ttggaagga ctactatgcc ctgcagactt	1200
30	tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg	1260
	agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgtc gtgaggaaat	1320

ES 2 583 639 T3

	tctctctcaa cactgagtac caggccaacc cggaaaccga actcataaac ctgaaagccg	1380
	aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt	1440
5	tgacgaaagc caacagctac aacgctgatc tttcgaatag caccggtaca cttgaatttg	1500
	aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct	1560
10	ccctctgggt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt	1620
	ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttggt aaggagaacc	1680
	catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt	1740
15	cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg	1800
	gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga	1860
20	acatgacgac ggggtgtggat aacctgttct acatcgacaa attccaggtg agggaagtca	1920
	agtgagatct gtcgatcgac aagctcgagg cagcagcagc tcggatagta tcgacacact	1980
	ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	2040
25	tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	2100
	gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt	2160
30	catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc	2220
	tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	2280

ES 2 583 639 T3

	gg tactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga	2340
	acacaaatgg aaagctt	2357
5		
	<210> 58	
	<211> 2335	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
15		
	<400> 58	
	ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg	60
	cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc	120
20		
	gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggccccga ttgcaaagac	180
	attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg	240
25		
	ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca	300
	caaccgcaa acggcgcgcc atgctgctgc aggccttcct gttcctgctg gccggcttcg	360
	ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca	420
30		
	cccccaaca gggctggatg aacgaccca acggcctgtg gtacgacgag aaggacgcca	480

ES 2 583 639 T3

agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg cccttgttct 540
 ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc 600
 5 cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct 660
 ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc tggacctaca 720
 acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca 780
 10 ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg 840
 tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca 900
 15 agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca 960
 acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc 1020
 aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg 1080
 20 gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc gaggccttcg 1140
 acaaccagtc ccgctgggtg gacttcggca aggactacta cgccctgcag accttcttca 1200
 25 acaccgacc caccctacggg agcgccttg gcatcgctg ggccccaac tgggagtact 1260
 ccgccttcgt gccaccaac ccctggcgct cctccatgct cctcgtgcgc aagttctccc 1320
 tcaacaccga gtaccaggcc aaccgggaga cggagctgat caacctgaag gccgagccga 1380
 30 tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga 1440

ES 2 583 639 T3

aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg 1500

tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct 1560

5 ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgcgt 1620

cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aaccctact 1680

tcaccaaccg catgagcgtg aacaaccagc cttcaagag cgagaacgac ctgtcctact 1740

10 acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg 1800

tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga 1860

15 cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgagag gtcaagtgat 1920

taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggctcgtg 1980

tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat 2040

20 caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt 2100

gtgctatttg cgaataccac ccccagcatc cccttcctc gtttcatatc gcttgcaccc 2160

25 caaccgcaac ttatctacgc tgtcctgcta tcctcagcg ctgctcctgc tcctgctcac 2220

tgcccctcgc acagccttgg tttgggctcc gcctgtattc tcctggact gcaacctgta 2280

aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atgga 2335

30

<210> 59

ES 2 583 639 T3

<211> 382

<212> PRT

<213> Cinnamomum camphorum

5 <400> 59

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

10 Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

15 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

20

Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln
65 70 75 80

25

Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Asn Pro Pro Gln Leu Leu
85 90 95

30

Asp Asp His Phe Gly Pro His Gly Leu Val Phe Arg Arg Thr Phe Ala
100 105 110

ES 2 583 639 T3

Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr Ser Ile Val Ala
115 120 125

5

Val Met Asn His Leu Gln Glu Ala Ala Leu Asn His Ala Lys Ser Val
130 135 140

10

Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu Met Ser Lys Arg
145 150 155 160

15

Asp Leu Ile Trp Val Val Lys Arg Thr His Val Ala Val Glu Arg Tyr
165 170 175

20

Pro Ala Trp Gly Asp Thr Val Glu Val Glu Cys Trp Val Gly Ala Ser
180 185 190

25

Gly Asn Asn Gly Arg Arg His Asp Phe Leu Val Arg Asp Cys Lys Thr
195 200 205

30

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

ES 2 583 639 T3

Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp Glu Glu Ile Lys
245 250 255

5

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

10

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

15

Glu Ser His His Ile Ser Ser Phe Thr Ile Glu Tyr Arg Arg Glu Cys
305 310 315 320

20

Thr Met Asp Ser Val Leu Gln Ser Leu Thr Thr Val Ser Gly Gly Ser
325 330 335

25

Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln Leu Glu Gly Gly
340 345 350

30

Ser Glu Val Leu Arg Ala Lys Thr Glu Trp Arg Pro Lys Leu Thr Asp
355 360 365

ES 2 583 639 T3

Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Ser Ser Val

370

375

380

5 <210> 60
 <211> 1240
 <212> ADN
 <213> Cinnamomum camphorum

10 <400> 60
 ggcgcgccat ggccaccacc tccttggcct ccgccttctg cagcatgaag gccgtgatgc 60
 tggcccgcga cggccgcggc atgaagcccc gctccagcga cctgcagctg cgcgccggca 120

15 acgcccagac ctccctgaag atgatcaacg gcaccaagtt ctctacacc gagagcctga 180
 agaagctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc tccgccgccg 240
 agaagcagtg gaccaacctg gagtgaagc ccaagcccaa cccccccag ctgctggacg 300

20 accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc agctacgagg 360
 tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag gaggccgccc 420

25 tgaaccacgc caagtccgtg ggcacccctg gcgacggctt cggcaccacc ctggagatgt 480
 ccaagcgcga cctgatctgg gtggtgaagc gcaccacgt ggccgtggag cgctaccccc 540
 cctggggcga caccgtggag gtggagtgtt ggggtggcgc ctccggcaac aacggccgcc 600

30 gccacgactt cctggtgctg gactgcaaga ccggcgagat cctgacccgc tgcacctccc 660

ES 2 583 639 T3

tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag gaggtgcgcg 720
 gcgagatcgg ccccgcccttc atcgacaacg tggccgtgaa ggacgaggag atcaagaagc 780
 5 cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga 840
 acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg atcctggaga 900
 ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc gagtaccgcc 960
 10 gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc ggctcctccg 1020
 aggccggcct ggtgtgagcag cacctgctgc agctggaggg cggcagcagag gtgctgcgcg 1080
 15 ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg 1140
 ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac cacgacatcg 1200
 actacaagga cgacgacgac aagtgactcg agttaattaa 1240

20

<210> 61

<211> 415

<212> PRT

25

<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

30

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu

ES 2 583 639 T3

20

25

30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val

5

35

40

45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser

10

50

55

60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro

15

65

70

75

80

Pro Pro Arg Thr Phe Leu His Gln Leu Pro Asp Trp Ser Arg Leu Leu

85

90

95

20

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

25

115

120

125

Glu Ser Thr Val Gln Asp Gly Leu Val Phe Arg Gln Ser Phe Ser Ile

30

130

135

140

Arg Ser Tyr Glu Ile Gly Thr Asp Arg Thr Ala Ser Ile Glu Thr Leu

ES 2 583 639 T3

	145		150		155		160									
	Met	Asn	His	Leu	Gln	Glu	Thr	Ser	Leu	Asn	His	Cys	Lys	Ser	Thr	Gly
5				165					170						175	
	Ile	Leu	Leu	Asp	Gly	Phe	Gly	Arg	Thr	Leu	Glu	Met	Cys	Lys	Arg	Asp
10				180					185						190	
	Leu	Ile	Trp	Val	Val	Ile	Lys	Met	Gln	Ile	Lys	Val	Asn	Arg	Tyr	Pro
15				195					200						205	
	Ala	Trp	Gly	Asp	Thr	Val	Glu	Ile	Asn	Thr	Arg	Phe	Ser	Arg	Leu	Gly
20				210					215						220	
	Lys	Ile	Gly	Met	Gly	Arg	Asp	Trp	Leu	Ile	Ser	Asp	Cys	Asn	Thr	Gly
25				225					230						235	
	Glu	Ile	Leu	Val	Arg	Ala	Thr	Ser	Ala	Tyr	Ala	Met	Met	Asn	Gln	Lys
30				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	

ES 2 583 639 T3

His Lys Phe Lys Val Lys Thr Gly Asp Ser Ile Gln Lys Gly Leu Thr
290 295 300

5

Pro Gly Trp Asn Asp Leu Asp Val Asn Gln His Val Ser Asn Val Lys
305 310 315 320

10

Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Thr Glu Val Leu Glu Thr
325 330 335

15

Gln Glu Leu Cys Ser Leu Ala Leu Glu Tyr Arg Arg Glu Cys Gly Arg
340 345 350

20

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

25

Val Arg Ser Gln Tyr Gln His Leu Leu Arg Leu Glu Asp Gly Thr Ala
370 375 380

30

Gly Ala Ile Ser Thr Gly Lys Thr Ser Asn Gly Asn Ser Val Ser
405 410 415

ES 2 583 639 T3

<210> 62

<211> 1339

<212> ADN

5 <213> *Cuphea hookeriana*

<400> 62

ggcgcgccat ggtggccgcc gccgcctcca gcgccttctt ccccgtgccc gccccggcg 60

10 cctcccccaa gcccggaag ttcggcaact ggccctccag cctgagcccc tccttcaagc 120

ccaagtccat ccccaacggc ggcttccagg tgaaggccaa cgacagcgcc caccccaagg 180

ccaacggctc cgccgtgagc ctgaagagcg gcagcctgaa caccaggag gacacctcct 240

15 ccagcccccc cccccgacc ttctgcacc agctgcccga ctggagccgc ctgctgaccg 300

ccatcaccac cgtgttcgtg aagtccaagc gccccgacat gcacgaccgc aagtccaagc 360

gccccgacat gctggtggac agcttcggcc tggagtccac cgtgcaggac ggcttgggtg 420

20 tccgccagtc cttctccatc cgctcctacg agatcggcac cgaccgcacc gccagcatcg 480

agacctgat gaaccacctg caggagacct ccctgaacca ctgcaagagc accggcatcc 540

25 tgctggacgg cttcggccgc accctggaga tgtgcaagcg cgacctgatc tgggtggtga 600

tcaagatgca gatcaaggtg aaccgctacc ccgcctgggg cgacaccgtg gagatcaaca 660

cccgttcag ccgcctgggc aagatcggca tgggccgca ctggctgatc tccgactgca 720

30 acaccggcga gatcctggtg cgcgccacca gcgcctacgc catgatgaac cagaagacct 780

ES 2 583 639 T3

gccgcctgtc caagctgccc tacgaggtgc accaggagat cgtgcccctg ttcgtggaca 840

gccccgtgat cgaggactcc gacctgaagg tgcacaagtt caaggtgaag accggcgaca 900

5 gcatccagaa gggcctgacc cccggctgga acgacctgga cgtgaaccag cacgtgtcca 960

acgtgaagta catcggctgg atcctggaga gcatgcccac cgaggtgctg gagaccagg 1020

agctgtgctc cctggccctg gagtaccgcc gcgagtgcgg ccgcgactcc gtgctggaga 1080

10 gcgtgaccgc catggacccc agcaaggtgg gcgtgcgctc ccagtaccag cacctgctgc 1140

gcctggagga cggcaccgcc atcgtgaacg gcgccaccga gtggcgcccc aagaacgccg 1200

15 gcgccaacgg cgccatctcc accggcaaga ccagcaacgg caactccgtg tccatggact 1260

acaaggacca cgacggcgac tacaaggacc acgacatcga ctacaaggac gacgacgaca 1320

agtgactcga gttaattaa 1339

20

<210> 63

<211> 382

<212> PRT

25

<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

30

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu

ES 2 583 639 T3

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

ES 2 583 639 T3

Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln His Val Asn Asn
275 280 285

5 Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro Asp Ser Ile Phe
290 295 300

10 Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr Arg Arg Glu Cys
305 310 315 320

15 Thr Arg Asp Ser Val Leu Arg Ser Leu Thr Thr Val Ser Gly Gly Ser
325 330 335

20 Ser Glu Ala Gly Leu Val Cys Asp His Leu Leu Gln Leu Glu Gly Gly
340 345 350

25 Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro Lys Leu Thr Asp
355 360 365

30 Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro Arg Val
370 375 380

<210> 64

<211> 1240

<212> ADN

30 <213> Umbellularia sp.

<400> 64

ES 2 583 639 T3

	ggcgcgccat ggccaccacc agcctggcct ccgccttctg ctccatgaag gccgtgatgc	60
	tggcccgcga cggccgcggc atgaagcccc gcagctccga cctgcagctg cgcgccggca	120
5	acgccccac ctccctgaag atgatcaacg gcaccaagtt cagctacacc gagagcctga	180
	agcgcctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc agcgcgccg	240
	agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag ctgctggacg	300
10	accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc tcctacgagg	360
	tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc	420
15	tgaaccacgc caagagcgtg ggcatacctgg gcgacggctt cggcaccacc ctggagatgt	480
	ccaagcgcga cctgatgtgg gtggtgcgcc gcaccacgt ggccgtggag cgctacccca	540
	cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac aacggcatgc	600
20	gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgc tgcacctccc	660
	tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac gaggtgcgcg	720
25	gcgagatcgg cccgccttc atcgacaacg tggccgtgaa ggacgacgag atcaagaagc	780
	tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgtgga	840
	acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga	900
30	ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg gagtaccgcc	960

ES 2 583 639 T3

	gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg	1020
	aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag gtgctgcgcg	1080
5	cccgaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg	1140
	ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac cagacatcg	1200
	actacaagga cgacgacgac aagtgactcg agttaattaa	1240
10		
	<210> 65	
	<211> 20	
	<212> ADN	
15	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Cebador sintético	
20		
	<400> 65	
	ccgccgtgct ggacgtggtg	20
25	<210> 66	
	<211> 20	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<223> Descripción de Secuencia Artificial: Cebador sintético	

<400> 66

ggtggcgggg tccaggggtg

20

5

<210> 67

<211> 20

<212> ADN

<213> Secuencia Artificial

10

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

15

<400> 67

cggccggcgg ctccttcaac

20

20

<210> 68

<211> 20

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

30

<400> 68

ggcgctcccg taggtcgggt

20

<210> 69

ES 2 583 639 T3

<211> 1335

<212> ADN

<213> *Chlorella sorokiniana*

5 <400> 69
 cgctgcaac gcaagggcag ccacagccgc tcccacccgc cgctgaaccg acacgtgctt 60
 gggcgcctgc cgctgcctg ccgcatgctt gtgctggtga ggctgggcag tgctgccatg 120
 10 ctgattgagg cttggttcat cgggtggaag cttatgtgtg tgctgggctt gcatgccggg 180
 caatgcgcat ggtggcaaga gggcggcagc acttgctgga gctgccgcgg tgcctccagg 240
 tggttcaatc gcggcagcca gagggatttc agatgatcgc gcgtacaggt tgagcagcag 300
 15 tgtcagcaaa ggtagcagtt tgccagaatg atcggttcag ctgttaatca atgccagcaa 360
 gagaaggggt caagtgcaaa cacgggcatg ccacagcacg ggcaccgggg agtggaatgg 420
 20 caccaccaag tgtgtgagcag ccagcatcgc cgctggctg tttcagctac aacggcagga 480
 gtcatccaac gtaaccatga gctgatcaac actgcaatca tcgggcgggc gtgatgcaag 540
 catgcctggc gaagacacat ggtgtgcgga tgctgccggc tgctgcctgc tgcgcacgcc 600
 25 gttgagttgg cagcaggctc agccatgcac tggatggcag ctgggctgcc actgcaatgt 660
 ggtggatagg atgcaagtgg agcgaatacc aaaccctctg gctgcttgct gggttgcatg 720
 30 gcatgcacc atcagcagga gcgcatgcga agggactggc cccatgcacg ccatgcaaaa 780
 ccggagcgc cagagtgtcc aactgtcac caggcccgc agctttgcag aacctgctc 840

ES 2 583 639 T3

atggacgcat gtagcgctga cgtcccttga cggcgctcct ctcgggtgtg ggaaacgcaa 900
 tgcagcacag gcagcagagg cggcggcagc agagcggcgg cagcagcggc gggggccacc 960
 5 cttcttgccg ggtcgcgccc cagccagcgg tgatgcgctg atcccaaacg agttcacatt 1020
 catttgcatt cctggagaag cgaggctggg gcctttgggc tgggtgcagcc cgcaatggaa 1080
 10 tgcgggaccg ccaggctagc agcaaaggcg cctcccctac tccgcatcga tgttccatag 1140
 tgcattggac tgcatttggg tggggcggcc ggctgtttct ttcgtgttgc aaaacgcgcc 1200
 agctcagcaa cctgtcccgt gggccccccg tgccgatgaa atcgtgtgca cgccgatcag 1260
 15 ctgattgccg ggctcgcgaa gtaggcgccc tcctttctgc tcgccctctc tccgtcccgc 1320
 cactagtggc gcgcc 1335
 20
 <210> 70
 <211> 1146
 <212> ADN
 <213> Umbellularia californica
 25
 <400> 70
 atggccacca ccagcctggc ctccgccttc tgctccatga aggccgtgat gctggccccg 60
 gacggccgcg gcatgaagcc ccgcagctcc gacctgcagc tgcgcgccgg caacgcccc 120
 30 acctccctga agatgatcaa cggcaccaag ttcagctaca ccgagagcct gaagcgcctg 180
 cccgactggt ccatgctgtt cgccgtgatc accaccatct tcagcggccg cgagaagcag 240

ES 2 583 639 T3

	tggaccaacc tggagtggaa gcccagccc aagctgcccc agctgctgga cgaccacttc	300
	ggcctgcacg gcctggtggt ccgccgcacc ttcgccatcc gtccttacga ggtgggcccc	360
5	gaccgcagca cctccatcct ggccgtgatg aaccacatgc aggaggccac cctgaaccac	420
	gccaagagcg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcgc	480
10	gacctgatgt ggggtggtgcg ccgcaccac gtggccgtgg agcgctaccc cacctggggc	540
	gacaccgtgg aggtggagtg ctggatcggc gccagcggca acaacggcat gcgccgcgac	600
	ttcctggtgc gcgactgcaa gaccggcgag atcctgacct gctgcacctc cctgagcgtg	660
15	ctgatgaaca cccgcaccgg ccgcctgagc accatccccg acgaggtgcg cggcgagatc	720
	ggccccgcct tcatcgacaa cgtggccgtg aaggacgacg agatcaagaa gctgcagaag	780
20	ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg	840
	gacgtgaacc agcacgtgaa caacctgaag tacgtggcct ggggtgttcga gaccgtgccc	900
	gacagcatct tcgagtccca ccacatcagc tccttcacct tggagtaccg ccgcgagtgc	960
25	accgcgact ccgtgctgcg cagcctgacc accgtgagcg gcggcagctc cgaggccggc	1020
	ctggtgtgcg accacctgct gcagctggag ggcggcagcg aggtgctgcg cgccgcacc	1080
30	gagtggcgcc ccaagctgac cgactccttc cgcgcatca gcgtgatccc cgccgagccc	1140
	cgcgctg	1146

ES 2 583 639 T3

<210> 71
 <211> 1146
 5 <212> ADN
 <213> Cinnamomum camphora

 <400> 71
 atggccacca cctccctggc ctccgccttc tgcagcatga aggccgtgat gctggccccg 60
 10 gacggccgcg gcatgaagcc ccgctccagc gacctgcagc tgcgcgccgg caacgccag 120
 acctccctga agatgatcaa cggcaccaag ttctcctaca ccgagagcct gaagaagctg 180
 15 cccgactggt ccatgctggt cgccgtgatc accaccatct tctccgccgc cgagaagcag 240
 tggaccaacc tggagtggaa gcccaagccc aacccccccc agctgctgga cgaccacttc 300
 ggccccacg gcctggtggt ccgccgcacc ttcgccatcc gcagctacga ggtgggcccc 360
 20 gaccgctcca ccagcatcgt ggccgtgatg aaccacctgc aggaggccgc cctgaaccac 420
 gccaaagtccg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcgc 480
 25 gacctgatct ggggtggtgaa gcgcaccac gtggccgtgg agcgctaccc cgctggggc 540
 gacaccgtgg aggtggagtg ctgggtgggc gcctccggca acaacggccg ccgccacgac 600
 ttctggtgc gcgactgcaa gaccggcgag atcctgacct gctgcacctc cctgagcgtg 660
 30 atgatgaaca cccgcacccg ccgcctgagc aagatccccg aggaggtgcg cggcgagatc 720

ES 2 583 639 T3

	ggccccgcct tcatcgacaa cgtggccgtg aaggacgagg agatcaagaa gccccagaag	780
	ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg	840
5	gacatcaacc agcacgtgaa caacatcaag tacgtggact ggatcctgga gaccgtgccc	900
	gacagcatct tcgagagcca ccacatctcc tccttcacca tcgagtaccg ccgcgagtgc	960
	accatggaca gcgtgctgca gtcctgacc accgtgagcg gcggctcctc cgaggccggc	1020
10	ctggtgtgcg agcacctgct gcagctggag ggcggcagcg aggtgctgcg cgccaagacc	1080
	gagtggcgcc ccaagctgac cgactccttc cgcgcatca gcgtgatccc cgccgagtcc	1140
	agcgtg	1146
15		
	<210> 72	
	<211> 72	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Oligonucleótido sintético	
25		
	<400> 72	
	atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	60
	gacgacaagt ga	72
30		
	<210> 73	

ES 2 583 639 T3

<211> 408

<212> ADN

<213> *Chlorella vulgaris*

5 <400> 73
ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggc gtgtgatgga 60

ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 120

10 cctcagtgtg tttgatcttg tgtgtacgcg cttttgagag ttgctagctg cttgtgctat 180

ttgcgaatac caccaccagc atccccttcc ctcgtttcat atcgcttgca tcccaaccgc 240

aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccct 300

15 cgcacagcct tggtttggc tccgctgta ttctcctggt actgcaacct gtaaaccagc 360

actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatgga 408

20 <210> 74

<211> 20

<212> ADN

<213> Secuencia Artificial

25 <220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 74

30 ctgggacgac gcttcggcac 20

<210> 75

<211> 20

<212> ADN

<213> Secuencia Artificial

5

<220>

<223> Descripción de Secuencia Artificial: Cebador
sintético

10

<400> 75

aagtcgcggc gcatgccgtt

20

<210> 76

15

<211> 20

<212> ADN

<213> Secuencia Artificial

<220>

20

<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 76

tacccgcct gggcgacac

20

25

<210> 77

<211> 20

<212> ADN

30

<213> Secuencia Artificial

<220>

ES 2 583 639 T3

<223> Descripción de Secuencia Artificial: Cebador
sintético

<400> 77

5 cttgctcagg cggcgggtgc 20
<210> 78
<211> 1317
<212> ADN
<213> Secuencia Artificial

10 <220>
<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

15 <400> 78
atggtggccg ccgccgctc cagcgcctc ttccccgtgc ccgccccgg cgctcccc 60
aagcccggca agttcggcaa ctggccctcc agcctgagcc cctccttcaa gccaagtcc 120
20 atccccaacg gcggttcca ggtgaaggcc aacgacagcg cccaccccaa ggccaacggc 180
tccgccgtga gcctgaagag cggcagcctg aacaccagg aggacacctc ctccagcccc 240
ccccccgca ccttctgca ccagctgccc gactggagcc gcctgctgac cgccatcacc 300
25 accgtgttcg tgaagtcaa gcgccccgac atgcacgacc gcaagtcaa gcgccccgac 360
atgctggtgg acagcttcgg cctggagtcc accgtgcagg acggcctggt gttccgccag 420
30 tccttctcca tccgctccta cgagatcggc accgaccgca ccgccagcat cgagaccctg 480
atgaaccacc tgcaggagac ctccctgaac cactgcaaga gcaccggcat cctgctggac 540

ES 2 583 639 T3

	ggcttcggcc gcaccctgga gatgtgcaag cgcgacctga tctgggtggt gatcaagatg	600
	cagatcaagg tgaaccgcta ccccgctgg ggcgacaccg tggagatcaa caccgcttc	660
5	agccgctgg gcaagatcgg catgggccgc gactggctga tctccgactg caacaccggc	720
	gagatcctgg tgcgcccac cagcgcctac gccatgatga accagaagac ccgccgctg	780
10	tccaagctgc cctacgaggt gcaccaggag atcgtgcccc tgttcgtgga cagccccgtg	840
	atcgaggact ccgacctgaa ggtgcacaag ttcaaggatga agaccggcga cagcatccag	900
	aagggcctga cccccggctg gaacgacctg gacgtgaacc agcacgtgtc caacgtgaag	960
15	tacatcggct ggatcctgga gagcatgccc accgaggtgc tggagacca ggagctgtgc	1020
	tccctggccc tggagtaccg ccgcgagtgc ggccgagact ccgtgctgga gagcgtgacc	1080
	gccatggacc ccagcaaggt gggcgtgctc tcccagtacc agcacctgct ggcctggag	1140
20	gacggcaccg ccatcgtgaa cggcgccacc gactggcgcc ccaagaacgc cggcgccaac	1200
	ggcgccatct ccaccggcaa gaccagcaac ggcaactccg tgtccatgga ctacaaggac	1260
25	cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtga	1317
	<210> 79	
	<211> 1170	
30	<212> ADN	
	<213> Secuencia Artificial	

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5 <400> 79
atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcccagctg 120
10 cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcgcc 180
gacatgcacg accgcaagtc caagcgcgcc gacatgctgg tggacagctt cggcctggag 240
tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc 300
15 ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg 360
aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc 420
20 aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc 480
tggggcgaca ccgtggagat caacaccgc ttcagccgcc tgggcaagat cggcatggg 540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcg caccagcgcc 600
25 tacgcatga tgaaccagaa gacccgccgc ctgtccaagc tgccttacga ggtgcaccag 660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac 720
30 aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac 780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg 840

ES 2 583 639 T3

cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag 900

5 tgcggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg 960

cgctcccagt accagcacct gctgcgctg gaggacggca ccgccatcgt gaacggcgcc 1020

accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc 1080

10 aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactaaa ggaccacgac 1140

atcgactaca aggacgacga cgacaagtga 1170

15 <210> 80
<211> 1170
<212> ADN
<213> Secuencia Artificial

20 <220>
<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 80

25 atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60

cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgcccagctg 120

cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc 180

30 gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag 240

ES 2 583 639 T3

	tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc	300
	ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg	360
5	aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc	420
	aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctacccccgcc	480
	tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggcatgggc	540
10	cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcgc caccagcgcc	600
	tacgccatga tgaaccagaa gacccgccgc ctgtccaagc tgccttacga ggtgcaccag	660
15	gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac	720
	aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac	780
	ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg	840
20	cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgag	900
	tgcggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg	960
25	cgctcccagt accagcacct gctgcgctg gaggacggca ccgccatcgt gaacggcgcc	1020
	accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc	1080
	aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac	1140
30	atcgactaca aggacgacga cgacaagtga	1170

ES 2 583 639 T3

<210> 81

<211> 1167

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

10

<400> 81

atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg 60

gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggdcgcg ccagctgccc 120

15

gactggagcc gcctgctgac cgccatcacc accgtgttcg tgaagtccaa gcgccccgac 180

atgcacgacc gcaagtccaa gcgccccgac atgctggtgg acagcttcgg cctggagtcc 240

20

accgtgcagg acggcctggt gttccgccag tccttctcca tccgctccta cgagatcggc 300

accgaccgca ccgccagcat cgagaccctg atgaaccacc tgcaggagac ctccctgaac 360

cactgcaaga gcaccggcat cctgctggac ggcttcggcc gcaccctgga gatgtgcaag 420

25

cgcgacctga tctgggtggt gatcaagatg cagatcaagg tgaaccgcta cccgcctgg 480

ggcgacaccg tggagatcaa caccgcttc agccgcctgg gcaagatcgg catgggcccg 540

30

gactggctga tctccgactg caacaccggc gagatcctgg tgcgcgccac cagcgcctac 600

gccatgatga accagaagac ccgccgctg tccaagctgc cctacgaggt gcaccaggag 660

ES 2 583 639 T3

	atcgtgcccc tgttcgtgga cagccccgtg atcgaggact ccgacctgaa ggtgcacaag	720
	ttcaaggtga agaccggcga cagcatccag aagggcctga cccccggctg gaacgacctg	780
5	gacgtgaacc agcacgtgtc caacgtgaag tacatcggct ggatcctgga gagcatgccc	840
	accgaggtgc tggagacca ggagctgtgc tccctggccc tggagtaccg ccgcgagtgc	900
10	ggccgcgact ccgtgctgga gagcgtgacc gccatggacc ccagcaaggt gggcgtgctgc	960
	tcccagtacc agcacctgct ggcctggag gacggcaccg ccatcgtgaa cggcgccacc	1020
	gagtggcgcc ccaagaacgc cggcgccaac ggcgcatct ccaccggcaa gaccagcaac	1080
15	ggcaactccg tgtccatgga ctacaaggac cagcagggcg actacaagga ccacgacatc	1140
	gactacaagg acgacgacga caagtga	1167
20		
	<210> 82	
	<211> 1149	
	<212> ADN	
	<213> Secuencia Artificial	
25		
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
30		
	<400> 82	
	atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg	60

ES 2 583 639 T3

	gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggcgcgc ccccgactgg	120
	tccatgctgt tcgccgtgat caccaccatc ttcagcgccg ccgagaagca gtggaccaac	180
5	ctggagtgga agcccaagcc caagctgccc cagctgctgg acgaccactt cggcctgcac	240
	ggcctggtgt tccgccgcac cttcgccatc cgctcctacg aggtgggccc cgaccgcagc	300
	acctccatcc tggccgtgat gaaccacatg caggaggcca ccctgaacca cgccaagagc	360
10	gtgggcatcc tgggcgacgg cttcggcacc accctggaga tgtccaagcg cgacctgatg	420
	tgggtggtgc gccgcaccca cgtggccgtg gagcgctacc ccacctgggg cgacaccgtg	480
15	gaggtggagt gctggatcgg cgccagcggc aacaacggca tgcgccgca cttcctggtg	540
	cgcgactgca agaccggcga gatcctgacc cgctgcacct ccctgagcgt gctgatgaac	600
	accgcaccc gccgcctgag caccatcccc gacgaggtgc gcggcgagat cggccccgcc	660
20	ttcatcgaca acgtggccgt gaaggacgac gagatcaaga agctgcagaa gctgaacgac	720
	tccaccgccg actacatcca gggcggcctg accccccgct ggaacgacct ggacgtgaac	780
25	cagcacgtga acaacctgaa gtacgtggcc tgggtgttcg agaccgtgcc cgacagcatc	840
	ttcgagtccc accacatcag ctccttcacc ctggagtacc gccgcgagtg caccgcgac	900
	tccgtgctgc gcagcctgac caccgtgagc ggcggcagct ccgaggccgg cctggtgtgc	960
30	gaccacctgc tgcagctgga gggcggcagc gaggtgctgc gcgcccgcac cgagtggcgc	1020

ES 2 583 639 T3

	cccaagctga ccgactcctt ccgcgggcatc agcgtgatcc ccgccgagcc ccgctgatg	1080
	gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac	1140
5	gacaagtga	1149
	<210> 83	
	<211> 1146	
10	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido	
15	sintético	
	<400> 83	
	atggcttccg cggcattcac catgtcggcg tgccccgca tgactggcag ggcccctggg	60
20	gcacgtcgct ccggacggcc agtcgccacc cgctgaggg ggcgcgcccc cgactggctc	120
	atgctgttcg ccgtgatcac caccatcttc agcgcgccg agaagcagtg gaccaacctg	180
	gagtggaagc ccaagcccaa gctgccccag ctgctggacg accacttcgg cctgcacggc	240
25	ctggtgttcc gccgcacctt cgccatccgc tcctacgagg tgggccccga ccgcagcacc	300
	tccatcctgg ccgtgatgaa ccacatgcag gaggccacc tgaaccacgc caagagcgtg	360
30	ggcatcctgg gcgacggctt cggcaccacc ctggagatgt ccaagcgca cctgatgtgg	420
	gtggtgcgcc gcaccacgt ggccgtggag cgctaccca cctggggcga caccgtggag	480

ES 2 583 639 T3

	gtggagtgct ggatcggcgc cagcggcaac aacggcatgc gccgcgactt cctggtgcb	540
	gactgcaaga ccggcgagat cctgacccgc tgcacctccc tgagcgtgct gatgaacacc	600
5	cgaccccgcc gcctgagcac catccccgac gaggtgcbgc gcgagatcgg cccgccttc	660
	atcgacaacg tggccgtgaa ggacgacgag atcaagaagc tgcagaagct gaacgactcc	720
10	accgccgact acatccaggc cggcctgacc ccccgctgga acgacctgga cgtgaaccag	780
	cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga ccgtgcccga cagcatcttc	840
	gagtcccacc acatcagctc cttcacctg gagtaccgcc gcgagtgcac ccgcgactcc	900
15	gtgctgcbca gcctgaccac cgtgagcggc ggcagctccg aggccggcct ggtgtgcbac	960
	cacctgctgc agctggaggc cggcagcagc gtgctgcbgc cccgcaccga gtggcgcctc	1020
20	aagctgaccg actccttcgc cggcatcagc gtgatccccg ccgagccccg cgtgatggac	1080
	tacaaggacc acgacggcga ctacaaggac cacgacatcg actacaagga cgacgacgac	1140
	aagtga	1146
25		
	<210> 84	
	<211> 1155	
	<212> ADN	
30	<213> Secuencia Artificial	
	<220>	

ES 2 583 639 T3

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 84

5	atggccaccg catccacttt ctcggcggtt aatgcccgct gcggcgacct gcgtcgctcg	60
	gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcy cgccccgac	120
	tggatcatgc tgttcgccgt gatcaccacc atcttcagcy ccgccgagaa gcagtggacc	180
10	aacctggagt ggaagcccaa gcccaagctg cccagctgc tggacgacca cttcggcctg	240
	cacggcctgg tgttccgccg caccttcgcc atccgctcct acgaggtggg ccccgaccgc	300
15	agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag	360
	agcgtgggca tcctgggca cggcttcggc accaccctgg agatgtcaa gcgcgacctg	420
	atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgt accccacctg gggcgacacc	480
20	gtggaggtgg agtgctgat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg	540
	gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg	600
25	aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc	660
	gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac	720
	gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg	780
30	aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgacagc	840

ES 2 583 639 T3

	atcttcgagt cccaccacat cagctccttc accctggagt accgccgca gtgcacccgc	900
	gactccgtgc tgcgcagcct gaccaccgtg agcggcggca gctccgaggc cggcctggtg	960
5	tgcgaccacc tgctgcagct ggagggcggc agcgagggtgc tgcgcgcccg caccgagtgg	1020
	cgccccaaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg	1080
	atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
10	gacgacaagt gatga	1155
	<210> 85	
15	<211> 1152	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
20	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 85	
	atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg	60
25	cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgccccgac	120
	tggatcatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc	180
30	aacctggagt ggaagcccaa gcccaagctg ccccagctgc tggacgacca cttcggcctg	240
	cacggcctgg tgttccgccg caccttcgcc atccgctcct acgaggtggg ccccgaccgc	300

ES 2 583 639 T3

	agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag	360
	agcgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg	420
5	atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgct accccacctg gggcgacacc	480
	gtggaggtgg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttctg	540
10	gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg	600
	aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc	660
	gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac	720
15	gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg	780
	aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgcacgc	840
20	atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc	900
	gactccgtgc tgcgcagcct gaccaccgtg agcggcggca gctccgaggc cggcctggtg	960
	tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccg caccgagtgg	1020
25	cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg	1080
	atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
30	gacgacaagt ga	1152

ES 2 583 639 T3

<210> 86

<211> 1155

<212> ADN

<213> Secuencia Artificial

5

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

10

<400> 86

atggccaccg catccacttt ctcggcggtc aatgcccgt gcggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccccgac 120

15

tggtccatgc tgttcgccgt gatcaccacc atcttctccg ccgccgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccaacccc ccccagctgc tggacgacca cttcggcccc 240

cacggcctgg tgttccgccg caccttcgcc atccgcagct acgaggtggg ccccgaccgc 300

20

tccaccagca tcgtggccgt gatgaaccac ctgcaggagg ccgccctgaa ccacgccaag 360

tccgtgggca tcctgggcca cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420

25

atctgggtgg tgaagcgcac ccacgtggcc gtggagcgct accccgcctg gggcgacacc 480

gtggaggtgg agtgctgggt gggcgcctcc ggcaacaacg gccgccgcca cgacttctctg 540

gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg 600

30

aacacccgca cccgccgcct gagcaagatc cccgaggagg tgcgcggcga gatcggcccc 660

ES 2 583 639 T3

	gccttcatcg acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac	720
	gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacatc	780
5	aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgacagc	840
	atcttcgaga gccaccacat ctctccttc accatcgagt accgccgca gtgcaccatg	900
	gacagcgtgc tgcagtcctt gaccaccgtg agcggcggct cctccgaggc cggcctggtg	960
10	tgcgagcacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcaa gaccgagtgg	1020
	cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gtccagcgtg	1080
15	atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
	gacgacaagt gatga	1155
20	<210> 87	
	<211> 1893	
	<212> ADN	
	<213> Secuencia Artificial	
25	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 87	
30	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	60
	tccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	120

ES 2 583 639 T3

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 5 cacaggccac tcgagcttgt gatcgcactc cgctaagggg ggcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 360
 10 cgctgcggcg acctgcgctg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 420
 cccgtgcgcy ggcgcccc cgactggctc atgctgttcg ccgtgatcac caccatcttc 480
 tccgccgcy agaagcagtg gaccaacctg gagggaagc ccaagcccaa cccccccag 540
 15 ctgctggagc accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc 600
 agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag 660
 20 gaggccgcc tgaaccacgc caagtccgtg ggcacccctg gcgacggctt cggcaccacc 720
 ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcaccacgt ggccgtggag 780
 cgctaccccg cctggggcga caccgtggag gtggagtgcg ggggtggcgc ctccggcaac 840
 25 aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 900
 tgcacctcc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag 960
 30 gaggtgcgcy gcgagatcgg cccgccttc atcgacaacg tggccgtgaa ggacgaggag 1020
 atcaagaagc cccagaagct gaacgactcc accgccgact acatccaggc cggcctgacc 1080

ES 2 583 639 T3

	ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg	1140
	atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc	1200
5	gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc	1260
	ggctcctccg aggccggcct ggtgtgagc cacctgctgc agctggaggg cggcagcgag	1320
10	gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc	1380
	gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac	1440
	cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	1500
15	atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg	1560
	ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt	1620
20	gtgtacgcbc ttttgcgagt tgctagctgc ttgtgctatt tgccaatacc acccccagca	1680
	tccccttccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc	1740
	tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct	1800
25	ccgcctgtat tctcctgcta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg	1860
	aagtagtggg atgggaacac aaatggaaag ctt	1893
30		
	<210> 88	
	<211> 1887	

ES 2 583 639 T3

<212> ADN

<213> Secuencia Artificial

<220>

5 <223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 88

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
10 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
15 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 300
cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcacatgctc ggcgtgcccc 360
20 gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg 420
agggggcgcg cccccgactg gtccatgctg ttcgccgtga tcaccacat cttctccgcc 480
25 gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaaccccc ccagctgctg 540
gacgaccact tcggccccca cggcctggtg ttccgccgca ccttcgcat ccgcagctac 600
gaggtgggcc ccgaccgctc caccagcatc gtggccgtga tgaaccacct gcaggaggcc 660
30 gccctgaacc acgccaagtc cgtgggcatc ctgggagcag gcttcggcac caccctggag 720

ES 2 583 639 T3

	atgtccaagc gcgacctgat ctgggtggtg aagcgcaccc acgtggccgt ggagcgctac	780
	cccgcctggg gcgacaccgt ggaggtggag tgctgggtgg gcgctccgg caacaacggc	840
5	cgccgccacg acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgctgcacc	900
	tccctgagcg tgatgatgaa caccgcacc cgccgcctga gcaagatccc cgaggaggtg	960
	cgcggcgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga ggagatcaag	1020
10	aagccccaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc	1080
	tggaacgacc tggacatcaa ccagcacgtg aacaacatca agtacgtgga ctggatcctg	1140
15	gagaccgtgc ccgacagcat cttcgagagc caccacatct cctccttcac catcgagtac	1200
	cgccgcgagt gcacatgga cagcgtgctg cagtcctga ccaccgtgag cggcggctcc	1260
	tccgaggccg gcctggtgtg cgagcacctg ctgcagctgg agggcggcag cgaggtgctg	1320
20	cgcgccaaga ccgagtggcg cccaagctg accgactcct tccgcggcat cagcgtgatc	1380
	cccgccgagt ccagcgtgat ggactacaag gaccacgacg gcgactaaa ggaccacgac	1440
25	atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta	1500
	tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga	1560
	cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac	1620
30	gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct	1680

ES 2 583 639 T3

	tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc	1740
	tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct	1800
5	gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag	1860
	tgggatggga acacaaatgg aaagctt	1887
10	<210> 89	
	<211> 3631	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 89	
20	gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtgaggct gggcagtgct	120
	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
25	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
30	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420

ES 2 583 639 T3

	gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
5	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
10	caatgtggtg gataggatgc aagtggagcg aatacacaac cctctggctg cttgctgggt	720
	tgcatggcat cgcacatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
	gccaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
15	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960
20	gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcatgcctg gagaagcag gctggggcct ttgggctggt gcagcccgca	1080
	atggaatgcy ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgaatgtt	1140
25	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
30	gatcagctga ttgccggct cgcaagtag gcgccctcct ttctgctcgc cctctctccg	1320
	tcccgcact agtatgctgc tgcaggcctt cctgttctcg ctggccggct tcgccgcaa	1380

ES 2 583 639 T3

	gatcagcgcc tccatgacga acgagacgtc cgaccgcccc ctggtgcact tcacccccaa	1440
	caagggctgg gggcgcgcca gccaccacgt gtacaagcgc ctgaccaga gcaccaacac	1500
5	caagtcccc agcgtgaacc agccctaccg caccggcttc cacttccagc cccccaagaa	1560
	ctggatgaac gacccaacg gcccctgat ctacaagggc atctaccacc tgttctacca	1620
10	gtggaacccc aagggcgccg tgtggggcaa catcgtgtgg gccactcca ccagcaccga	1680
	cctgatcaac tgggaccccc accccccgc catcttcccc agcgccccct tcgacatcaa	1740
	cggctgctgg tccggcagcg ccaccatcct gcccaacggc aagcccgtga tcctgtacac	1800
15	cggcatcgac cccaagaacc agcaggtgca gaacatcgcc gagcccaaga acctgtccga	1860
	cccctacctg cgcgagtgga agaagagccc cctgaacccc ctgatggccc ccgacgccgt	1920
20	gaacggcatc aacgcctcca gcttccgca ccccaccacc gcctggctgg gccaggacia	1980
	gaagtggcgc gtgatcatcg gctccaagat ccaccgccgc ggctggcca tcacctacac	2040
	cagcaaggac ttctgaagt gggagaagtc ccccgagccc ctgcactacg acgacggcag	2100
25	cggcatgtgg gagtgccccg acttcttccc cgtgaccgc ttcggcagca acggcgtgga	2160
	gacctccagc ttcggcgagc ccaacgagat cctgaagcac gtgctgaaga tctccctgga	2220
30	cgacaccaag cacgactact acaccatcgg cacctacgac cgcgtgaagg acaagttcgt	2280
	gcccgacaac ggcttcaaga tggacggcac cgccccccgc tacgactacg gcaagtacta	2340

ES 2 583 639 T3

	cgccagcaag accttcttcg actccgcca gaaccgccgc atcctgtggg gctggaccaa	2400
	cgagtcctcc agcgtggagg acgacgtgga gaagggctgg tccggcatcc agaccatccc	2460
5	ccgcaagatc tggctggacc gcagcggcaa gcagctgac cagtggcccg tgcgagaggt	2520
	ggagcgcctg cgcaccaagc aggtgaagaa cctgcgcaac aaggtgctga agtccggcag	2580
10	ccgcctggag gtgtacggcg tgaccgccgc ccaggccgac gtggaggtgc tgttcaaggt	2640
	gcgcgacctg gagaaggccg acgtgatcga gccctcctgg accgaccccc agctgatctg	2700
	cagcaagatg aacgtgtccg tgaagtccgg cctgggcccc ttcggcctga tgggtctggc	2760
15	cagcaagaac ctggaggagt acacctccgt gtacttccgc atcttcaagg cccgccagaa	2820
	cagcaacaag tacgtggtgc tgatgtgctc cgaccagtcc cgcagctccc tgaaggagga	2880
20	caacgacaag accacctacg gcgccttcgt ggacatcaac ccccaccagc ccctgagcct	2940
	gcgcgccctg atcgaccact ccgtggtgga gagcttcggc ggcaagggcc gcgcctgcat	3000
	cacctccgc gtgtaccca agctggccat cggcaagtcc agccacctgt tcgccttcaa	3060
25	ctacggctac cagtccgtgg acgtgctgaa cctgaacgcc tggagcatga actccgcca	3120
	gatcagcatg gactacaagg accacgacgg cgactacaag gaccacgaca tcgactaaa	3180
30	ggacgacgac gacaagtgat taattaaccg gctcgaggca gcagcagctc ggatagtatc	3240
	gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc tgccttgacc	3300

ES 2 583 639 T3

5
 10
 15
 20
 25
 30

tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatcct gtgtgtacgc 3360
 gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag catccccttc 3420
 cctcgtttca tatcgcttgc atccaaccg caacttatct acgctgtcct gctatccctc 3480
 agcgctgctc ctgctcctgc tcaactgcccc tcgcacagcc ttggtttggg ctccgcctgt 3540
 attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg 3600
 ggatgggaac acaaatggaa agcttgagct c 3631

15 <210> 90
 <211> 621
 <212> PRT
 <213> Secuencia Artificial
 20 <220>
 <223> Descripción de Secuencia Artificial: Polipéptido
 sintético

25 <400> 90
 Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
 1 5 10 15

30 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

ES 2 583 639 T3

35

40

45

Arg Leu Thr Gln Ser Thr Asn Thr Lys Ser Pro Ser Val Asn Gln Pro

5

50

55

60

Tyr Arg Thr Gly Phe His Phe Gln Pro Pro Lys Asn Trp Met Asn Asp

10

65

70

75

80

Pro Asn Gly Pro Met Ile Tyr Lys Gly Ile Tyr His Leu Phe Tyr Gln

15

85

90

95

Trp Asn Pro Lys Gly Ala Val Trp Gly Asn Ile Val Trp Ala His Ser

20

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

25

130

135

140

Ile Leu Pro Asn Gly Lys Pro Val Ile Leu Tyr Thr Gly Ile Asp Pro

30

145

150

155

160

Lys Asn Gln Gln Val Gln Asn Ile Ala Glu Pro Lys Asn Leu Ser Asp

ES 2 583 639 T3

165

170

175

Pro Tyr Leu Arg Glu Trp Lys Lys Ser Pro Leu Asn Pro Leu Met Ala

5

180

185

190

Pro Asp Ala Val Asn Gly Ile Asn Ala Ser Ser Phe Arg Asp Pro Thr

10

195

200

205

Thr Ala Trp Leu Gly Gln Asp Lys Lys Trp Arg Val Ile Ile Gly Ser

210

215

220

15

Lys Ile His Arg Arg Gly Leu Ala Ile Thr Tyr Thr Ser Lys Asp Phe

225

230

235

240

20

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

25

Asn Gly Val Glu Thr Ser Ser Phe Gly Glu Pro Asn Glu Ile Leu Lys

275

280

285

30

His Val Leu Lys Ile Ser Leu Asp Asp Thr Lys His Asp Tyr Tyr Thr

290

295

300

ES 2 583 639 T3

Ile Gly Thr Tyr Asp Arg Val Lys Asp Lys Phe Val Pro Asp Asn Gly
305 310 315 320

5

Phe Lys Met Asp Gly Thr Ala Pro Arg Tyr Asp Tyr Gly Lys Tyr Tyr
325 330 335

10

Ala Ser Lys Thr Phe Phe Asp Ser Ala Lys Asn Arg Arg Ile Leu Trp
340 345 350

15

Gly Trp Thr Asn Glu Ser Ser Ser Val Glu Asp Asp Val Glu Lys Gly
355 360 365

20

Trp Ser Gly Ile Gln Thr Ile Pro Arg Lys Ile Trp Leu Asp Arg Ser
370 375 380

25

Gly Lys Gln Leu Ile Gln Trp Pro Val Arg Glu Val Glu Arg Leu Arg
385 390 395 400

30

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

ES 2 583 639 T3

Leu Phe Lys Val Arg Asp Leu Glu Lys Ala Asp Val Ile Glu Pro Ser
 435 440 445

5

Trp Thr Asp Pro Gln Leu Ile Cys Ser Lys Met Asn Val Ser Val Lys
 450 455 460

10

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

15

Ser Asn Lys Tyr Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser
 500 505 510

20

Leu Lys Glu Asp Asn Asp Lys Thr Thr Tyr Gly Ala Phe Val Asp Ile
 515 520 525

25

Asn Pro His Gln Pro Leu Ser Leu Arg Ala Leu Ile Asp His Ser Val
 530 535 540

30

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

ES 2 583 639 T3

565

570

575

Tyr Gly Tyr Gln Ser Val Asp Val Leu Asn Leu Asn Ala Trp Ser Met

5

580

585

590

Asn Ser Ala Gln Ile Ser Met Asp Tyr Lys Asp His Asp Gly Asp Tyr

10

595

600

605

Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

15

610

615

620

<210> 91

<211> 997

<212> ADN

<213> Prototheca moriformis

20

<400> 91

cctgtcgatc gaagagaagg agacatgtgt acattattgg tgtgagggcg ctgaatcggc 60

cattttttaa aatgatcacg ctcatgccaa tagacgcggc acataacgac gttcaaacc 120

25

ccgcaaagc cgcggacaac cccatccctc cacaccccc acacaaagaa cccgccaccg 180

cttaccttgc ccacgagta ggcctttcgt tgcgcaaac cggcctcggg gatgaatgca 240

30

tgcccgttcc tgacgagcgc tgcccgggcc aacacgctct tttgctgcgt ctctcaggc 300

ttgggggcct ccttgggctt ggggtgccgc atgatctgcg cgcacagag aaacgttgct 360

ES 2 583 639 T3

	ggtaaaaagg agcgcccggc tgcgcaatat atatataggc atgccaacac agcccaacct	420
	cactcgggag cccgtcccac caccccaag tcgctgcct tgacggcata ctgctgcaga	480
5	agcttcatga gaatgatgcc gaacaagagg ggcacgagga cccaatcccg gacatccttg	540
	tcgataatga tctcgtgagt ccccatcgtc cgcccagcgc tccggggagc ccgccgatgc	600
10	tcaagacgag agggccctcg accaggagg gctggcccgg gcgggactg gcgtcgaagg	660
	tgcgcccgtc gttcgcctgc agtcctatgc cacaaaaca gtcttctgac ggggtgcgtt	720
	tgctcccgtg cgggcaggca acagaggtat tcaccctggt catggggaga tcggcgatcg	780
15	agctgggata agagatactt ctggcaagca atgacaactt gtcaggaccg gaccgtgcca	840
	tatatttctc acctagcgcc gcaaaaccta acaatttggg agtcactgtg cactgagtt	900
20	cgactggtag ctgaatggag tcgctgctcc actaaacgaa ttgtcagcac cgccagccgg	960
	ccgaggacc gagtcatagc gagggtagta gcgcgcc	997
25	<210> 92	
	<211> 753	
	<212> ADN	
	<213> Prototheca moriformis	
30	<400> 92	
	actaattgca atcgtgcagt aatcatcgat atggtcacaa gtagatcccc tactgacacc	60

ES 2 583 639 T3

	ctctcgtagc	120
	agcagacccg ggccgatctg ggatacgagc cggccctcca cctgcgctcg aggtggaatc	180
	aagtaaataa ccaatacact tttcgacacc acacagagtt gcacggacgg tggcgtacct	240
5	ctacgctcgc gctcttcacg cgctggacga ccgcacgcat gagcccgggt ggcttggctt	300
	gggctgcaaa aatgcacaac aaacaagtat cagacgctca tggatgcaca cgcgctccca	360
10	agcacgctca gactaaatat tacagtagct cgtatctgat aagatatcga gacataccgc	420
	tcaactcacc cgcaactgc gccccgccag gtgatgca cagggcccca ccatgcatc	480
	catcgcacgc ctctcgagg gcgctatcac gtggccggag agcgttcaca gcgtaccca	540
15	ctgtatctgg gcggtatgcg gtccgtcaac atggagacag ataccgcac caccacctg	600
	caagctcttc catattggaa gtagaaaatt gtaattgtat catcgcacga ggggccaact	660
20	tgccgtcggc gagctgggag acgaacacca cctggacggt gtcgagactc gctcgtgccg	720
	tgcgcccggc cgctgggtat ccagaccgct gcc	753
25	<210> 93	
	<211> 1122	
	<212> ADN	
	<213> Prototheca moriformis	
30	<400> 93	
	caacgacaac cagcaggcaa ctcggtcagc gaccaaacac gcgagtcaaa ttggtgctg	60

ES 2 583 639 T3

	ttcttgcctt gtctatttac tgtgatagca agactgtcgg tcagtcaata ccgcggtgcg	120
	cacgtcgggg tgccaagcct agcagagcac gggacggctg gtgctgtgcg ccagctcagc	180
5	tcgcttcgcg accaattgta ggaccggcaa agtcaccaaa acatgccagc ggtgcgattc	240
	aattggatcat gagctctaca aaattgtttt gtgctgtgcg caggtatcca acggcgcggc	300
	agagaaagt tgacagctct cgatttcac tcggaaaaat ggggagaatt tatgacacac	360
10	aagtgcgag gcggccagc cgccagcat attctggcgt gacctgggcc gccacaaaa	420
	tgcttgatg cactctaaaa taattatatt tgccatgaac aaggaagag ttaccgcacc	480
	cagccctaga cttgggcgcc cgagcaaggt tacgtcaagc caccttcgcc catcgcccaa	540
15	ctccgtattc cccgacagcc gcacgtggcc ctcgccgaa tgaaccctga atcgccatca	600
	cgccacgcgt tcgccaatcg ttccgctctc tggcttcac ggcttgcgcc ttcacgtcgt	660
20	ggtcacgaca gtgcattcat acttccattt gcacctcggc acacactttt acgcatcgcc	720
	tacccttgct gcggcagtct agggtcactt tgcagccatg ggacagtgct acaccaccgt	780
	cgggtgcgcaa agctatttca agtgaaccgt gggcggaaaa aaggaatgta cactgtctca	840
25	accgactcct acaattgttt accatgcaga tcagagctcg acggccatca tcgagcaggt	900
	gtggggcctt ggtggcgcg cgcggggcc cagggcgtcg caggcattga tggcactctg	960
30	agactttcgc acgcgcatga gggaccccat caagagaaga gtgtgtcttt atgtcccat	1020
	tcatgatgat gtatcttctg attgtcgcag tttggcaagt ttaaccggat cgccgctcca	1080

ES 2 583 639 T3

	ggtgtggcgt ggcggatfff tctaggggtg cttgagcagt cg	1122
5	<210> 94 <211> 574 <212> ADN <213> Prototheca moriformis	
10	<400> 94 ggcccagggc cctgcggatg gcccacacca gatctagcct ctcttatgcc atgcccgcct	60
	cgctgcccgt cgtatcccc cgccgatccg cgcgtagggg accgcccct gaccacgcc	120
15	acgaaagagc tttgctctc aatttctcgc caacagaacc gtatcaaagc ctcaacgcct	180
	atcccgaaca atccgtattc acaccaaadc gagtataccg gactggtttg cctagtcttg	240
	aaggaaatga tcccgtccat gctcgggaagg gggagcgggc ggaggatcct actcatctct	300
20	gaaatgggat tgggccgaag atggggtggg caagcacgtg ccaaaccaca gcgagttgct	360
	gacgagcagg ctcatccaat cccccggcga atcctccctc acgccccgca tgcatacaag	420
	tccctccac acgccccctc ccatccattt tcgcctggtc cgaacgcgag cggcgtcgag	480
25	gcgaccact tgctccgag cgccgtctgg gtctccacc cacagcggct ttgctgccag	540
	aggcaccacc cttgccccac ctctcttgc agcc	574
30	<210> 95 <211> 1096	

ES 2 583 639 T3

<212> ADN

<213> Prototheca moriformis

<400> 95

5 ccaggcaggc ggtagggttg ccgattgctt gagcgaattg gaagatataa ttttttgtgg 60

tgctccctgga cgctgtttgt ggcgctcctt tttggagaag attgcgtggg ggagctttcc 120

atgtaccacg cttccttctg aaaggattct ggccgagtcc tgatgagccc aaagaaaaca 180

10 cctgcctttc agtgctggca ctctgaaaac gtcaacagat gattatacat gtcacaaaag 240

gcagccgatt aggaacggga gctctggccg ttcgtttggc tgctgggct gattgaagtg 300

15 atccaccctg ttcgaatgaa ggcggtcgag tcgaattatc gaccggagct gtcgggaagg 360

cgctccggggc agagtgaggt gctgcggcct ggttgctggt caaaaagacc ccggtagccc 420

aacaatcacg aacgaaagga atataattgc ttgcatacta tacattcagt ttctatgtgg 480

20 cgggtagaca agtctcatgg gcttctaaag gctgtccctt gaaggctact tataaaaact 540

tgctgcgcca tggcacggat cgcgcttgcg caggctgcaa ccctgcgcgc aagggtcaaat 600

acacagcaaa agatactaac agaatttcta aaaacattta aatatttggt tcgaccagcc 660

25 aattgtggtc gtaggcacgc aaaagacttt gttttgcgcc caccgagcat ccacgctggc 720

agtcaagcca gtccgatgtg cattgcgtgg cagcatcgag gagcatcaaa aacctcgtgc 780

30 acgcttttct gtcaatcatc atcaaccact ccaccatgta taccgatgc atcgcgggtgc 840

gcagcgcgcc acgctccca gaccgccc aaaaccagc agcggcgaaa gcaaattctt 900

ES 2 583 639 T3

acttgcccga aacccccgagc agcggcattc acacgtgggc gaaaaccca cttgccctaa 960

caggcgtatg tctgctgtca cgatgcctga caacggtatt atagatatac actgattaat 1020

5 gtttgagtgt gtgcgagtcg cgaatcagga atgaattgct agtaggcact ccgaccgggc 1080

gggggcccag ggacca 1096

10 <210> 96
<211> 1075
<212> ADN
<213> Prototheca moriformis

15 <400> 96

ggcccacagg acgcgcgtca aaggtgctgg gcgtgtatgc cctggtcggc aggtcgttgc 60

tgttgctgcg ctcgtggttc cgcaaccctg attttggcgt cttattctgg cgtggcaagc 120

20 gctgacgcc gcgagccggg ccggcggcga tgcggtgtct cacggctgcc gagctccaag 180

ggaggcaaga ggcgccgat cagctgaagg gctttacacg caaggtacag ccgctcctgc 240

25 aaggctgctg ggtggacttg aacctgtagg tcctctgctg aagttcctcc actacctcac 300

caggcccagc agaccaaagc acaggctttt caggctcgtg tcatccactc taaaacactc 360

gactacgacc tactgatggc cctagattct tcatcaaaa tgcctgagac acttgctcag 420

30 aattgaaact ccctgaagg accaccagag gccctgagtt gttccttccc cccgtggcga 480

gctgccagcc aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca 540

ES 2 583 639 T3

ggtcatggga ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc 600
 taatcagcta tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt 660
 5 gatccttcgt gtacggggccc ttcctcaac cctaggtatg cgcgcatgcg gtcgccgcg 720
 aactcgcgcg agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat 780
 10 gcgcggcgcc tttcttgca taatttatgc aatggactgc tctgcaaatt tctgggtctg 840
 tcgccaacc taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta 900
 ccgactacc taatcagc cggctgcct gacgccagcg tccacttttg cgtacacatt 960
 15 ccattcgtgc ccaagacatt tcatttggt gcgaagcgtc cccagttacg ctcacctgtt 1020
 tcccacctc cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagcc 1075
 20
 <210> 97
 <211> 772
 <212> ADN
 <213> Prototheca moriformis
 25
 <400> 97
 tcaccagcgg acaaagcacc ggtgtatcag gtccgtgtca tccactctaa agagctcgac 60
 tacgacctac tgatggcct agattcttca tcaaaaacgc ctgagacact tgcccaggat 120
 30 tgaaactccc tgaaggacc accaggggcc ctgagttgtt cttcccccc gtggcgagct 180

ES 2 583 639 T3

gccagccagg ctgtacctgt gatcggggct ggcgggaaaa caggcttcgt gtgctcaggt 240
 tatgggaggt gcaggacagc tcattaaacg ccaacaatcg cacaattcat ggcaagctaa 300

tcagttatth cccattaacg agctataatt gtcccaaaat tctgggtctac cgggggtgat 360

5 ccttcgtgta cgggcccttc cctcaaccct aggtatgctc acatgctggtc gccgctgcaac 420

gcgctgaggt gccgaggggt tgggacgggc cgtcccgaat tgcagttgca cccggatgctg 480

10 tggcacctth tttgctgataa tttatgcaat ggactgctct gcaaaattct ggctctgctg 540

ccaaccctag gatcagctgt gtaggatttc gtaatcattc gtcctgatgg ggagctaccg 600

actgctctag tatcagctcg actgctctgac gccagctgctc acttttctgctc acacattcca 660

15 ttcgtgctca agacatttca ttgtggtgctg aagctgctcc agttactgctc acctgatccc 720

caacctctt attgttctgt ctgacagctg ggcccagaggt ccggtctgctg cc 772

20

<210> 98
 <211> 991
 <212> ADN
 <213> *Protoheca moriformis*

25

<400> 98

cgaaggggtc tgcactgatt ctgctggtct ggaggctgctc gctgaaatgc 60

tctgctctgt cgggtctctg ctggggctgctc cagagatctc acctgctcac acgcaactgc 120

30 cgcactctgt gccctgctc tggctgctc atgctgctc tctccctgctc taccctctcc 180

ES 2 583 639 T3

	tcatgtgatc tttccacacg agtgacgcag gtgcgcggag tggaggggaat caggacgttt	240
	tcaaggtacc tgctcgagcc gtaccaacag ctgccgcccc gcaaggaaga gatcgaggca	300
5	gagattgccc ggctggaggc ccggataacg gagctcaaga gcaagctgtc cgagtgagac	360
	cgcccaggtg cacgtgtcga ctcgctatga catgtactcg acacaacatg aggaattcat	420
	cgaatttgta ggaagcgggc attggtacgg gagtgggaaa gcgaaaaaac ctccctccgg	480
10	cagtgccatc tgccggagtc gaacgttgat agggttctcg tgacagggtg tgaccttca	540
	gccttgcatc aattaaacgc tatagacatt atcagtaacc gtgaatcccc cattggatgc	600
	cacccgcgcg accattgggg acctgcatta cagatctagg tgagatgaca gcgaggcaac	660
15	ttcggccccg ggcccagctt gcggcgcacc aatattggtc acgggaagcc acacaccgac	720
	cataaatgaa tacttgtaag ctatgtcaac cgatcaatgg cgtcgaaagt gtgccacgag	780
20	gatccatctg gcggggcggc gtggcgcaca agcgcagtcg caatttctcg gaccatctg	840
	acctaggccc agcgcgcgg gagaaatccc cggcgggtcc tccacgcagt aaccctaag	900
	agtatcgagc gccgaccatt tacaccatcg cccccgaaat ccttccgaca ttattattat	960
25	cttttagatc ttggaacaga ctctgccaac c	991
	<210> 99	
30	<211> 1347	
	<212> ADN	
	<213> Prototheca moriformis	

ES 2 583 639 T3

<400> 99

	agagagcggg ggtgggggtg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg	60
5	cgactcgcac acggtccagt tcccccccc tccgcccaaa cgcaagcctc ccatcttgat	120
	gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata	180
	ataacaatgc cctgtgtcaa gtgcattcct aaaaaaattc tgtccaacc aacaatcca	240
10	cctgaaatac caccagcct gccagtaga ctcttccaat accatctccc tacctccag	300
	cgcaagcgac ccccatgcbc gaccaggctc gaaagtgatt tatgacttga gacgagcgag	360
	tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt	420
15	gcacggcaac gcacggctc gccaaaccga ccagccagga cctcgactac tctaccgca	480
	attgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact	540
20	tgatgccaat tttgacattc gggttgctcg ttggctgcbc ccacatcggc cgtgagtgca	600
	gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta	660
	catcaacacg acgtgttgcg taatcccgcc cggctgcbc tcgtgccaac ccattcgcga	720
25	tggatggtcg gaaaatggtg tgccaactgc cctgaggag gctctcgcga aacgggcacg	780
	tccctgaaac cgaaactgtg gccttgctgt cggccacgca agcacgtgga ccctaaacac	840
30	caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc	900
	atcgcacact gccattataa tgcatttagc tcggcgacaa gtttagaaaa ggcaggctgc	960

ES 2 583 639 T3

attgttccat ttcgccgtgg cggcgtgggt gccatttta cgaggtttgg gctcccgggc 1020

agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080

5 gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140

acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200

10 ccgcgacgacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260

cacaattgcy ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat 1320

atthttccagt cgtcagcgaat atcaagt 1347

15

<210> 100

<211> 1180

<212> ADN

20 <213> *Prototheca moriformis*

<400> 100

gatggtgggg tgtctgcctt gggctgggtg atggaggctg gtggtgcgcg ggthttcctga 60

25 tgcattctat ctacgcagtg tcatgggtgac cattccacac accagtacac ctttacacta 120

aggatccatc cctccttccc tcttcaggac tacatggacc ccacgagcta ccgaccgggc 180

tttctcaaaa acgtcaaggt catgtttgac atgcgggacg tgggtggacga cgtgcaaggt 240

30 gcgtccggag tgcgcgcaaa tgagcaagtc gggcaatgtg tccgggtggg caccggggct 300

ES 2 583 639 T3

	ggagatccgc gatccccgag aaaacgccgt accaccccc gcgctattcc ctcgattgcg	360
	cgcagatgtg gtgaccgaca cgggggacaa cctggcggac atggggcgcc ggacctggaa	420
5	gcacgccaag tcgcacacgg ggaggctcgt gcagtcccc ccatcgtacc tcaaggtct	480
	ctttggtcgc gatccaaagt acgctggtgg catggcatgc ccgaaatgaa catcatgtgt	540
	gatctccgat tgccaatggc cacctccacg gaccaccttg caggcggag cgcaatccag	600
10	ggcccagacc tgacgaggac ggagactcct cgtccagcgc ggggtcccc acccgacga	660
	gcagccgacc cctgctaacc cggcaacgat cggaccagca accttgctgt agttccgatc	720
15	cgtgatgacg ggcatgccc ccgctcgatc cgctttgatg actgtctatt atttgcgcg	780
	agccccctcg gaacctacc ccgctcttgc aagccccctg catcggagat cctcgtgcg	840
	ccgcatgac cccactggat tgccaacat ccttctttat cgtgtaaaat gtgattcctc	900
20	ggctgcaatc gactggcctt cgcttctggc cccaagagg ctcgaacgtg cggcagcgag	960
	ggcgctgaca cacccaagcc ctagggcttt caacgtcggc tgccaggccg gataggggga	1020
	tcgcctcctt tccaccacc acctacgagg gattcgagtc ggcttccagc tcagctattc	1080
25	ggccgcgccc ccggccctgc agacgtcctc cagtttccga acaggtcgct ctcagaacac	1140
	ctgccgcggc tgcgatacgg caggctctca aagcgtcgc	1180
30		
	<210> 101	
	<211> 1263	

ES 2 583 639 T3

<212> ADN

<213> Prototheca moriformis

<400> 101

5 cgcgtaggagc ggtgtagtgcg gatgccgagc gcctgccaag gccttttgta tgcctggcct 60

gggaagtttc ctgactgaag catcttcaag atgctctctc acgaccagcg acaccaacac 120

cgtcactttt tgcccctcct gccgcaggtg ccactttcta ctttgacgtc ttctccaggc 180

10 ggtacattgc gggactgagc gccaattcgg ccaagaacag cgctgtcgac ttgaggaggc 240

aggggtccgt cgactctgcc gagtgacacg ccttcgacct gactgtacta cggcctgctg 300

15 aagagtgggt ctgcccggcc ggcgtgaccg gccctgtgcc cacaatcgac catctattcg 360

ctccttgtca tctggcgccg tcaattgccc gcgacttgac ggcaactggc tcgatcgagt 420

cgatttgaag aagcacgttt tgcctacag ggccgaggtc cgttaccaac gtggttctcg 480

20 ttaggttttc gtcgggaggt ggtgagcga cgtgccgatg ccatcccggc aaacccagc 540

aaggtcgcca gtctggttct gacgcaatag agtgagtttt gggccagtct aaaaattcgt 600

25 ctggcatgac gtggctccac atcgtagccg gagcctgcct tggtaatgtg aggcaccggc 660

gccaactcca ttatggcagg catcgagcgc gcaggtgagt acatgacctt ccgtgaattg 720

ggaaggcgag cttgtgtaac gcctgagatc gtgccagtga ggcatcgtaa actcaaaata 780

30 tttttagaaa agtgtctgat gcctggtgag gctgagtagg gcaagggcaa gcccttggca 840

gatgggtaat gggtagcagc ctcaaacag caacccgagc tcccccttag ggcccctgag 900

ES 2 583 639 T3

gctcgatggc agggccagcg agcccgcggc caaagggcgc catcccacgg tcgcccacg 960

actccacggg tcctatacct catcttgaat ggcactaaaa actatagaat atcgggcact 1020

5 ggtgggctgc tggggtacag ctggccgagc gcagtggcaa accctagggtc ccgcctcaag 1080

ggcgattccc gggatcaatga cacgcaagca agatcacatg gcgcgggtccg cctcgcggct 1140

10 ccacaccag gccctagttt cgcaacccat aaatatcgcc ccgataccat cataagccag 1200

caaataattt tttatcagag ttccaaacct cctcagctgt gggaaaccag cccactctga 1260

acg 1263

15

<210> 102

<211> 1400

<212> ADN

20 <213> Prototheca moriformis

<400> 102

ccgagcagtt catggccaag tacaaggact agagaccgga ggtcggtagg ctgaatggag 60

25 ctggcgtcgt cgtgcgcgac gtgcacgcga tgcgatacta cgaccccaca aacgcatgcc 120

tcccatcttg atgcctttcc ggccatttat actatttctc atttcgctgt aacatcttga 180

ataatagaat tgcctgtgt caagtggatt ccaagaaata ttctgtcca acaaaacaac 240

30 ccaacctgaa aacaacctca aataccacca gccctgcca cctgcccagt acacttttcc 300

ES 2 583 639 T3

	aataccatct ccctaccttc acgcgcaagc ggcacccatg cgcgaccagg ctcgaaagga	360
	tttcacgact caggacgagc gagtggcggc gcgaccgcct gcctgttcgt cacgtgccgt	420
	acgtcggcga ccgctagagc tttgcctggc aacccccggc ttcgtcaacc cggccagcca	480
5	ggatctcgac cactctaccg cgaaatcgcc tcaagaagtc gccaaaagtg ccgtacacca	540
	tgcttcgcag cgctgttcaa acttgatgcc aatcttgaca atcaggttgc tcgttggctg	600
10	cgccacatc ggccgtgatt gcagcaggcg gggatcggac acggaggacg cggcgtcacg	660
	ccgcgaacgc agcccgtaac tctacatcaa cgcgatatgt tgcgtaatcc cgcccggctg	720
	cgcattgtga caaccattc gcgatggatg gtcggaaaat ggtgtgccaa ctgccctgag	780
15	ggactctctc gcgaaacggg cacgtccctg tatccgaaac tgtggcatgg ccttgtcgac	840
	cacgcaagca cgtggaccct aacaccacga aaataagtaa aaaaggttga catcctctac	900
20	gagcgaattg tttgctcgac cttcatcgc aactgtcat tataatgcat ctagctcggc	960
	gacaagtta aaaaaggcag gctgcattat tccatthtgc cgtggcggca tgggtgccca	1020
	ttttatgagg tttgggctct tgggcagcga ccgagccagg ttgagtcctt ctcgcccgtc	1080
25	gacaacgttc caaagcccat aagtggctaa taaacaactt gatggtacct gtacactgcc	1140
	agttccttct tccccggccg aggtttacac gtgatggcca tggcttcgcg tttcaggctg	1200
30	acttccatt ccgactttcc agagggtccg cggacgccgg gggttggctg cgtgaggccc	1260
	acccttggtt ccccgcgtcc cgacaaacac aattgcgta cataaggggg aagccgcccc	1320

ES 2 583 639 T3

	ccgttcagag tgcaaacatc tttcattata tttttcagtc gtcagcgaaa tcaagtatgt	1380
	cgctgacagg catgaaggcc	1400
5		
	<210> 103	
	<211> 3681	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
15		
	<400> 103	
	gccctttgtc atcgttggca tgctttttgc gtatgtacca tatgttgaat gtataatagc	60
	aacggttgac cgtctgagat gcgagctttg ggtcttgca aatgcgtggc cgcacggctc	120
20		
	cctcgcacc agccccgagg cgtcgcgcac ctggcgagga gcagaccac gccaagaaag	180
	tctagtccag catgtaacaa catcaggcaa tgtgacgttt tcggttcccg atttctctgc	240
25		
	cgctctttga cggcaggcac gggcgagcaa ccggcggcgc tcgcgtcagg cacgatggat	300
	ggggcgctgc ccacctgtca atgtaccca ccagtctgtc gatcgtaca agcaaccttg	360
	tgctccacat tcccacttgc agacagtcta gtcgattttg ccaagctgga tgtgaggatt	420
30		
	ggccatatct tggaggccaa gattcacccg gatgctgatg ggtacgtacg cgagccaggc	480

ES 2 583 639 T3

aggcagctgc gttgactttc tgattggcac aaagctttgg ctactctcaa taccaaccac 540

gtgccccttc tgcacacctg cttccttctg atgaccactc gccacgcatg tcgcagtctg 600

5 tacgtcgagc agatcgacct cggcgaggag gggggccctc gcaccatcgt gagtggcctg 660

gtccggcacg tgaccctgga ggaccttgtc ggccggcggg tgggtggtgct ggccaacctc 720

aagcctcggg gcatgcgcg ggtcaaatcg gctgggatgc tgctctgcmc cgccaacgcm 780

10 gatcacaccg cgggtggagcc gctgcmggtc ccggacgcmc ccgtgacggg ggagcmggtc 840

tgggcmgggg acgaggcact cctgtccacg gagcctgcca cacccaatca ggtaaggaca 900

15 cgttattggt gcmcatggtg cgaatgcmgt gtctgacctg ctgtgggtat gtggttggg 960

attggaacc gaatgaggc cgttcaggat tgagcccttg gccccacct gctcctcctc 1020

tcacgcccgc aggtccagaa gaagaaaatc tgggaggcag tacagccgct gctgagagtg 1080

20 aacgcccagg ggatcmctac tgtggcagga gaggctatgg tgaccagtgc ggggacctg 1140

accgcmcca cmctggttga cmccgcmatt tcctgacgcm agcmactgat tcttgacctt 1200

25 tgagaagcca ccacagcacc attttcattg ttcctccttg atttcagtac gacttctcac 1260

catttcagta ctgtaggacc cccaaaatag tgtgatcacg ctcgcaaggc acctgtgtga 1320

tcacggggaa gggcgaattc ctttcttgcg ctatgacact tccagcaaaa ggtagggcmg 1380

30 gctgcmgac ggcttcccgm cmctgcmatgc aacaccgatg atgcttcmgac cccccgaagc 1440

ES 2 583 639 T3

	tccttcgggg ctgcatgggc gctccgatgc cgctccaggg cgagcgctgt ttaaataagcc	1500
	aggccccga ttgcaaagac attatagcga gctaccaaag ccatattcaa acacctagat	1560
5	cactaccact tctacacagg ccactcgagc ttgtgatcgc actccgctaa gggggcgcct	1620
	cttcctcttc gtttcagtca caacccgcaa acggcgcgcc atgctgctgc aggccttcct	1680
	gttcctgctg gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga	1740
10	ccgccccctg gtgcacttca cccccaacaa gggctggatg aacgaccca acggcctgtg	1800
	gtacgacgag aaggacgcca agtggcacct gtacttccag tacaaccgga acgacaccgt	1860
15	ctgggggacg cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga	1920
	ccagcccatc gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt	1980
	ggtggactac aacaacacct ccggcttctt caacgacacc atcgaccgc gccagcgctg	2040
20	cgtggccatc tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct	2100
	ggacggcggc tacaccttca ccgagtagca gaagaacccc gtgctggccg ccaactccac	2160
25	ccagttccgc gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc	2220
	ggccaagtcc caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa	2280
	gctggagtcc gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct	2340
30	gatcgaggtc cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat	2400

ES 2 583 639 T3

	caaccccggc gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg	2460
	caccacttc gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta	2520
5	cgccctgcag accttcttca acaccgaccc gacctacggg agcgccttg gcatcgcgtg	2580
	ggcctccaac tgggagtact ccgccttcgt gccaccaac ccctggcgct cctccatgtc	2640
	cctcgtgcg aagttctccc tcaacaccga gtaccaggcc aaccggaga cggagctgat	2700
10	caacctgaag gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc	2760
	caccaacacc acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg	2820
15	caccctggag ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt	2880
	gttcgcgga ctctccctct ggttcaagg cctggaggac cccgaggagt acctccgat	2940
	gggcttcgag gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt	3000
20	cgtgaaggag aaccctact tcaccaaccg catgagcgtg aacaaccagc cttcaagag	3060
	cgagaacgac ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct	3120
25	gtacttcaac gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc	3180
	cctgggctcc gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca	3240
	ggtgcgag gtcaagtgat taattaactc gaggcagcag cagctcggat agtatcgaca	3300
30	cactctggac gctggctgtg tgatggactg ttgccgccac acttgctgcc ttgacctgtg	3360

ES 2 583 639 T3

	aatatccctg ccgcttttat caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt	3420
	ttgcgagttg ctagctgctt gtgctatttg cgaataccac ccccagcatc cccttcctc	3480
5	gtttcatatc gcttgcaccc caaccgcaac ttatctacgc tgtcctgcta tccctcagcg	3540
	ctgctcctgc tctgctcac tgcccctcgc acagccttgg tttgggctcc gcctgtattc	3600
	tcttggact gcaacctgta aaccagcact gcaatgctga tgcacgggaa gtagtgggat	3660
10	gggaacacaa atggaaagct t	3681
	<210> 104	
15	<211> 3850	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
20	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 104	
	tttgccccg ctttcagct ccgatctgc tggcgtccgc cgcgagacgt gacatcgac	60
25	gtcgccgga gcgccagctt gatcacttg cagggggccg tgctctacaa ataccaggcc	120
	ccgcggcggc cagttcgac atccaatacc tgccgagcca tcttgcttac actttttatc	180
30	gactcctcta ctctgttcgc gagagcgtc ggtccaggct tggaattcgc cgaattcagc	240
	tcgatcagtc gcttctgcaa ctgatctcgg ccgttcgcag actgcctttt ctcagcttgt	300

ES 2 583 639 T3

caggtagcga gttgttgttt tatatttatt cgatttcac tgtgttgcac gtcttgttcg 360
 tgctgtgcgt tctttctggg ccgcgctgtc gggtcgcatg ggctagctgt actcatgtta 420
 5 gtcacgcccg tccgacctg ttcgaggaag gccccacact gagcgtgccc tctttctaca 480
 ccccttgctc agaaattaga tagaaagcag aattcctttc ttgcgctatg acacttccag 540
 10 caaaaggtag ggcgggctgc gagacggctt cccggcgctg catgcaacac cgatgatgct 600
 tcgaccccc gaagtcctt cggggctgca tgggcgctcc gatgccgctc cagggcgagc 660
 gctgtttaa tagccaggcc cccgattgca aagacattat agcgagctac caaagccata 720
 15 ttcaaacacc tagatcacta ccacttctac acaggccact cgagcttggt atcgactcc 780
 gctaagggg cgctcttcc tcttcgtttc agtcacaacc cgcaaacggc gcgccatgct 840
 20 gctgcaggcc ttctgttcc tgctggccgg cttcgccgcc aagatcagcg cctccatgac 900
 gaacgagacg tccgaccgcc ccctggtgca cttcaccgcc aacaagggt ggatgaacga 960
 cccaacggc ctgtggtacg acgagaagga cgccaagtgg cacctgtact tccagtacaa 1020
 25 cccgaacgac accgtctggg ggacgccctt gttctggggc cacgccacgt cgcacgacct 1080
 gaccaactgg gaggaccagc ccatcgccat cgccccgaag cgcaacgact ccggcgcctt 1140
 30 ctccggctcc atggtggtgg actacaacaa cacctccggc ttcttcaacg acaccatcga 1200
 cccgcgccag cgctgctgg ccatctggac ctacaacacc ccggagtccg aggagcagta 1260

ES 2 583 639 T3

	catctcctac agcctggacg gcggtacac cttcaccgag taccagaaga accccgtgct	1320
	ggccgccaac tccaccagt tccgcgaccc gaaggtcttc tggtagagc cctcccagaa	1380
5	gtggatcatg accgcgcca agtcccagga ctacaagatc gagatctact cctccgacga	1440
	cctgaagtcc tggaagctgg agtccgcgtt cgccaacgag ggcttcctcg gctaccagta	1500
10	cgagtgcccc ggcctgatcg aggtccccac cgagcaggac cccagcaagt cctactgggt	1560
	gatgttcac tccatcaacc ccggcgcccc ggccggcggc tccttcaacc agtacttcgt	1620
	cggcagcttc aacggcacc acttcgaggc cttcgacaac cagtcccgcg tggtagactt	1680
15	cggaaggac tactacgcc tgcagacctt cttcaacacc gacccgacct acgggagcgc	1740
	cctgggcatc gcgtgggcct ccaactggga gtactccgcc ttcgtgcca ccaaccctg	1800
20	gcgctcctcc atgtccctcg tgcgcaagtt ctccctcaac accgagtacc aggccaacc	1860
	ggagacggag ctgatcaacc tgaaggcca gccgatcctg aacatcagca acgcccggcc	1920
	ctggagccgg ttcgccacca acaccagtt gacgaaggcc aacagctaca acgtcgacct	1980
25	gtccaacagc accggcacc tggagttcga gctggtgtac gccgtcaaca ccaccagac	2040
	gatctccaag tccgtgttcg cggacctctc cctctggttc aaggcctgg aggaccccga	2100
30	ggagtacctc cgcattggct tcgaggtgtc cgcgtcctcc ttcttctgg accgcccga	2160
	cagcaagggtg aagttcgtga aggagaacct ctacttcacc aaccgatga gcgtgaacaa	2220

ES 2 583 639 T3

ccagccctc aagagcgaga acgacctgtc ctactacaag gtgtacggct tgctggacca 2280

gaacatcctg gagctgtact tcaacgacgg cgacgtcgtg tccaccaaca cctacttcat 2340

5 gaccaccggg aacgccctgg gctccgtgaa catgacgacg ggggtggaca acctgttcta 2400

catcgacaag ttccaggtgc gcgaggtcaa gtgattaatt aactcgaggc agcagcagct 2460

10 cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 2520

ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 2580

tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accacccccca 2640

15 gcatccccct ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 2700

tgctatccct cagcgctgct cctgctcctg ctactgccc ctgcacagc cttggtttgg 2760

20 gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcaactgcaat gctgatgcac 2820

gggaagtagt gggatgggaa cacaaatgga ccgacacgcc cccggcccag gtccagttct 2880

cctgggtcct ccagaggccc gtcgccatgt aaagtggcag agattggcgc ctgattcgat 2940

25 ttgatccaa ggatctcaa tcggtgatgg ggactgagtg cccaactacc acccttgcac 3000

tatcgtcctc gcactattta ttcccactt ctgctcgcct tgccgggcga ttgcgggcgt 3060

30 ttctgccctt gacgtatcaa tttcgcccct gctggcgcga ggattcttca ttctaataag 3120

aactcactcc cgccagctct gtacttttcc tgcggggccc ctgcatggct tgttcccaat 3180

ES 2 583 639 T3

gcttgctcga tcgacggcgc ccattgccca cggcgctgcc gcatccatgt gaagaaacac 3240

ggaagagtgc gaagactgga agtgaattaa gagtataaga agaggtacca agggattctc 3300

5 aggtgctctt aggaacggct tttccttcgc caagagaaac tgctactgct cgtgtcgcca 3360

cggtgggtcaa gccgccccat ctgcatcca ccaggccccat ccgaggactc gcgatcagcc 3420

10 tgctggatcc ggactgccga cctgaccgct cgcattcacc attacaacc tccaattgga 3480

caccactccc acgtcctaaa gttcacatg caagctgatc gatcgcattc gccgatgcac 3540

tcgcctgcca cagaggtgtg cgcttcggac tagcgtgcag gcgccccgag gccaccagca 3600

15 tgcaccgatg gaagcgggca cggccgctgc tccaggctgc tggctcgctc agaccatag 3660

caacctccgc tgcgtcccta aatgtcacac agagcgtctt tgatgggtac ggatgggaga 3720

20 gaatctgatt gggcattgct ggtgcagtgc aggaagatgg caagtcaca gtcagtcatg 3780

ctgtacaaac tgggtgcctcg tagtattgac tcgtatagtg catagtatca tgcattgctg 3840

ttacttgcaa 3850

25

<210> 105

<211> 3108

<212> ADN

30 <213> Secuencia Artificial

<220>

ES 2 583 639 T3

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 105

5	tttggccccg ctttccagct ccggatctgc tggcgtccgc cgcgagacgt gacatcgcac	60
	gtcgccggga gcgccagctt gatcacttgg cagggggccg tgctctacaa ataccaggcc	120
	ccgcggcggg cagttcgcac atccaatacc tgccgagcca tcttgccctac actttttatc	180
10	gactcctcta ctctgttcgc gagagcgctc ggtccaggct tggaattcgc cgaattcagc	240
	tcgatcagtc gcttctgcaa ctgatctcgg ccgttcgcag actgcctttt ctcagcttgt	300
15	caggtagcga gttggttgtt tatatttatt cgatttcac tgtggttgcg gtcttggttcg	360
	tgctgtgctg tctttctggg ccgcgctgctc gggctcgcag ggctagctgt actcatgtta	420
	gtcatgccgg tccgaccttg ttcgaggaag gccccacact gagcgtgccc tctttctaca	480
20	ccccttgctc agaaattaga tagaaagcaa tgctgctgca ggccttcctg ttctgctgg	540
	ccggcttcgc cgccaagatc agcgcctcca tgacgaacga gacgtccgac cgccccctgg	600
25	tgcacttcac cccaacaag ggctggatga acgaccccaa cggcctgtgg tacgacgaga	660
	aggacgcaa gtggcacctg tacttccagt acaaccgaa cgacaccgctc tgggggacgc	720
	ccttggtctg gggccacgcc acgtccgacg acctgaccaa ctgggaggac cagcccatcg	780
30	ccatcgcccc gaagcgcaac gactccggcg ccttctccgg ctccatggtg gtggactaca	840

ES 2 583 639 T3

acaacacctc cggcttcttc aacgacacca tcgacccgcg ccagcgctgc gtggccatct 900
 ggacctataa caccgccgag tccgaggagc agtacatctc ctacagcctg gacggcggct 960
 5 acaccttcac cgagtaccag aagaacccccg tgctggccgc caactccacc cagttccgcg 1020
 acccgaaggt cttctggtac gagccctccc agaagtggat catgaccgcg gccaaagtccc 1080
 aggactataa gatcgagatc tactcctccg acgacctgaa gtcctggaag ctggagtccg 1140
 10 cgttcgccaa cgagggcttc ctcggctacc agtacgagtg ccccggcctg atcgaggtcc 1200
 ccaccgagca ggaccccagc aagtcctact gggatgatgtt catctccatc aaccccggcg 1260
 15 ccccgcccg cggctccttc aaccagtact tcgtcggcag cttcaacggc acccacttcg 1320
 aggccttcga caaccagtcc cgcgtggtgg acttcggcaa ggactactac gccctgcaga 1380
 ctttcttcaa caccgaccgc acctacggga gcgcctggg catcgcgtgg gcctccaact 1440
 20 gggagtactc cgccttcgtg cccaccaacc cctggcgctc ctccatgtcc ctcgtgcgca 1500
 agttctccct caacaccgag taccaggcca acccggagac ggagctgatc aacctgaagg 1560
 25 ccgagccgat cctgaacatc agcaacgccg gccctggag ccggttcgcc accaacacca 1620
 cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa cagcaccggc accctggagt 1680
 tcgagctggt gtacgccgtc aacaccacc agacgatctc caagtccgtg ttcgcgacc 1740
 30 tctccctctg gttcaagggc ctggaggacc ccgaggagta cctccgatg ggcttcgagg 1800

ES 2 583 639 T3

	tgccgcgtc ctccttcttc ctggaccgcg ggaacagcaa ggtgaagttc gtgaaggaga	1860
	accctactt caccaaccgc atgagcgtga acaaccagcc cttcaagagc gagaacgacc	1920
5	tgctacta caaggtgtac ggcttgctgg accagaacat cctggagctg tacttcaacg	1980
	acggcgacgt cgtgtccacc aacacctact tcatgaccac cgggaacgcc ctgggctccg	2040
	tgaacatgac gacgggggtg gacaacctgt tctacatcga caagttccag gtgcgagagg	2100
10	tcaagtgacc gacacgcccc cggcccaggt ccagttctcc tgggtcttcc agaggcccgt	2160
	cgccatgtaa agtggcagag attggcgcct gattcgattt ggatccaagg atctccaatc	2220
15	ggtgatgggg actgagtgcc caactaccac cttgcacta tcgtcctcgc actatattt	2280
	cccaccttct gctcgccttg ccgggcgatt gcgggcgttt ctgcccttga cgtatcaatt	2340
	tcgcccctgc tggcgcgagg attcttcatt ctaataagaa ctcaactccg ccagctctgt	2400
20	acttttcttg cggggcccct gcatggcttg ttcccaatgc ttgctcgatc gacggcgccc	2460
	attgcccacg gcgctgccgc atccatgtga agaaacacgg aagagtgcga agactggaag	2520
25	tgaattaaga gtataagaag aggtaccaag ggattctcag gtgctcttag gaacggcttt	2580
	tccttcgcca agagaaactg ctactgctcg tgtcgccacg gtggtcaagc cgccccatct	2640
	gcgatccacc aggccatcc gcggactcgc gatcagcctg ctggatccgg actgccgacc	2700
30	tgaccgctcg catccacat tacaaccctc caattggaca ccaactcccac gtcctaaagt	2760

ES 2 583 639 T3

	tcaccatgca agctgatcga tcgcattcgc cgatgcactc gcctgccaca gaggtgtgcg	2820
	cttcggacta gcgtgcaggc gccccgaggc caccagcatg caccgatgga agcgggcacg	2880
5	gccgctgctc caggtcgctg gctcgctcag acccatagca acctccgctg cgtcacctaaa	2940
	tgtcacacag agcgtctttg atgggtacgg atgggagaga atctgattgg gcattgctgg	3000
	tgcagtgcag gaagatggca agtgcacagt cagtcatgct gtacaaactg gtgcctcgta	3060
10	gtattgactc gtatagtgca tagtatcatg catggtcggt acttgcaa	3108
	<210> 106	
15	<211> 559	
	<212> ADN	
	<213> Chlorella luteoviridis	
	<400> 106	
20	tgttgaagaa tgagccggcg acttatagga agtggcgtgg ttaaggaatt ttccgaagcc	60
	caagcgaag caagttttaa aaatagcgat atttgtcact ttttatggac ccgaaccgg	120
	gtgatctaac cgtgaccagg atgaagcttg ggtaacacca agtgaaggtc cgaactcttc	180
25	gatctttaa aatcgtgaga tgagttgcgg ttagtaggtg aaatgccaat cgaactcgga	240
	gctagctggt tctccccgaa atgtgttgag ggcagcgat gaatgacaaa acaaatagta	300
30	cggtgtaggg gtaaagcact gtttcggtgc gggctgcgaa agcggtagca aatcgtggca	360
	aactcagaat actacgcttg tatacattc atcagtgaga ctgtggggga taagctccat	420

ES 2 583 639 T3

agtcaagagg gaaacagccc agatcaccag ttaaggcccc aaaatgacag ctaagtggca 480
 aaggaggtga aagtgcagaa acaaccagga ggtttgccca gaagcagcca tcctttaaag 540
 5 agtgcgtaat agctcactg 559
 <210> 107
 <211> 1841
 10 <212> ADN
 <213> *Cuphea palustris*
 <400> 107
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 15 tcccggcgct gcatgcaaca cccgatgatgc ttcgaccccc cgaagtcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 20 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgcactc cgctaagggg ggcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 360
 25 cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgccagc gaggcccctc 420
 cccgtgcgcy ggcgcgccag catgctgctg tcggcggtga ccacggtctt cggcgtggcc 480
 30 gagaagcagt ggcccatgct ggaccgcaag tccaagcgcc ccgacatgct ggtcgagccc 540
 ctgggcgtgg accgcatcgt ctacgacggc gtgagcttcc gccagtcggt ctccatccgc 600

ES 2 583 639 T3

	agctacgaga tcggcgccga ccgcaccgcc tcgatcgaga cgctgatgaa catgttccag	660
	gagacctccc tgaaccactg caagatcatc ggcttgctga acgacggctt cggccgcacg	720
5	cccgagatgt gcaagcgcga cctgatctgg gtcgtgacca agatgcagat cgaggtgaac	780
	cgctacccca cgtggggcga caccatcgag gtcaacacgt gggtgagcgc ctcgggcaag	840
	cacggcatgg gccgcgactg gctgatctcc gactgccaca ccggcgagat cctgatccgc	900
10	gcgacgagcg tctgggggat gatgaaccag aagacccgcc gcctgtcgaa gatcccctac	960
	gaggtgcgcc aggagatcga gccccagttc gtcgactccg cccccgtgat cgtggacgac	1020
15	cgcaagttcc acaagctgga cctgaagacg ggcgacagca tctgcaacgg cctgaccccc	1080
	cgctggacgg acctggacgt gaaccagcac gtcaacaacg tgaagtacat cggctggatc	1140
	ctgcagtcgg tccccaccga ggtgttcgag acgcaggagc tgtgcggcct gaccctggag	1200
20	taccgccgcy agtgccggccy cgactccgtg ctggagagcy tcacggccat ggaccctcgy	1260
	aaggagggcg accgctccct gtaccagcac ctgctgcgcc tggaggacgg cgcggacatc	1320
25	gtgaagggcc gcaccgagtg gcgccccaaag aacgccggcg ccaagggcgc catcctgacg	1380
	ggcaagacca gcaacggcaa ctcgatctcc tgactcgagt taattaactc gaggcagcag	1440
	cagctcggat agtatcgaca cactctggac gctggctcgtg tgatggactg ttgccgccac	1500
30	acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	1560

ES 2 583 639 T3

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatattg cgaataccac 1620

ccccagcatc cccttcctc gtttcatatc gcttgcaccc caaccgcaac ttatctagc 1680

5 tgcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1740

tttgggctcc gcctgtattc tcctggact gcaacctgta aaccagcact gcaatgctga 1800

tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 1841

10

<210> 108

<211> 1010

<212> ADN

15 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

20

<400> 108

ggcgcgccca gctgcccgc tggagcatgc tgctggccgc gatcaccacc ctgttcctgg 60

cggccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgc atgctggtgg 120

25 accccttcgg cctgggccgc ttcgtgcagg acggcctggt gttccgcaac aacttcagca 180

tccgcagcta cgagatcggc gcggaccgca ccgccagcat cgagaccctg atgaaccacc 240

30 tgcaggagac cgccctgaac cacgtgaaga gcgtgggcct gctggaggac ggcctgggca 300

gcacccgcga gatgagcctg cgcaacctga tctgggtggt gaccaagatg caggtggcgg 360

ES 2 583 639 T3

	tggaccgcta ccccacctgg ggcgacgagg tgcaggtgag cagctgggcg accgccatcg	420
	gcaagaacgg catgcgccgc gagtggatcg tgaccgactt cgcaccggc gagaccctgc	480
5	tgcgcgccac cagcgtgtgg gtgatgatga acaagctgac ccgccgcatc agcaagatcc	540
	ccgaggaggt gtggcacgag atcggcccca gcttcatcga cgcgccccctgcccaccg	600
10	tggaggacga cggccgcaag ctgaccgct tcgacgagag cagcggcgac ttcattccgca	660
	agggcctgac cccccgctgg agcgacctgg acatcaacca gcacgtgaac aacgtgaagt	720
	acatcggctg gctgctggag agcgcgcccc ccgagatcca cgagagccac gagatcgcca	780
15	gcctgaccct ggagtaccgc cgcgagtgcg gccgcgacag cgtgctgaac agcgccacca	840
	aggtgagcga cagcagccag ctgggcaaga gcgccgtgga gtgcaaccac ctggtgcgcc	900
20	tgcagaacgg cggcgagatc gtgaagggcc gcaccgtgtg gcgcccccaag cgccccctgt	960
	acaacgacgg cgccgtggtg gacgtgcccc ccaagaccag ctgactcgag	1010
25	<210> 109	
	<211> 5472	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	

ES 2 583 639 T3

<400> 109

	ggtagccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	60
5	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	120
	atgggvcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	180
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
10	cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	300
	cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg	360
15	gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg	420
	gtgcacttca cccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag	480
	aaggacgcca agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg	540
20	cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc	600
	gccatcggcc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac	660
25	aacaacacct ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc	720
	tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc	780
	tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc	840
30	gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc	900

ES 2 583 639 T3

	caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc	960
	gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc	1020
5	cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc	1080
	gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc	1140
	gaggccttcg acaaccagtc ccgctgggtg gacttcggca aggactacta cgccctgcag	1200
10	accttcttca acaccgacc gacctacggg agcgccttg gcatcgcgtg ggctccaac	1260
	tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc	1320
15	aagttctccc tcaacaccga gtaccaggcc aaccggaga cggagctgat caacctgaag	1380
	gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc	1440
	acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag	1500
20	ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac	1560
	ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgat gggcttcgag	1620
25	gtgtccgct cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag	1680
	aaccctact tcaccaaccg catgagcgtg aacaaccagc cttcaagag cgagaacgac	1740
	ctgtcctact acaagggtga cggcttgctg gaccagaaca tcctggagct gtacttcaac	1800
30	gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc	1860

ES 2 583 639 T3

gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgag 1920

gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggctg 1980

5 tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2040

atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2100

ttgtgctatt tgcaataacc acccccagca tccccttccc tcgtttcata tcgcttgcatt 2160

10 cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2220

actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 2280

15 taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 2340

tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 2400

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 2460

20 aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 2520

gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca 2580

25 gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggagcctg ccgcctgcct 2640

gccgcatgct tgtgctggtg aggctgggca gtgctgcat gctgattgag gcttgggtca 2700

tcgggtggaa gcttatgtgt gtgctgggct tgcattgccg gcaatgcgca tgggtggcaag 2760

30 agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc 2820

ES 2 583 639 T3

	agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
	ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940
5	acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga	3000
	gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatcaa cgtaaccatg	3060
	agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca	3120
10	tggtgtgcgg atgctgccgg ctgctgcctg ctgctgcacgc cgttgagttg gcagcaggct	3180
	cagccatgca ctggatggca gctgggctgc cactgcaatg tggtggatag gatgcaagtg	3240
15	gagcgaatac caaacctct ggctgcttgc tgggttgcac ggcatcgcac catcagcagg	3300
	agcgcgatcg aagggactgg ccccatgcac gccatgcaa accggagcgc accgagtgtc	3360
	cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca thtagcgtg	3420
20	acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
	gcggcgagcag cagagcggcg gcagcagcgg cgggggccac ctttcttgcg gggtcgcgcc	3540
25	ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcac gcctggagaa	3600
	gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
	cagcaaaggc gcctccccta ctccgatcg atgttcata gtgcattgga ctgcatttgg	3720
30	gtggggcggc cggctgtttc tttcgtgttg caaacgcgc cagctcagca acctgtcccg	3780

ES 2 583 639 T3

	tgggtcccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
	agtaggcgcc ctcttttctg ctgcccctct ctccgtcccc ccactagtat ggccaccgca	3900
5	tccactttct cggcgttcaa tgcccgtgc ggcgacctgc gtcgctcggc gggctccggg	3960
	ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg cccagctgcc cgactggagc	4020
	cgcttgctga ccgcatcac caccgtgttc gtgaagtcca agcgccccga catgcacgac	4080
10	cgcaagtcca agcgccccga catgctggtg gacagcttcg gcctggagtc caccgtgcag	4140
	gacggcctgg tgttccgcca gtccttctcc atccgctcct acgagatcgg caccgaccgc	4200
15	accgccagca tcgagaccct gatgaaccac ctgcaggaga cctccctgaa ccaactgcaag	4260
	agcaccggca tcctgctgga cggcttcggc cgcaccctgg agatgtgcaa gcgcgacctg	4320
	atctgggtgg tgatcaagat gcagatcaag gtgaaccgct accccgcctg gggcgacacc	4380
20	gtggagatca acacccgctt cagccgcctg ggcaagatcg gcatgggccg cgactggctg	4440
	atctccgact gcaacaccgg cgagatcctg gtgcgcgcca ccagcgccta cgccatgatg	4500
25	aaccagaaga cccgccgcct gtccaagctg ccctacgagg tgcaccagga gatcgtgccc	4560
	ctgttcgtgg acagccccgt gatcgaggac tccgacctga aggtgcacaa gttcaaggtg	4620
	aagaccggcg acagcatcca gaagggcctg acccccggt ggaacgacct ggacgtgaac	4680
30	cagcacgtgt ccaacgtgaa gtacatcggc tggatcctgg agagcatgcc caccgaggtg	4740

ES 2 583 639 T3

	ctggagacc	aggagctgtg	ctccctggcc	ctggagtacc	gccgcgagtg	cggccgagac	4800
	tccgtgctgg	agagcgtgac	cgccatggac	cccagcaagg	tgggcgtgcg	ctcccagtac	4860
5	cagcacctgc	tgcgcctgga	ggacggcacc	gccatcgtga	acggcgccac	cgagtggcgc	4920
	cccaagaacg	ccggcgccaa	cggcgccatc	tccaccggca	agaccagcaa	cggcaactcc	4980
	gtgtccatgg	actacaagga	ccacgacggc	gactacaagg	accacgacat	cgactacaag	5040
10	gacgacgacg	acaagtgact	cgaggcagca	gcagctcgga	tagtatcgac	acactctgga	5100
	cgctggtcgt	gtgatggact	gttgccgcca	cacttgctgc	cttgacctgt	gaatatccct	5160
15	gccgctttta	tcaaacagcc	tcagtgtggt	tgatcttgtg	tgtacgcgct	tttgcgagtt	5220
	gctagctgct	tgtgctat	gcgaatacca	ccccagcat	ccccttcct	cgtttcatat	5280
	cgcttgcatc	ccaaccgcaa	cttatctacg	ctgtcctgct	atccctcagc	gctgctcctg	5340
20	ctcctgctca	ctgcccctcg	cacagccttg	gtttgggctc	cgctgtatt	ctcctggtag	5400
	tgcaacctgt	aaaccagcac	tgcaatgctg	atgcacggga	agtagtggga	tggaacaca	5460
25	aatggaaagc	tt					5472
	<210>	110					
	<211>	5451					
30	<212>	ADN					
	<213>	Secuencia Artificial					

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5 <400> 110
 ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 10 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 300
 15 cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 360
 gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 420
 20 gtgcacttca ccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag 480
 aaggacgcca agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg 540
 cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 600
 25 gccatcggcc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 660
 aacaacacct ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc 720
 30 tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 780
 tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 840

ES 2 583 639 T3

gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 900

caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 960

5 gcgttcgcca acgagggcct cctcggctac cagtacgagt gccccggcct gatcgaggtc 1020

cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1080

10 gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc 1140

gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1200

accttcttca acaccgacc gacctacggg agcgccttg gcatcgcgtg ggcctccaac 1260

15 tgggagtact ccgccttcgt gccaccaac ccctggcgct cctccatgtc cctcgtgcgc 1320

aagttctccc tcaacaccga gtaccaggcc aaccggaga cggagctgat caacctgaag 1380

20 gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 1440

acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 1500

ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 1560

25 ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 1620

gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 1680

30 aaccctact tcaccaaccg catgagcgtg aacaaccagc ctttcaagag cgagaacgac 1740

ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 1800

ES 2 583 639 T3

	gacggcgacg tcgtgtccac caacacctac ttcacgacca ccggaacgc cctgggctcc	1860
	gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgag	1920
5	gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggctg	1980
	tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt	2040
10	atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcg ttttgcgagt tgctagctgc	2100
	ttgtgctatt tgcaataacc acccccagca tccccttccc tcgtttcata tcgcttgc	2160
	cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc	2220
15	actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg	2280
	taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga	2340
20	tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc	2400
	agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg	2460
	aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga	2520
25	gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca	2580
	gccacagccg ctcccaccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct	2640
30	gccgcatgct tgtgctggtg aggctgggca gtgctgcat gctgattgag gcttggttca	2700
	tcgggtggaa gcttatgtgt gtgctgggct tgcatgccgg gcaatgcgca tgggtggcaag	2760

ES 2 583 639 T3

	agggcggcag cacttgctgg agctgccg cggtcctccag gtggttcaat cgcggcagcc	2820
	agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
5	ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940
	acacgggcat gccacagcac gggcaccggg gagtggaaatg gcaccaccaa gtgtgtgcga	3000
10	gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatcaa cgtaaccatg	3060
	agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca	3120
	tggtgtgcgg atgctgccgg ctgctgcctg ctgctgcacgc cgttgagttg gcagcaggct	3180
15	cagccatgca ctggatggca gctgggctgc cactgcaatg tggatgatag gatgcaagt	3240
	gagcgaatac caaacctct ggctgcttgc tgggttgcat ggcacgcac catcagcagg	3300
20	agcgcgatcg aaggactgg ccccatgcac gccatgcaa accggagcgc accgagtgtc	3360
	cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca ttagcgcctg	3420
	acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
25	gcggcggcag cagagcggcg gcagcagcgg cgggggcccac ctttcttgcg gggtcgcgcc	3540
	ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcac gcctggagaa	3600
30	gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
	cagcaaaggc gcctccccta ctccgatcg atgttccata gtgcattgga ctgcatttgg	3720

ES 2 583 639 T3

	gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccc	3780
	tgggtcccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
5	agtaggcgcc ctcttttctg ctcgccctct ctccgtcccc cactagtat gacgttcggg	3900
	gtcgccctcc cggccatggg ccgcggtgtc tcccttcccc ggcccagggt cgcggtgcgc	3960
10	gcccagtcgg cgagtcagg tttggagagc gggcgcgccc ccgactggtc catgctgttc	4020
	gccgtgatca ccaccatctt cagcgcgcc gagaagcagt ggaccaacct ggagtggaag	4080
	cccaagccca agctgcccc gctgctggac gaccacttcg gcctgcacgg cctggtgttc	4140
15	cgccgcacct tcgccatccg ctctacgag gtgggcccc accgcagcac ctccatcctg	4200
	gccgtgatga accacatgca ggaggccacc ctgaaccacg ccaagagcgt gggcatcctg	4260
20	ggcgacggct tcggcaccac cctggagatg tccaagcgcg acctgatgtg ggtggtgcgc	4320
	cgcaccacg tggccgtgga gcgctacccc acctggggcg acaccgtgga ggtggagtgc	4380
	tggatcggcg ccagcggcaa caacggcatg cgccgcgact tcctggtgcg cgactgcaag	4440
25	accggcgaga tcctgacctg ctgcacctcc ctgagcgtgc tgatgaacac ccgcacccg	4500
	cgctgagca ccatccccga cgaggtgcgc ggcgagatcg gccccgcctt catcgacaac	4560
30	gtggccgtga aggacgacga gatcaagaag ctgcagaagc tgaacgactc caccgccgac	4620
	tacatccagg gggcctgac cccccgctgg aacgacctgg acgtgaacca gcacgtgaac	4680

ES 2 583 639 T3

aacctgaagt acgtggcctg ggtgttcgag accgtgcccc acagcatcctt cgagtcccac 4740
 cacatcagct ccttcaccct ggagtaccgc cgcgagtgca cccgcgactc cgtgctgcgc 4800
 5 agcctgacca ccgtgagcgg cggcagctcc gaggccggcc tgggtgtgca ccacctgctg 4860
 cagctggagg gcggcagcga ggtgctgcgc gcccgcaccg agtggcgccc caagctgacc 4920
 10 gactccttcc gcggcatcag cgtgatcccc gccgagcccc gcgtgatgga ctacaaggac 4980
 cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtgactc 5040
 gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg 5100
 15 ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct 5160
 cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatattg 5220
 20 cgaataccac ccccagcatc cccttcctc gtttcatatc gcttgcatcc caaccgcaac 5280
 ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc 5340
 acagccttgg tttgggctcc gcctgtattc tcctggact gcaacctgta aaccagcact 5400
 25 gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atggaaagct t 5451

<210> 111

30 <211> 5454

<212> ADN

<213> Secuencia Artificial

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5 <400> 111
 ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 10 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 300
 15 cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 360
 gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 420
 20 gtgcacttca cccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag 480
 aaggacgcca agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg 540
 cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 600
 25 gccatcggcc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 660
 aacaacacct ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc 720
 30 tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 780
 tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 840

ES 2 583 639 T3

	gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc	900
	caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc	960
5	gcgttcgcca acgagggcct cctcggctac cagtacgagt gccccggcct gatcgaggtc	1020
	cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc	1080
	gccccggccg gcggtcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc	1140
10	gaggccttcg acaaccagtc ccgctgggtg gacttcggca aggactacta cgccctgcag	1200
	accttcttca acaccgacc gacctacggg agcgccttg gcatcgcgtg ggctccaac	1260
15	tgggagtact ccgccttcgt gccaccaac ccctggcgct cctccatgtc cctcgtgcgc	1320
	aagttctccc tcaacaccga gtaccaggcc aaccggaga cggagctgat caacctgaag	1380
	gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc	1440
20	acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag	1500
	ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac	1560
25	ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag	1620
	gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag	1680
	aaccctact tcaccaaccg catgagcgtg aacaaccagc cttcaagag cgagaacgac	1740
30	ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac	1800

ES 2 583 639 T3

	gacggcgacg tcgtgtccac caacacctac ttcacgacca ccggaacgc cctgggctcc	1860
	gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgag	1920
5	gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggctg	1980
	tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt	2040
	atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc	2100
10	ttgtgctatt tgcaataacc acccccagca tccccttccc tcgtttcata tcgcttgc	2160
	cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc	2220
15	actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg	2280
	taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga	2340
	tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc	2400
20	agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg	2460
	aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga	2520
25	gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca	2580
	gccacagccg ctcccaccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct	2640
	gccgcatgct tgtgctggg aggctgggca gtgctgcat gctgattgag gcttggttca	2700
30	tcgggtgga gcttatgtgt gtgctgggct tgcacgccg gcaatgcgca tgggtggcaag	2760

ES 2 583 639 T3

	agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc	2820
	agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
5	ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940
	acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga	3000
	gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg	3060
10	agctgatcaa cactgcaatc atcggggcgg cgatgatgcaa gcatgcctgg cgaagacaca	3120
	tggtgtgcgg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct	3180
15	cagccatgca ctggatggca gctgggctgc cactgcaatg tggtggatag gatgcaagtg	3240
	gagcgaatac caaacctct ggctgcttgc tgggttgcat ggcacgcac catcagcagg	3300
	agcgcacgag aagggactgg ccccatgcac gccatgcaa accggagcgc accgagtgtc	3360
20	cacactgtca ccaggcccgc aagctttgca gaaccatgct catggacgca tgtagcgtg	3420
	acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
25	gcggcggcag cagagcggcg gcagcagcgg cgggggcccac ccttcttgcg gggtcgcgcc	3540
	ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcat gcctggagaa	3600
	gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
30	cagcaaaggc gcctccccta ctccgatcg atgtccata gtgcattgga ctgcatttgg	3720

ES 2 583 639 T3

	gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccg	3780
	tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
5	agtaggcgcc ctcctttctg ctcgccctct ctccgtcccg ccactagtat ggccaccgca	3900
	tccactttct cggcgttcaa tgcccgtgc ggcgacctgc gtcgctcggc gggctccggg	3960
	ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg cccccgactg gtccatgctg	4020
10	ttcgccgtga tcaccacat cttctccgcc gccgagaagc agtggaccaa cctggagtgg	4080
	aagcccaagc ccaaccccc ccagctgctg gacgaccact tcggccccca cggcctggtg	4140
15	ttccgccgca ctttcgcat ccgcagctac gaggtgggcc ccgaccgctc caccagcatc	4200
	gtggccgtga tgaaccacct gcaggaggcc gccctgaacc acgccaagtc cgtgggcatc	4260
	ctgggcgacg gcttcggcac caccctggag atgtccaagc gcgacctgat ctgggtggtg	4320
20	aagcgcacc acgtggccgt ggagcgctac cccgcctggg gcgacaccgt ggaggtggag	4380
	tgctgggtgg ggcctccgg caacaacggc cgccgccacg acttcctggt gcgcgactgc	4440
25	aagaccggcg agatcctgac ccgctgcacc tccctgagcg tgatgatgaa caccgcacc	4500
	cgccgcctga gcaagatccc cgaggagggtg cgcggcgaga tcggccccgc cttcatcgac	4560
	aacgtggccg tgaaggacga ggagatcaag aagccccaga agctgaacga ctccaccgcc	4620
30	gactacatcc agggcggcct gacccccgc tggaacgacc tggacatcaa ccagcacgtg	4680

ES 2 583 639 T3

	aacaacatca agtacgtgga ctggatcctg gagaccgtgc ccgacagcat cttcgagagc	4740
	caccacatct cctccttcac catcgagtac cgccgcgagt gcacatgga cagcgtgctg	4800
5	cagtccctga ccaccgtgag cggcggctcc tccgaggccg gcctggtgtg cgagcacctg	4860
	ctgcagctgg agggcggcag cgaggtgctg cgcgccaaga ccgagtggcg cccaagctg	4920
	accgactcct tccgcggcat cagcgtgatc cccgccgagt ccagcgtgat ggactacaag	4980
10	gaccacgacg gcgactacaa ggaccacgac atcgactaca aggacgacga cgacaagtga	5040
	ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga	5100
15	ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag	5160
	cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat	5220
	ttgcgaatac cacccccagc atccccctcc ctcgtttcat atcgcttgca tccaaccgc	5280
20	aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccctt	5340
	cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc	5400
25	actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggaaa gctt	5454
	<210> 112	
	<211> 2933	
30	<212> ADN	
	<213> Secuencia Artificial	

ES 2 583 639 T3

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

5 <400> 112
agagagcgga ggtgggggtg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg 60
cgactcgcac acggtccagt tcccccccc tccgccaaa cgcaagcctc ccatcttgat 120
10 gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata 180
ataacaatgc cctgtgtcaa gtgcattcct aaaaaattc tgtccaacc aacaatccca 240
cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg 300
15 cgcaagcgac ccccatgcmc gaccaggctc gaaagtgatt tatgacttga gacgagcgag 360
tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt 420
20 gcacggcaac gcacggcttc gccaaaccga ccagccagga cctcgactac tctaccgca 480
attcgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact 540
tgatgccaat tttgacattc gggttgctcg ttggctgcmc ccacatcggc cgtgagtgca 600
25 gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta 660
catcaacacg acgtgttgcm taatccccgc cggctgcmc tcgtgccaac ccattcgcga 720
30 tggatggctg gaaaatggtg tgccaactgc cctgaggag gctctcgcga aacgggcacg 780
tccctgaaac cgaaactgtg gccttgtcgt cggccacgca agcacgtgga ccctaaacac 840

ES 2 583 639 T3

caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc 900

atcgcacact gccattataa tgcatttagc tcggcgacaa gtttagaaaa ggcaggctgc 960

5 attgttccat ttcgccgtgg cggcgtgggt gccatttta cgaggtttgg gctcccgggc 1020

agcgaccgag ccaggctgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080

10 gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140

acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200

ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260

15 cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttctactat 1320

atthttccagt cgtcagcgaa atcaagtact agtatggcca ccgcatccac tttctcggcg 1380

20 ttcaatgccc gctgcggcga cctgcgtcgc tcggcgggct ccgggccccg gcgcccagcg 1440

aggccccctcc ccgtgcgcgg gcgcgcccc gactggtcca tgctgttcgc cgtgatcacc 1500

accatcttct ccgccgccga gaagcagtgg accaacctgg agtggaagcc caagcccaac 1560

25 cccccccagc tgctggacga ccacttcggc ccccacggcc tgggtttccg ccgcaccttc 1620

gccatccgca gctacgaggt gggccccgac cgctccacca gcatcgtggc cgtgatgaac 1680

30 cacctgcagg aggccgccct gaaccacgcc aagtccgtgg gcatcctggg cgacggcttc 1740

ggcaccacc tggagatgtc caagcgcgac ctgatctggg tggatgaagcg caccacgtg 1800

ES 2 583 639 T3

gccgtggagc gctacccccg ctggggcgac accgtggagg tggagtgctg ggtgggcgcc 1860
 tccggcaaca acggccgccg ccacgacttc ctggtgcgcg actgcaagac cggcgagatc 1920
 5 ctgacccgct gcacctccct gagcgtgatg atgaacaccc gcacccgccg cctgagcaag 1980
 atccccgagg aggtgcgcgg cgagatcggc cccgccttca tcgacaacgt ggccgtgaag 2040
 10 gacgaggaga tcaagaagcc ccagaagctg aacgactcca ccgccgacta catccagggc 2100
 ggctgacct cccgctggaa cgacctggac atcaaccagc acgtgaacaa catcaagtac 2160
 gtggactgga tcctggagac cgtgccccgac agcatcttcg agagccacca catctcctcc 2220
 15 ttcaccatcg agtaccgccg cgagtgcacc atggacagcg tgctgcagtc cctgaccacc 2280
 gtgagcggcg gctcctccga ggccggcctg gtgtgcgagc acctgctgca gctggagggc 2340
 20 ggcagcgagg tgctgcgcgc caagaccgag tggcgccccca agctgaccga ctccttccgc 2400
 ggcacagcg tgatccccgc cgagtccagc gtgatggact acaaggacca cgacggcgac 2460
 tacaaggacc acgacatcga ctacaaggac gacgacgaca agtgactcga gttaattaac 2520
 25 tcgaggcagc agcagctcgg atagtatcga cacactctgg acgctggctg tgtgatggac 2580
 tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc 2640
 30 ctcagtgtgt ttgatcttgt gtgtacgcg ttttgcgagt tgctagctgc ttgtgctatt 2700
 tgccaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat cccaaccgca 2760

ES 2 583 639 T3

acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc actgcccctc 2820

gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca 2880

5 ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaag ctt 2933

<210> 113

<211> 4817

10 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido

15 sintético

<400> 113

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

20 gtgcttgggc gcctgccgcc tgctgcccgc atgcttgtgc tggtaggct gggcagtgc 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

25 tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

30 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggtgtttc agctacaacg 480

ES 2 583 639 T3

	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
5	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
	caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
10	tgcatggcat cgacccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gtcctctcg ggtgtgggaa	900
15	acgcaatgca gcacaggcag cagagggcggc ggcagcagag cggcggcagc agcggcggg	960
	gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagccccga	1080
20	atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
25	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
	gatcagctga ttgccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320
	tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
30	ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440

ES 2 583 639 T3

cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

5 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

10 ccaccaggcc aagcaccgca tcgagcgcgc ccgacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

15 gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcacg gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040

20 cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcac 2100

cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

25 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340

30 ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400

ES 2 583 639 T3

	ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggtt gggctccgcc	2460
	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
5	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
	gaaggctctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
10	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
15	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
20	cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
	cagtcacaac ccgcaaacgg cgcgccatgg tggccgccgc gcctccagc gccttcttcc	3120
25	ccgtgcccgc ccccgggcgcc tcccccaagc ccggcaagtt cggcaactgg ccctccagcc	3180
	tgagcccctc cttcaagccc aagtccatcc ccaacggcgg cttccaggtg aaggccaacg	3240
	acagcgccca cccaaggcc aacggctccg ccgtgagcct gaagagcggc agcctgaaca	3300
30	cccaggagga cacctcctcc agcccccccc cccgcacctt cctgcaccag ctgcccgact	3360

ES 2 583 639 T3

	ggagccgcct gctgaccgcc atcaccaccg tgttcgtgaa gtccaagcgc cccgacatgc	3420
	acgaccgcaa gtccaagcgc cccgacatgc tgggtggacag cttcggcctg gagtccaccg	3480
5	tgcaggacgg cctggtgttc cgccagtcct tctccatccg ctcttacgag atcggcaccg	3540
	accgcaccgc cagcatcgag accctgatga accacctgca ggagacctcc ctgaaccact	3600
	gcaagagcac cggcatcctg ctggacggct tcggccgcac cctggagatg tgcaagcgcg	3660
10	acctgatctg ggtggtgatc aagatgcaga tcaaggtgaa ccgctacccc gcctggggcg	3720
	acaccgtgga gatcaacacc cgcttcagcc gcctgggcaa gatcggcatg ggccgcgact	3780
15	ggctgatctc cgactgcaac accggcgaga tcctggtgcg cgccaccagc gcctacgcc	3840
	tgatgaacca gaagaccgc cgctgtcca agctgcccta cgaggtgcac caggagatcg	3900
	tgcccctgtt cgtggacagc cccgtgatcg aggactccga cctgaaggtg cacaagttca	3960
20	aggtgaagac cggcgacagc atccagaagg gcctgacccc cggctggaac gacctggacg	4020
	tgaaccagca cgtgtccaac gtgaagtaca tcggctggat cctggagagc atgcccaccg	4080
25	aggtgctgga gaccaggag ctgtgctccc tggccctgga gtaccgccgc gagtgcgcc	4140
	gcgactccgt gctggagagc gtgaccgcca tggaccccag caaggtgggc gtgcgctccc	4200
	agtaccagca cctgctgctc ctggaggacg gcaccgcat cgtgaacggc gccaccgagt	4260
30	ggcgcccaaa gaacgccggc gccaacggcg ccattctccac cggcaagacc agcaacggca	4320

ES 2 583 639 T3

	actccgtgtc catggactac aaggaccacg acggcgacta caaggaccac gacatcgact	4380
	acaaggacga cgacgacaag tgactcgagg cagcagcagc tcggatagta tcgacacact	4440
5	ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	4500
	tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	4560
	gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt	4620
10	catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc	4680
	tctgtctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	4740
15	gg tactgcaa cctgtaaacc agcactgcaa tgctgatgca cg ggaagtag tgggatggga	4800
	acacaaatgg aaagctt	4817
20	<210> 114	
	<211> 4665	
	<212> ADN	
	<213> Secuencia Artificial	
25	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 114	
30	ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgct	120

ES 2 583 639 T3

	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat	180
	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
5	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
10	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctgta	540
15	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
20	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcatggcat cgcacatca gcaggagcgc atgcgaaggg actggccccca tgcacgcat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
25	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gtcctctctg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
30	gccacccttc ttgccccgtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagccccga	1080

ES 2 583 639 T3

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catc gatgtt 1140
 ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
 5 cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgctcgc cctctctccg 1320
 10 tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
 15 cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 20 gggcgaggctg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggctgc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgacccgc atggaggccg gcctggtgga 1800
 25 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccccct 1860
 gaaggccccg atgccccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 30 caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcgcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040

ES 2 583 639 T3

cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gcccccgact cccagcgcac 2100

cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

5 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

10 cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggtt gggctccgcc 2460

15 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta 2520

gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct 2580

20 gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

25 gaattcctt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

30 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

ES 2 583 639 T3

cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacac tagtatgacg ttcggggtcg ccctcccggc catgggccgc 3120
 5 ggtgtctccc ttccccggcc cagggctcgc gtgcgcgcc agtcggcgag tcaggttttg 3180
 gagagcgggc gcgccagct gcccgactgg agccgcctgc tgaccgcat caccaccgtg 3240
 10 ttcgtgaagt ccaagcggcc cgacatgcac gaccgcaagt ccaagcggcc cgacatgctg 3300
 gtggacagct tcggcctgga gtccaccgtg caggacggcc tgggtttccg ccagtccttc 3360
 tccatccgct cctacgagat cggcaccgac cgcaccgcca gcatcgagac cctgatgaac 3420
 15 cacctgcagg agacctcct gaaccactgc aagagcaccg gcatcctgct ggacggcttc 3480
 ggccgcaccc tggagatgtg caagcgcgac ctgatctggg tggatgataa gatgcagatc 3540
 20 aaggatgaacc gctaccccg cctggggcgac accgtggaga tcaacacccg cttcagccgc 3600
 ctgggcaaga tcggcatggg ccgagactgg ctgatctccg actgcaacac cggcgagatc 3660
 ctggtgcgag ccaccagcgc ctacgcatg atgaaccaga agaccgccc cctgtccaag 3720
 25 ctgccctacg aggtgcacca ggagatcgtg cccctgttcg tggacagccc cgtgatcgag 3780
 gactccgacc tgaaggtgca caagttcaag gtgaagaccg ggcagagcat ccagaagggc 3840
 30 ctgacccccg gctggaacga cctggacgtg aaccagcacg tgtccaacgt gaagtacatc 3900
 ggctggatcc tggagagcat gccaccgag gtgctggaga cccaggagct gtgctccctg 3960

ES 2 583 639 T3

gccctggagt accgccgca gtgcggccgc gactccgtgc tggagagcgt gaccgccatg 4020
 gaccccagca aggtgggcgt gcgctcccag taccagcacc tgctgvcct ggaggacggc 4080
 5 accgccatcg tgaacggcgc caccgagtgg cgccccaaga acgccggcgc caacggcgcc 4140
 atctccaccg gcaagaccag caacggcaac tccgtgtcca tggactacaa ggaccacgac 4200
 10 ggcgactaca aggaccacga catcgactac aaggacgacg acgacaagtg actcgaggca 4260
 gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 4320
 ccacacttgc tgccttgacc tgtgaatata cctgccgctt ttatcaaaca gcctcagtgt 4380
 15 gtttgatctt gtgtgtacgc gcttttgca gttgctagct gcttvtgcta tttgcaata 4440
 ccacccccag catccccttc cctcgtttca tctgcttgc atcccaaccg caacttatct 4500
 20 acgctgtcct gctatccctc agcgtgctc ctgctcctgc tcaactgcccc tcgcacagcc 4560
 ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 4620
 ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agctt 4665
 25
 <210> 115
 <211> 4668
 <212> ADN
 30 <213> Secuencia Artificial
 <220>

ES 2 583 639 T3

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 115

5	ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgcctgccgc atgcttgctg tggtgaggct gggcagtgct	120
	gcatgctga ttgaggcttg gttcatcggg tgaagctta tgtgtgtgct gggcttgc	180
10	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
15	cagcagtgct agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
	cagcaagaga aggggtcaag tgcaaacacg ggcagccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
20	gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgta	540
	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
25	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
	caatgtggtg gataggatgc aagtggagcg aatacacaac cctctggctg cttgctgggt	720
	tgcatggcat cgacccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
30	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840

ES 2 583 639 T3

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

5 gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

10 ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

15 gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgctcgc cctctctccg 1320

tccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

20 cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

25 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

30 ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800

ES 2 583 639 T3

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

5 caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcac 2100

10 cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

15 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400

20 ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggtt gggctccgcc 2460

tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

25 gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

30 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

ES 2 583 639 T3

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

5 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 3060

10 cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

cctccgttca cgatcgggac gctgcgcaag gccatccccg cgactgttt cgagcgctcg 3180

15 gcgcttcgtg ggcgcgccca gctgccccgac tggagccgcc tgctgaccgc catcaccacc 3240

gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300

ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtggt cgcagatcc 3360

20 ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420

aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480

25 ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540

atcaaggatga accgctaccc cgctggggc gacaccgtgg agatcaacac ccgcttcagc 3600

cgctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660

30 atcctggtgc gcgccaccag cgctacgcc atgatgaacc agaagaccg ccgcctgtcc 3720

ES 2 583 639 T3

	aagctgccct acgaggtgca ccaggagatc gtgccctgt tcgtggacag ccccgatc	3780
	gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag	3840
5	ggcctgacct ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac	3900
	atcggctgga tcctggagag catgcccacc gaggtgctgg agaccagga gctgtgctcc	3960
	ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc	4020
10	atggaccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac	4080
	ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccg cgccaacggc	4140
15	gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac	4200
	gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag	4260
	gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg	4320
20	ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag	4380
	tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgct ctatttgca	4440
25	ataccacccc cagcatcccc ttccctcggt tcatatcgct tgcacccaa ccgcaactta	4500
	tctacgctgt cctgctatcc ctcagcgtg ctctgctcc tgctcactgc ccctcgaca	4560
	gccttggttt gggctccgcc tgtattctcc tggactgca acctgtaaac cagcactgca	4620
30	atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt	4668

ES 2 583 639 T3

<210> 116

<211> 4668

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

10

<400> 116

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgcccgc tgccctgccc atgcttgtgc tggtagggct gggcagtgct 120

15

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

gccgggcaat gcgcatggcg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

20

tccaggtggt tcaatcgccg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

cagcagtgct agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcagccac agcacgggca ccggggagtg 420

25

gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgta 540

30

tgcaagcatg cctggcgaag acacatggcg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

ES 2 583 639 T3

	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
5	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
10	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960
	gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcatgcctg gagaagcag gctggggcct ttgggctggt gcagccccgca	1080
15	atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt	1140
	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
20	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
	gatcagctga ttgcccggct cgcaagtag gcgcccctcct ttctgctcgc cctctctccg	1320
	tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgcccggct cccccgccgc	1380
25	ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
	cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc	1500
30	cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
	cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620

ES 2 583 639 T3

5
 10
 15
 20
 25
 30

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttca 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
 gaaggcccg atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040
 cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcatt 2100
 cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
 cgcgcttttg cgagttgcta gctgcttgg ctatttgca ataccacccc cagcatcccc 2340
 ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggttt gggctccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct 2580

ES 2 583 639 T3

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

5 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

10 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

15 cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120

20 cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180

cccgtgcgcy ggcgcgcca gctgcccgac tggagccgcc tgctgaccgc catcaccacc 3240

gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300

25 ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtggt ccgccagtcc 3360

ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420

30 aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480

ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540

ES 2 583 639 T3

	atcaaggtga accgctaccc cgcttggggc gacaccgtgg agatcaacac ccgcttcagc	3600
	cgcttgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag	3660
5	atcctgggtgc gcgccaccag cgcttacgcc atgatgaacc agaagaccgg ccgctgtcc	3720
	aagctgccct acgaggtgca ccaggagatc gtgccctgt tcgtggacag ccccgtgatc	3780
10	gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag	3840
	ggcctgacct ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac	3900
	atcggctgga tcctggagag catgcccacc gaggtgctgg agaccagga gctgtgctcc	3960
15	ctggccctgg agtaccgccc cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc	4020
	atggaccca gcaaggtggg cgtgctgctcc cagtaccagc acctgctgcg cctggaggac	4080
20	ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc	4140
	gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac	4200
	gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgactcgag	4260
25	gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg	4320
	ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag	4380
30	tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttggt ctatttgca	4440
	ataccacccc cagcatcccc ttccctcggt tcatatcgct tgcatcccaa ccgcaactta	4500

ES 2 583 639 T3

	tctacgctgt cctgctatcc ctcagcgctg ctctgctcc tgctcactgc ccctcgaca	4560
	gccttggttt gggctccgcc tgtattctcc tggactgca acctgtaaac cagcactgca	4620
5	atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt	4668
	<210> 117	
10	<211> 4656	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
15	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 117	
	ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
20	gtgcttgggc gcctgccgcc tgccctgccgc atgcttgtgc tggtaggct gggcagtgc	120
	gcatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc	180
25	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
30	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420

ES 2 583 639 T3

	gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
5	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
10	tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
15	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
	gccacccttc ttgcggggtc gcgccccagc cagcggatg gcgctgatcc caaacgagtt	1020
20	cacattcatt tgcctgcctg gagaagcag gctggggcct ttgggctggt gcagccccga	1080
	atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgaatgt	1140
25	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
	gatcagctga ttgcccggct cgcaagtag gcgcctcct ttctgctcgc cctctctccg	1320
30	tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380

ES 2 583 639 T3

	ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
	cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc	1500
5	cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
	cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
	gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc	1680
10	catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
	ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga	1800
15	ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
	gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
	caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt	1980
20	ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
	cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgc	2100
25	cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
	atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
	acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
30	cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc	2340

ES 2 583 639 T3

	tccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
	ctcagcgctg ctctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc	2460
5	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
	gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
10	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
15	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
20	aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta	3000
	cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
25	cagtcacaac ccgcaaacac tagtatggct tccgggcat tcaccatgtc ggcgtgcccc	3120
	gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg	3180
	agggggcgcg cccccgactg gagccgcctg ctgaccgcca tcaccaccgt gttcgtgaag	3240
30	tccaagcgcc ccgacatgca cgaccgcaag tccaagcgcc ccgacatgct ggtggacagc	3300

ES 2 583 639 T3

	ttcggcctgg agtccaccgt gcaggacggc ctggtgttcc gccagtcctt ctccatccgc	3360
	tcctacgaga tcggcaccga ccgcaccgcc agcatcgaga ccctgatgaa ccacctgcag	3420
5	gagacctccc tgaacctg caagagcacc ggcacccctgc tggacggcctt cggccgcacc	3480
	ctggagatgt gcaagcgca cctgatctgg gtggtgatca agatgcagat caaggtgaac	3540
	cgctaccccg cctggggcga caccgtggag atcaacacc gcttcagccg cctgggcaag	3600
10	atcggcatgg gccgcgactg gctgatctcc gactgcaaca ccggcgagat cctggtgcgc	3660
	gccaccagcg cctacgccat gatgaaccag aagaccgcc gcctgtccaa gctgcctac	3720
15	gaggtgcacc aggagatcgt gcccctgttc gtggacagcc ccgtgatcga ggactccgac	3780
	ctgaaggtgc acaagttaa ggtgaagacc ggcgacagca tccagaagg cctgaccccc	3840
	ggctggaacg acctggacgt gaaccagcac gtgtccaacg tgaagtacat cggctggatc	3900
20	ctggagagca tgcccaccga ggtgctggag acccaggagc tgtgctccct ggccctggag	3960
	taccgccgag agtgcgccg cgactccgtg ctggagagcg tgaccgcat ggacccagc	4020
25	aaggtgggag tgcgctcca gtaccagcac ctgctgcgcc tggaggacgg caccgcatc	4080
	gtgaacggcg ccaccgagtg gcgccccaa aacgccggcg ccaacggcg catctccacc	4140
	ggcaagacca gcaacggcaa ctccgtgtcc atggactaca aggaccagc cggcgactac	4200
30	aaggaccagc acatcgacta caaggacgac gacgacaagt gactcgaggc agcagcagct	4260

ES 2 583 639 T3

	cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg	4320
	ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct	4380
5	tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accaccccca	4440
	gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc	4500
	tgctatccct cagcgtgct cctgctcctg ctactgccc ctgcacagc cttggtttgg	4560
10	gctccgctg tattctcctg gtactgcaac ctgtaaacca gcaactgcaat gctgatgcac	4620
	gggaagtagt gggatgggaa cacaaatgga aagctt	4656
15	<210> 118	
	<211> 4721	
	<212> ADN	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
25	<400> 118	
	ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgcttgcgc atgcttgtgc tggtaggct gggcagtgct	120
30	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240

ES 2 583 639 T3

	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
5	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg	480
10	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctgta	540
	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgctg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
15	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
20	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
25	gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggt gcagccccga	1080
30	atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt	1140
	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200

ES 2 583 639 T3

cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgcccggct cggaagtag gcgccctcct ttctgctcgc cctctctccg 1320
 5 tccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 10 cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
 cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 15 gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 20 ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgcccgcct 1860
 gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 25 caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggctg 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggctg 2040
 30 cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcac 2100
 cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

ES 2 583 639 T3

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
 5
 cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340
 ttcctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc 2400
 10
 ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
 15
 gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 20
 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
 25
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
 30
 cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacgg cgcgccatgg ccaccaccag cctggcctcc gccttctgct 3120

ES 2 583 639 T3

ccatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc agctccgacc 3180
 tgcagctgcg cgccggcaac gccccacct ccctgaagat gatcaacggc accaagttca 3240
 5 gctacaccga gagcctgaag cgctgccccg actggtccat gctgttcgcc gtgatcacca 3300
 ccatcttcag cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaagc 3360
 10 tgccccagct gctggacgac cacttcggcc tgcacggcct ggtgttcgcg cgcaccttcg 3420
 ccatccgctc ctacgagtg ggccccgacc gcagcacctc catcctggcc gtgatgaacc 3480
 acatgcagga ggccaccctg aaccacgcca agagcgtggg catcctgggc gacggcttcg 3540
 15 gcaccaccct ggagatgtcc aagcgcgacc tgatgtgggt ggtgcgccgc acccacgtgg 3600
 ccgtggagcg ctaccccacc tggggcgaca ccgtggaggt ggagtgctgg atcggcgcca 3660
 20 gcggcaacaa cggcatgctc cgcgacttcc tgggtgcgca ctgcaagacc ggcgagatcc 3720
 tgacccgctg cacctccctg agcgtgctga tgaacacccg cacccgccgc ctgagcacca 3780
 tccccgacga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg 3840
 25 acgacgagat caagaagctg cagaagctga acgactccac cgccgactac atccagggcg 3900
 gcctgacccc ccgctggaac gacctggacg tgaaccagca cgtgaacaac ctgaagtacg 3960
 30 tggcctgggt gttcgagacc gtgcccgaca gcatcttcga gtcccaccac atcagctcct 4020
 tcaccctgga gtaccgccgc gagtgcaccc gcgactccgt gctgcgagc ctgaccaccg 4080

ES 2 583 639 T3

	tgagcggcgg cagctccgag gccggcctgg tgtgcgacca cctgctgcag ctggagggcg	4140
	gcagcgaggt gctgcgcgcc cgcaccgagt ggcgccccaa gctgaccgac tccttccgcg	4200
5	gcatcagcgt gatccccgcc gagccccgcg tgatggacta caaggaccac gacggcgact	4260
	acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag	4320
10	cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	4380
	acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	4440
	gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	4500
15	ccccagcatc cccttccctc gtttcatatc gcttgcaccc caaccgcaac ttatctacgc	4560
	tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg	4620
20	tttgggctcc gcctgtattc tcctggact gcaacctgta aaccagcact gcaatgctga	4680
	tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	4721
25	<210> 119	
	<211> 4650	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	

ES 2 583 639 T3

<400> 119

	ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
5	gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgc	120
	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc	180
	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
10	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
15	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
20	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
25	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
	gccaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
30	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900

ES 2 583 639 T3

acgcaatgca gcacaggcag cagagggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

5 cacattcatt tgcattgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

10 cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgcctcct ttctgctcgc cctctctccg 1320

15 tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

20 cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

25 gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800

30 ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgccccct 1860

ES 2 583 639 T3

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

5 ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcac 2100

cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

10 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

15 cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340

ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggttt gggctccgcc 2460

20 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

25 gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttgggt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

30 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

ES 2 583 639 T3

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

5 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttcggggctg ccctcccggc catgggccgc 3120

10 ggtgtctccc ttccccggcc cagggctcgc gtgcgcgcc agtcggcgag tcaggttttg 3180

gagagcgggc gcgccccga ctggtccatg ctgttcgcc tgatcaccac catcttcagc 3240

15 gccgccgaga agcagtggac caacctggag tggaagccca agcccaagct gccccagctg 3300

ctggacgacc acttcggcct gcacggcctg gtgttcgcc gcaccttcgc catccgctcc 3360

tacgaggtgg gccccgaccg cagcacctcc atcctggccg tgatgaacca catgcaggag 3420

20 gccaccctga accacgcaa gagcgtggc atcctggcg acggcttcgg caccaccctg 3480

gagatgtcca agcgcgacct gatgtgggtg gtgcgccga cccacgtggc cgtggagcgc 3540

25 taccacact gggcgacac cgtggaggtg gagtgctgga tcggcgccag cggcaacaac 3600

ggcatgcgcc gcgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gaccgctgc 3660

acctccctga gcgtgctgat gaacaccgc acccgccgc tgagcacat ccccgacgag 3720

30 gtgcgcggcg agatcggccc cgccttcac gacaacgtg ccgtaagga cgacgagatc 3780

ES 2 583 639 T3

aagaagctgc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc 3840

cgctggaacg acctggacgt gaaccagcac gtgaacaacc tgaagtacgt ggcttgggtg 3900

5 ttcgagaccg tgcccgcacag catcttcgag tcccaccaca tcagctcctt caccctggag 3960

taccgccgag agtgcacccg cgactccgtg ctgctgcagc tgaccaccgt gagcggcggc 4020

agctccgagg ccggcctggt gtgctgaccac ctgctgcagc tggagggcgg cagcggagtg 4080

10 ctgctgcagc gcaccgagtg gcgcccccaag ctgaccgact ccttccgagg catcagcgtg 4140

atccccgccg agccccgcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200

15 gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260

gtatcgacac actctggacg ctggctcgtg gatggactgt tgccgccaca cttgctgcct 4320

tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg 4380

20 tacgcgcttt tgcgagttgc tagctgcttg tgctatctgc gaataccacc cccagcatcc 4440

ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat 4500

25 ccctcagcgc tgctcctgct cctgctcact gccctcgcga cagccttggg ttgggctccg 4560

cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620

tagtgggatg ggaacacaaa tggaaagctt 4650

30

<210> 120

ES 2 583 639 T3

<211> 4653

<212> ADN

<213> Secuencia Artificial

5 <220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 120

10 ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgc 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180
15 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
20 cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
25 gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgta 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
30 cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt 720

ES 2 583 639 T3

	tgcatggcat cgcacatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
5	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
10	gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggt gcagcccga	1080
	atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt	1140
15	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
20	gatcagctga ttgccggct cgccaagtag gcgccctcct ttctgctcgc cctctctccg	1320
	tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
	ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
25	cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga cgcacctgtc	1500
	cggcgcctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
30	cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
	gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc	1680

ES 2 583 639 T3

	catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga	1740
	ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga	1800
5	ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
	gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
10	caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt	1980
	ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
	cgagtgggcc gaccgcttcc tggctgtgta cggcatcgcc gccccgact cccagcgcac	2100
15	cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
	atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
20	acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
	cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc	2340
	ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc	2400
25	ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggtt gggctccgcc	2460
	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
30	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
	gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640

ES 2 583 639 T3

	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatc	2760
5	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagtcct tcggggctgc	2880
10	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
	cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt	3060
15	cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
	cctccgttca cgatcgggac gctgcgcaag gccatccccg cgactgttt cgagcgctcg	3180
20	gcgcttcgtg ggcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc	3240
	agcgcgccg agaagcagtg gaccaacctg gagtgaagc ccaagcccaa gctgccccag	3300
	ctgctggacg accacttcg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc	3360
25	tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag	3420
	gaggccacc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc	3480
30	ctggagatgt ccaagcgca cctgatgtgg gtggtgcgcc gcaccacgt ggccgtggag	3540
	cgctacccca cctggggcga caccgtggag gtggagtgc ggatcggcgc cagcggcaac	3600

ES 2 583 639 T3

	aacggcatgc gccgcgactt cctggtgctg gactgcaaga ccggcgagat cctgacccgc	3660
	tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac	3720
5	gaggtgctgc gcgagatcgg ccccgcttc atcgacaacg tggccgtgaa ggacgacgag	3780
	atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc	3840
10	ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg	3900
	gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg	3960
	gagtaccgcc gcgagtgcac ccgcgactcc gtgctgctca gcctgaccac cgtgagcggc	4020
15	ggcagctccg aggccggcct ggtgtgctgc cacctgctgc agctggaggg cggcagcgg	4080
	gtgctgctgc cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc	4140
20	gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
	cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
	atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg	4320
25	ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt	4380
	gtgtacgctc ttttgagagt tgctagctgc ttgtgctatt tgccaatacc acccccagca	4440
30	tccccttccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc	4500
	tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct	4560

ES 2 583 639 T3

	ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg	4620
	aagtagtggg atgggaacac aaatggaaag ctt	4653
5		
	<210> 121	
	<211> 4653	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
15		
	<400> 121	
	ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgc	120
20		
	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc	180
	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
25		
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
30		
	gaatggcacc accaagtgtg tgcgagccag catcggccg tggctgtttc agctacaacg	480

ES 2 583 639 T3

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctgta 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

5 cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt 720

tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

10 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

15 acgcaatgca gcacaggcag cagagggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcattcctg gagaagcag gctggggcct ttgggctggt gcagcccgca 1080

20 atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgaatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

25 cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgccggct cgcaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcagca ggacggcctc cacgccggct cccccgccgc 1380

30 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

ES 2 583 639 T3

cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

5 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

10 ccaccaggcc aagcaccgca tcgagcgcgc ccgacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

15 gaaggccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcacg gactgcgcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040

20 cgagtgggcc gaccgcttcc tggctgtgta cggcatcgcc gccccgact cccagcgcac 2100

cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

25 atcgacacac tctggacgct ggtcgtgtga tggactggtg ccgccacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340

30 ttcctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400

ES 2 583 639 T3

	ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggtt gggctccgcc	2460
	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta	2520
5	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
	gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
10	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
15	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
20	cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
	cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc	3120
25	cgctgcggcg acctgcgtcg ctggcgggc tccgggcccc ggcgcccagc gaggcccctc	3180
	cccgtgcgcy ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc	3240
	agcgcgcccg agaagcagtg gaccaacctg gagtgaagc ccaagcccaa gctgccccag	3300
30	ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc	3360

ES 2 583 639 T3

	tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag	3420
	gaggccacc tgaaccacgc caagagcgtg ggcacccctg gcgacggctt cggcaccacc	3480
5	ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag	3540
	cgctaccca cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac	3600
	aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc	3660
10	tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac	3720
	gaggtgcgcg gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgacgag	3780
15	atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc	3840
	ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg	3900
	gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg	3960
20	gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc	4020
	ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag	4080
25	gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc	4140
	gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
	cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
30	atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg	4320

ES 2 583 639 T3

ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgccaatacc acccccagca 4440

5 tccccttccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

10 aagtagtggg atgggaacac aatggaaag ctt 4653

<210> 122

15 <211> 4647

<212> ADN

<213> Secuencia Artificial

<220>

20 <223> Descripción de Secuencia Artificial: Polinucleótido
sintético

<400> 122

25 ggtaccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgc 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

30 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300

ES 2 583 639 T3

	cagcagtgtc agcaaaghta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
5	gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgta	540
10	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcy	600
	cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
	caatgtggtg gataggatgc aagtggagcy aatacctaac cctctggctg cttgctgggt	720
15	tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
20	atgctcatgg acgcatgtag cyctgacgty ccttgacggc gctcctctcy ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
	gccacccttc ttgcggggty gcgccccagc cagcggatg gcgctgatcc caaacgagtt	1020
25	cacattcatt tgcatgcctg gagaagcag gctggggcct ttgggctggt gcagcccgca	1080
	atggaatgcy ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
30	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcy tgttgcaaaa	1200
	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcy tgtgcacgcc	1260

ES 2 583 639 T3

gatcagctga ttgcccggct cgcaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380

5 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

10 cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

15 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

ccaccaggcc aagcaccgca tcgagcgcgc ccgacccgc atggaggccg gcctggtgga 1800

20 ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860

gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

25 ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcac 2100

30 cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220

ES 2 583 639 T3

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgcttgtg ctatttgcca ataccacccc cagcatcccc 2340

5 ttcctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc 2400

ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc 2460

10 tgtattctcc tggactgca acctgtaa acgactgca atgctgatgc acggaagta 2520

gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

15 gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

20 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

25 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 3060

30 cagtcacaac ccgcaaacac tagtatggct tccgcgcat tcaccatgtc ggcgtgcccc 3120

gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg 3180

ES 2 583 639 T3

agggggcgcg cccccgactg gtccatgctg ttcgccgtga tcaccacat cttcagcgcc 3240
 gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaagctgcc ccagctgctg 3300
 5 gacgaccact tcggcctgca cggcctggtg ttccgccgca ctttcgcat ccgctcctac 3360
 gaggtgggccc ccgaccgag cacctccatc ctggccgtga tgaaccacat gcaggaggcc 3420
 10 accctgaacc acgccaagag cgtgggcatc ctgggagcag gcttcggcac caccctggag 3480
 atgtccaagc gcgacctgat gtgggtggtg cgccgcacc acgtggccgt ggagcgctac 3540
 cccacctggg gcgacaccgt ggaggtggag tgctggatcg gcgccagcgg caacaacggc 3600
 15 atgcgccgag acttcctggt gcgagactgc aagaccggcg agatcctgac ccgctgcacc 3660
 tccctgagcg tgctgatgaa caccgcacc cgccgcctga gcaccatccc cgacgaggtg 3720
 20 cgcgggcaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga cgagatcaag 3780
 aagctgcaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc 3840
 tggaacgacc tggacgtgaa ccagcacgtg aacaacctga agtacgtggc ctgggtgttc 3900
 25 gagaccgtgc ccgacagcat cttcagatcc caccacatca gctccttcac cctggagtac 3960
 cgccgcgagt gcacccgca ctccgtgctg cgagcctga ccaccgtgag cggcggcagc 4020
 30 tccgaggccg gcctggtgtg cgaccacctg ctgcagctgg agggcggcag cgaggtgctg 4080
 cgcgccccga ccgagtggcg cccaagctg accgactcct tccgggcat cagcgtgatc 4140

ES 2 583 639 T3

	cccgccgagc cccgcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac	4200
	atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta	4260
5	tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga	4320
	cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac	4380
10	gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct	4440
	tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc	4500
	tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct	4560
15	gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cggaagtag	4620
	tgggatggga acacaaatgg aaagctt	4647
20		
	<210> 123	
	<211> 4721	
	<212> ADN	
	<213> Secuencia Artificial	
25		
	<220>	
	<223> Descripción de Secuencia Artificial: Polinucleótido sintético	
30		
	<400> 123	
	ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60

ES 2 583 639 T3

	gtgcttgggc gcctgccgcc tgcttccgc atgcttgtgc tggtaggct gggcagtgc	120
	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc	180
5	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaagta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
10	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg	480
15	gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga	540
	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
20	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
25	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagagggcggc ggcagcagag cggcggcagc agcggcggg	960
30	gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020

ES 2 583 639 T3

cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt 1140
 5 ccatagtgca ttgactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgcccgct cggaagtag gcgccctcct ttctgctcg cctctctccg 1320
 10 tccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 15 cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
 cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 20 gggcgaggty cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 25 ccaccaggcc aagcaccgca tcgagcgcgc ccgacccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccccct 1860
 gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 30 caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980

ES 2 583 639 T3

	ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggagg	2040
	cgagtgggccc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgc	2100
5	cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
	atcgacacac tctggacgct ggtcgtgta tggactgttg ccgccacact tgctgccttg	2220
	acctgtgaat atccctgccc cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
10	cgcgcttttg cgagttgcta gctgcttgct ctatttgcga ataccacccc cagcatcccc	2340
	ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
15	ctcagcgtg ctctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc	2460
	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta	2520
	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
20	gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
	gcgcttgggt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
25	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
30	atgggagctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940

ES 2 583 639 T3

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
 cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 3060
 5 cagtcacaac ccgcaaacgg cgcgccatgg ccaccacctc cctggcctcc gccttctgca 3120
 gcatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc tccagcgacc 3180
 tgcagctgcg cgccggcaac gccagacct ccctgaagat gatcaacggc accaagttct 3240
 10 cctacaccga gagcctgaag aagctgcccg actggtccat gctgttcgcc gtgatcacca 3300
 ccatcttctc cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaacc 3360
 15 cccccagct gctggacgac cacttcggcc cccacggcct ggtgttcgcg cgcaccttcg 3420
 ccatccgag ctacgagtg ggccccgacc gctccaccag catcgtggcc gtgatgaacc 3480
 acctgcagga ggccgccctg aaccacgcca agtccgtggg catcctgggc gacggcttcg 3540
 20 gcaccaccct ggagatgtcc aagcgcgacc tgatctgggt ggtgaagcgc acccacgtgg 3600
 ccgtggagcg ctacccccgcc tggggcgaca ccgtggaggt ggagtgctgg gtgggagcct 3660
 25 ccggcaaaa cggccgccgc cagcacttcc tgggtgcgca ctgcaagacc ggcgagatcc 3720
 tgaccgctg cacctccctg agcgtgatga tgaacacccg caccgcccgc ctgagcaaga 3780
 tccccgagga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg 3840
 30 acgaggagat caagaagccc cagaagctga acgactccac cgccgactac atccagggcg 3900

ES 2 583 639 T3

	gcctgacccc ccgctggaac gacctggaca tcaaccagca cgtgaacaac atcaagtacg	3960
	tggactggat cctggagacc gtgcccgaca gcatcttcga gagccaccac atctcctcct	4020
5	tcaccatcga gtaccgccgc gagtgcacca tggacagcgt gctgcagtcc ctgaccaccg	4080
	tgagcggcgg ctctccgag gccggcctgg tgtgagca cctgctgag ctggagggcg	4140
	gcagcgaggt gctgcgcc aagaccgagt ggcgccccaa gctgaccgac tccttccgcg	4200
10	gcatcagcgt gatccccgcc gaggccagc tgatggacta caaggaccac gacggcgact	4260
	acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgatgactc gaggcagcag	4320
15	cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	4380
	acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	4440
	gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	4500
20	ccccagcatc cccttccctc gtttcatatc gcttgcaccc caaccgcaac ttatctacgc	4560
	tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg	4620
25	tttgggctcc gcctgtattc tcctgttact gcaacctgta aaccagcact gcaatgctga	4680
	tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	4721
30	<210> 124	
	<211> 4650	
	<212> ADN	

ES 2 583 639 T3

<213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido

5 sintético

<400> 124

	ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
10	gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgtc	120
	gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
15	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
20	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg	480
	gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgta	540
25	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
30	caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggctg cttgctgggt	720
	tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780

ES 2 583 639 T3

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

5 acgcaatgca gcacaggcag cagagggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020

10 cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcy ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140

ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

15 cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

20 tcccgcctct agaatatcaa tgatcgagca ggacggcctc cagcgggct cccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440

cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500

25 cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

30 gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc 1680

catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

ES 2 583 639 T3

	ccaccaggcc aagcaccgca tcgagcgcgc cgcacccgc atggaggccg gcctggtgga	1800
	ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
5	gaaggcccg atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
	caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggctg	1980
10	ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggctg	2040
	cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcct	2100
	cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
15	atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
	acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
20	cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc	2340
	ttccctcgtt tcatatcgtt tgcatcccaa ccgcaactta tctacgctgt cctgctatcc	2400
	ctcagcgtg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc	2460
25	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta	2520
	gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
30	gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700

ES 2 583 639 T3

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

5 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

10 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttcggggctg ccctcccggc catgggcccgc 3120

15 ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgcc agtcggcgag tcaggttttg 3180

gagagcgggc gcgccccga ctggtccatg ctgttcgccg tgatcaccac catcttctcc 3240

20 gccgccgaga agcagtggaac caacctggag tggaagccca agcccaacc ccccagctg 3300

ctggacgacc acttcggccc ccacggcctg gtgttcgcc gcaccttcgc catccgcagc 3360

tacgaggtgg gccccgaccg ctccaccagc atcgtggccg tgatgaacca cctgcaggag 3420

25 gccgccctga accacgcaa gtccgtggc atcctggcg acggcttcgg caccaccctg 3480

gagatgtcca agcgcgacct gatctgggtg gtgaagcgca cccacgtggc cgtggagcgc 3540

30 taccgccct ggggcgacac cgtggaggtg gagtgctggg tgggcgcctc cggcaacaac 3600

ggccgccgcc acgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gaccgcctgc 3660

ES 2 583 639 T3

	acctccctga gcgtgatgat gaacacccgc acccgccgcc tgagcaagat ccccgaggag	3720
	gtgcgcggcg agatcggccc cgccttcacg gacaacgtgg ccgatgaagga cgaggagatc	3780
5	aagaagcccc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc	3840
	cgctggaacg acctggacat caaccagcac gtgaacaaca tcaagtacgt ggactggatc	3900
10	ctggagaccg tgcccgacag catcttcgag agccaccaca tctcctcctt caccatcgag	3960
	taccgcccgc agtgcacat ggacagcgtg ctgcagtccc tgaccaccgt gagcggcggc	4020
	tcctccgagg ccggcctggt gtgagcagc ctgctgcagc tggagggcgg cagcgaggtg	4080
15	ctgcgcgcca agaccgagtg gcgccccaaag ctgaccgact ccttccgagg catcagcgtg	4140
	atccccgccg agtccagcgt gatggactac aaggaccacg acggcgacta caaggaccac	4200
20	gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata	4260
	gatatgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct	4320
	tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg	4380
25	tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc	4440
	ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat	4500
30	ccctcagcgc tgctcctgct cctgctcact gccctcgca cagccttggg ttgggctccg	4560
	cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag	4620

ES 2 583 639 T3

	tagtgggatg ggaacacaaa tggaaagctt	4650
5	<210> 125 <211> 4653 <212> ADN <213> Secuencia Artificial	
10	<220> <223> Descripción de Secuencia Artificial: Polinucleótido sintético	
	<400> 125	
15	ggtaccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
	gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct	120
	gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgcac	180
20	gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
	tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
25	cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
	cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
	gaatggcacc accaagtgtg tgcgagccag catcgcgcc tggctgtttc agctacaacg	480
30	gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga	540

ES 2 583 639 T3

	tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
	cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
5	caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
	tgcatggcat cgcacatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
	gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
10	atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
	acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
15	gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt	1020
	cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagccccga	1080
	atggaatgcy ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt	1140
20	ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
	cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
25	gatcagctga ttgccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320
	tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
	ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
30	cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc	1500

ES 2 583 639 T3

	cggcgcctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
	cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
5	gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc	1680
	catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttca	1740
	ccaccaggcc aagcaccgca tcgagcgcgc cgcacccgc atggaggccg gcctggtgga	1800
10	ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccgcct	1860
	gaaggcccc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
15	caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggctg	1980
	ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggctg	2040
	cgagtgggcc gaccgcttcc tggctgctgta cggcatcgcc gccccgact cccagcgc	2100
20	cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
	atcgacacac tctggacgct ggtcgtgtga tggactggtg ccgccacact tgctgccttg	2220
25	acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
	cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc	2340
	ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
30	ctcagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggtt gggctccgcc	2460

ES 2 583 639 T3

	tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acggaagta	2520
	gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
5	gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
	gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
	gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatac	2760
10	gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
	tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
15	atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
	aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
	cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt	3060
20	cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
	cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg	3180
25	gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc	3240
	tccgccgccg agaagcagtg gaccaacctg gagtggaaagc ccaagcccaa cccccccag	3300
	ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc	3360
30	agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag	3420

ES 2 583 639 T3

	gaggccgcc tgaaccacgc caagtccgtg ggcatacctgg gcgacggctt cggcaccacc	3480
	ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcaccacgt ggccgtggag	3540
5	cgctaccccg cctggggcga caccgtggag gtggagtgct gggtgggcgc ctccggcaac	3600
	aacggccgcc gccacgactt cctggtgcmc gactgcaaga ccggcgagat cctgacccgc	3660
	tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag	3720
10	gaggtgcmc gcgagatcgg cccgccttc atcgacaacg tggccgtgaa ggacgaggag	3780
	atcaagaagc cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc	3840
15	ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg	3900
	atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc	3960
	gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc	4020
20	ggctcctccg aggccggcct ggtgtgcmg cacctgctgc agctggaggg cggcagcmg	4080
	gtgctgcmc ccaagaccga gtggcgcctc aagctgaccg actccttccg cggcatcagc	4140
25	gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
	cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
	atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg	4320
30	ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt	4380

ES 2 583 639 T3

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgccaatacc acccccagca 4440

tccccttccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc 4500

5 tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctgcta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

10

<210> 126

<211> 3669

<212> ADN

15 <213> Secuencia Artificial

<220>

<223> Descripción de Secuencia Artificial: Polinucleótido
sintético

20

<400> 126

cccgtgatca cacaggtgcc ttgcgagcgt gatcacacta ttttgggggt cctacagtac 60

tgaaatggtg agaagtcgta ctgaaatcaa ggatgaacaa tgaaaatggt gctgtggtgg 120

25 cttctcaaag gtcaagaatc agtcgctcgc gtcaggaaat cgcggcgtca accagcgtgg 180

gcgcggtcag tggccccgca ctggtcacca tagcctctcc tgccacagta gcgatcccct 240

30 gggcgttcac tctcagcagc ggctgtactg cctcccagat tttcttcttc tggacctgcg 300

ggcgtgagag gatgagcagg gtggggccaa gggctcaatc ctgaacggcc ctcatcgggt 360

ES 2 583 639 T3

	ttccaatccc acaacacata cccacagcag gtcagaccac gcattcgcac catgcgacc	420
	aaataacgtg tccttacctg attgggtgtg gcaggctccg tggacaggag tgccctgtcc	480
5	cccgccaga cccgctcccc cgtcacggcg gcgtccggga cccgcagcgg ctccaccgcg	540
	gtgtgatccg cgttggcggc gcagagcagc atcccagccg atttgacccc gcgcatgctc	600
10	cgaggcttga ggttggccag caccaccacc cgccggccga caaggtcctc cagggtcacg	660
	tgccggacca ggccactcac gatggtgcga gggccccct cctcgccgag gtcgatctgc	720
	tcgacgtaca gactgcgaca tgcgtggcga gtggtcatca gaaggaagca ggtgtgcaga	780
15	aggggcacgt ggttggtatt gagagtagcc aaagctttgt gccaatcaga aagtcaacgc	840
	agctgcctgc ctggctcgcg tacaattcct ttcttgcgct atgacacttc cagcaaaagg	900
20	tagggcgggc tgcgagacgg cttcccggcg ctgcatgcaa caccgatgat gcttcgaccc	960
	cccgaagctc cttcggggct gcatgggcbc tccgatgccg ctccagggcg agcgctgttt	1020
	aaatagccag gccccgatt gcaaagacat tatagcgagc taccaaagcc atattcaaac	1080
25	acctagatca ctaccacttc tacacaggcc actcgagctt gtgatcgcac tccgctaagg	1140
	gggcgctctt tcctcttcgt ttcagtcaca acccgcaaac ggcgcgccat gctgctgcag	1200
30	gccttcctgt tcctgctggc cggcttcgcc gccaaatca gcgcctccat gacgaacgag	1260
	acgtccgacc gccccctggt gcacttcacc cccaacaagg gctggatgaa cgaccccaac	1320

ES 2 583 639 T3

	ggcctgtggt acgacgagaa ggacgccaag tggcacctgt acttccagta caacccgaac	1380
	gacaccgtct gggggacgcc cttgttctgg ggccacgcca cgtccgacga cctgaccaac	1440
5	tgggaggacc agcccatcgc catcgccccg aagcgcaacg actccggcgc cttctccggc	1500
	tccatggtgg tggactacaa caacacctcc ggcttcttca acgacacat cgacccgcgc	1560
10	cagcgctgcg tggccatctg gacctacaac accccggagt ccgaggagca gtacatctcc	1620
	tacagcctgg acggcggcta caccttcacc gagtaccaga agaaccccgt gctggccgcc	1680
	aactccacc agttccgcga cccgaaggtc ttctggtacg agccctcca gaagtggatc	1740
15	atgaccgcgg ccaagtcca ggactacaag atcgagatct actcctccga cgacctgaag	1800
	tcctggaagc tggagtccgc gttcgccaac gagggcttcc tcggctacca gtacgagtgc	1860
20	cccggcctga tcgaggtccc caccgagcag gacccagca agtcctactg ggtgatgttc	1920
	atctccatca accccggcgc cccggccggc ggctccttca accagtactt cgtcggcagc	1980
	ttcaacggca cccacttcga ggccttcgac aaccagtccc gcgtggtgga cttcggcaag	2040
25	gactactacg ccctgcagac cttcttcaac accgaccga cctacgggag cgccttggc	2100
	atcgcgtggg cctccaactg ggagtactcc gccttcgtgc ccaccaacc ctggcgctcc	2160
30	tccatgtccc tcgtgcgcaa gttctccctc aacaccgagt accaggccaa cccggagacg	2220
	gagctgatca acctgaaggc cgagccgatc ctgaacatca gcaacgccgg cccctggagc	2280

ES 2 583 639 T3

	cggttcgcca ccaacaccac gttgacgaag gccaacagct acaacgtcga cctgtccaac	2340
	agcaccggca ccctggagtt cgagctggtg tacgccgtca acaccacca gacgatctcc	2400
5	aagtccgtgt tcgcggaact ctccctctgg ttcaagggcc tggaggacct cgaggagtac	2460
	ctccgcatgg gcttcgaggt gtccgctcc tccttcttcc tggaccgagg gaacagcaag	2520
10	gtgaagtctg tgaaggagaa cccctacttc accaaccgca tgagcgtgaa caaccagccc	2580
	ttcaagagcg agaacgacct gtcctactac aaggtgtacg gcttgctgga ccagaacatc	2640
	ctggagctgt acttcaacga cggcgacgtc gtgtccacca acacctactt catgaccacc	2700
15	gggaacgccc tgggctccgt gaacatgacg acgggggtgg acaacctggt ctacatcgac	2760
	aagttccagg tgcgagaggt caagtgatta attaactcga ggcagcagca gctcggatag	2820
20	tatcgacaca ctctggacgc tggctgtgtg atggactggt gccgccacac ttgctgcctt	2880
	gacctgtgaa tatccctgcc gcttttatca aacagcctca gtgtgtttga tcttgtgtgt	2940
	acgcgctttt gcgagttgct agctgcttgt gctatcttgc aataccaccc ccagcatccc	3000
25	cttccctcgt ttcatatcgc ttgcatcca accgcaactt atctacgctg tcttgcctatc	3060
	cctcagcgt gctcctgctc ctgctcactg cccctcgcac agccttgggt tgggctccgc	3120
30	ctgtattctc ctggtactgc aacctgtaaa ccagcactgc aatgctgatg cacgggaagt	3180
	agtgggatgg gaacacaaat ggaaagcttg agctcggtag ccgtacccat cagcatccgg	3240

ES 2 583 639 T3

gtgaatcttg gcctccaaga tatggccaat cctcacatcc agcttgcaaa aatcgactag 3300

actgtctgca agtgggaatg tggagcaciaa ggttgcttgt agcgatcgac agactggtgg 3360

5 ggtacattga caggtgggca gcgccgcac ccatcgtgcct gacgcgagcg ccgccggttg 3420

ctcgcccgtg cctgccgtca aagagcggca gagaaatcgg gaaccgaaaa cgtcacattg 3480

10 cctgatgttg ttacatgctg gactagactt tcttggcgtg ggtctgctcc tcgccaggtg 3540

cgcgacgcct cggggctggg tgcgaggag cctgctggcc acgcatttga caagacccaa 3600

agctcgcac tcagacggtc aaccgttcgt attatacatt caacatatgg tacatagca 3660

15 aaaagcatg 3669

<210> 127

20 <211> 39

<212> PRT

<213> *Prototheca moriformis*

<400> 127

25 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

30 20 25 30

ES 2 583 639 T3

Glu Ser Gly Arg Ala Gln Leu

35

5 <210> 128

<211> 40

<212> PRT

<213> Prototheca moriformis

10 <400> 128

Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr

1

5

10

15

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

20

35

40

<210> 129

<211> 36

25 <212> PRT

<213> Prototheca moriformis

<400> 129

Met Ala Ser Ala Ala Phe Thr Met Ser Ala Cys Pro Ala Met Thr Gly

30

1

5

10

15

ES 2 583 639 T3

Arg Ala Pro Gly Ala Arg Arg Ser Gly Arg Pro Val Ala Thr Arg Leu
20 25 30

5

Arg Gly Arg Ala
35

10

<210> 130
<211> 40
<212> PRT
<213> *Chlorella protothecoides*

15

<400> 130
Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15

20

Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30

25

Pro Val Arg Gly Arg Ala Gln Leu
35 40

30

<210> 131
<211> 87
<212> PRT
<213> *Cuphea hookeriana*

ES 2 583 639 T3

<400> 131

Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

5

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

10

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

15

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

20

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

25

<211> 60

<212> PRT

<213> *Umbellularia californica*

<400> 132

30

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

ES 2 583 639 T3

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

5

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

10

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu
50 55 60

<210> 133

15

<211> 60

<212> PRT

<213> Cinnamomum camphora

<400> 133

20

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

25

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

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

ES 2 583 639 T3

50

55

60

<210> 134

5

<211> 1104

<212> ADN

<213> Prototheca moriformis

<400> 134

10

atggcaccga ccagcctgct tgccagtact ggcgtctctt ccgcttctct gtggtcctct 60

gcgcgctcca gcgcggtgcg ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg 120

ccgctgcca tgcagcgcg ctgcttccga acagtggcgg tcagggccgc acccgcgta 180

15

gccgtccgtc cggaaccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg 240

gaggacaagc gcatcttctt ggaggagcac cgcattcggg gcaacgaggt gggcccctcg 300

20

cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg 360

gtggccatgt ggggccggag ctcggagggt ttcgacgag acccggagct gcaggaggcg 420

ggtctcatct ttgtgatgac gcgcatgcag atccaaatgt accgctaccc gcgctggggc 480

25

gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag 540

tgggtgctgc gcgacaagct gaccggcag gcgctgggcg cggccacctc cagctgggtc 600

30

atgatcaaca tccgcacgcg ccggccgtgc cgcgatgccg agctcgtccg cgtcaagtcg 660

gccttcttcg gcgcgagcc gccgcgctg gcgctgccgc ccacggtcac gcgcgccaag 720

ES 2 583 639 T3

ctgccaaca tcgacgacc ggcgccgtg cgcgggcacc gccaggtcgc gcgccgacc 780
 gacatggaca tgaacgggca cgtgaacaac gtggcctacc tggcctggtg cctggaggcc 840
 5
 gtgcccgagc acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc 900
 gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg 960
 10
 ctcacgcaca acggcgccgg ccgcaacccc tcctgcttcg tccatagcat tctgcgcgcc 1020
 gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc 1080
 aagccgcca aggcgagcca ctga 1104
 15
 <210> 135
 <211> 1104
 <212> ADN
 <213> Prototheca moriformis
 20
 <400> 135
 atggcaccga ccagcctgct tgcccgtact ggcgtctctt ccgcttctct gtgctcctct 60
 acgcgctccg gcgcgtgcmc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg 120
 25
 ccgctgcca tgcagcggcg ctgcttccga acagtggctg tcagggccgc acccgagta 180
 gccgtccgtc cggaaccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg 240
 30
 gaggacaagc gcatcttcct ggaggagcac cgcattcgtg gcaacgaggt gggcccctcg 300
 cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg 360

ES 2 583 639 T3

gtggccatgt ggggtcggag ctcgagggt ttcgcgacgg acccgagct gcaggaggcg 420

ggcctcatct ttgtgatgac gcgcatgcag atccaaatgt accgctaccc gcgctggggc 480

5 gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag 540

tgggtgctgc gcgacaagct gaccggcgag gcgctggggc cggccacctc cagctgggtc 600

10 atgatcaaca tccgcacgcg ccggccgtgc cgcattgccc agctcgtccg cgtcaagtcg 660

gccttcttcg cgcgcgagcc gccgcgcctg gcgctgccgc ccgcggtcac gcgtgccaag 720

ctgcccaca tccgcacgcg gccgcgcctg cgcgggcacc gccaggtcgc gcgccgcacc 780

15 gacatggaca tgaacggcca cgtgaacaac gttgcctacc tggcctggtg cctggaggcc 840

gtgcccgagc acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc 900

20 gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg 960

ctcacgcaca acggcgccgg ccgcaacccc tctgcttcg tccatagcat tctgcgcgcc 1020

gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc 1080

25 aagccgcca aggcgagcca ctga 1104

<210> 136

30 <211> 367

<212> PRT

<213> *Prototheca moriformis*

ES 2 583 639 T3

<400> 136

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

5

Leu Trp Ser Ser Ala Arg Ser Ser Ala Cys Ala Phe Pro Val Asp His
20 25 30

10

Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
35 40 45

15

Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
50 55 60

20

Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65 70 75 80

25

Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
85 90 95

30

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

ES 2 583 639 T3

Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
130 135 140

5

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145 150 155 160

10

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

15

Gly Ala Ala Thr Ser Ser Trp Val Met Ile Asn Ile Arg Thr Arg Arg
195 200 205

20

Pro Cys Arg Met Pro Glu Leu Val Arg Val Lys Ser Ala Phe Phe Ala
210 215 220

25

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Thr Val Thr Arg Ala Lys
225 230 235 240

30

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
245 250 255

ES 2 583 639 T3

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
260 265 270

5 Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
275 280 285

10 Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
290 295 300

15 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

20 Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
340 345 350

25 Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
355 360 365

<210> 137

30 <211> 367

<212> PRT

<213> Prototheca moriformis

ES 2 583 639 T3

<400> 137

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

5

Leu Cys Ser Ser Thr Arg Ser Gly Ala Cys Ala Phe Pro Val Asp His
20 25 30

10

Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
35 40 45

15

Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
50 55 60

20

Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65 70 75 80

25

Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
85 90 95

30

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

ES 2 583 639 T3

Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
130 135 140

5

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145 150 155 160

10

Asp Leu Met Gln Val Glu Thr Trp Phe Gln Thr Ala Gly Lys Leu Gly
165 170 175

15

Ala Gln Arg Glu Trp Val Leu Arg Asp Lys Leu Thr Gly Glu Ala Leu
180 185 190

20

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

25

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Ala Val Thr Arg Ala Lys
225 230 235 240

30

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
245 250 255

ES 2 583 639 T3

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
260 265 270

5 Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
275 280 285

10 Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
290 295 300

15 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

20 Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
340 345 350

25 Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
355 360 365

<210> 138

30 <211> 328

<212> PRT

<213> Cuphea hookeriana

ES 2 583 639 T3

<400> 138

Gln Leu Pro Asp Trp Ser Arg Leu Leu Thr Ala Ile Thr Thr Val Phe

1 5 10 15

5

Val Lys Ser Lys Arg Pro Asp Met His Asp Arg Lys Ser Lys Arg Pro

20 25 30

10

Asp Met Leu Val Asp Ser Phe Gly Leu Glu Ser Thr Val Gln Asp Gly

35 40 45

15

Leu Val Phe Arg Gln Ser Phe Ser Ile Arg Ser Tyr Glu Ile Gly Thr

50 55 60

20

Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu Thr

65 70 75 80

25

Ser Leu Asn His Cys Lys Ser Thr Gly Ile Leu Leu Asp Gly Phe Gly

85 90 95

30

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

ES 2 583 639 T3

Ile Asn Thr Arg Phe Ser Arg Leu Gly Lys Ile Gly Met Gly Arg Asp
130 135 140

5

Trp Leu Ile Ser Asp Cys Asn Thr Gly Glu Ile Leu Val Arg Ala Thr
145 150 155 160

10

Ser Ala Tyr Ala Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys Leu
165 170 175

15

Pro Tyr Glu Val His Gln Glu Ile Val Pro Leu Phe Val Asp Ser Pro
180 185 190

20

Val Ile Glu Asp Ser Asp Leu Lys Val His Lys Phe Lys Val Lys Thr
195 200 205

25

Gly Asp Ser Ile Gln Lys Gly Leu Thr Pro Gly Trp Asn Asp Leu Asp
210 215 220

30

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

ES 2 583 639 T3

Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Glu Ser Val
260 265 270

5 Thr Ala Met Asp Pro Ser Lys Val Gly Val Arg Ser Gln Tyr Gln His
275 280 285

10 Leu Leu Arg Leu Glu Asp Gly Thr Ala Ile Val Asn Gly Ala Thr Glu
290 295 300

15 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

20

<210> 139

<211> 322

<212> PRT

<213> Umbellularia californica

25

<400> 139

Pro Asp Trp Ser Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala
1 5 10 15

30

Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Lys Leu
20 25 30

ES 2 583 639 T3

Pro Gln Leu Leu Asp Asp His Phe Gly Leu His Gly Leu Val Phe Arg
35 40 45

5

Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr
50 55 60

10

Ser Ile Leu Ala Val Met Asn His Met Gln Glu Ala Thr Leu Asn His
65 70 75 80

15

Ala Lys Ser Val Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu
85 90 95

20

Met Ser Lys Arg Asp Leu Met Trp Val Val Arg Arg Thr His Val Ala
100 105 110

25

Val Glu Arg Tyr Pro Thr Trp Gly Asp Thr Val Glu Val Glu Cys Trp
115 120 125

30

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

ES 2 583 639 T3

Leu Met Asn Thr Arg Thr Arg Arg Leu Ser Thr Ile Pro Asp Glu Val

165 170 175

5

Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp

180 185 190

10

Asp Glu Ile Lys Lys Leu Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr

195 200 205

15

Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln

210 215 220

20

His Val Asn Asn Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro

225 230 235 240

25

Asp Ser Ile Phe Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr

245 250 255

30

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

ES 2 583 639 T3

Leu Glu Gly Gly Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro
290 295 300

5

Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro
305 310 315 320

10

Arg Val

<210> 140

<211> 345

15

<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

20

Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Asn Pro
20 25 30

25

Pro Gln Leu Leu Asp Asp His Phe Gly Pro His Gly Leu Val Phe Arg
35 40 45

30

Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr
50 55 60

ES 2 583 639 T3

Ser Ile Val Ala Val Met Asn His Leu Gln Glu Ala Ala Leu Asn His
65 70 75 80

5

Ala Lys Ser Val Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu
85 90 95

10

Met Ser Lys Arg Asp Leu Ile Trp Val Val Lys Arg Thr His Val Ala
100 105 110

15

Val Glu Arg Tyr Pro Ala Trp Gly Asp Thr Val Glu Val Glu Cys Trp
115 120 125

20

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

25

Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val
145 150 155 160

30

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

ES 2 583 639 T3

Glu Glu Ile Lys Lys Pro Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr
195 200 205

5

Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Ile Asn Gln
210 215 220

10

His Val Asn Asn Ile Lys Tyr Val Asp Trp Ile Leu Glu Thr Val Pro
225 230 235 240

15

Asp Ser Ile Phe Glu Ser His His Ile Ser Ser Phe Thr Ile Glu Tyr
245 250 255

20

Arg Arg Glu Cys Thr Met Asp Ser Val Leu Gln Ser Leu Thr Thr Val
260 265 270

25

Ser Gly Gly Ser Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln
275 280 285

30

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

ES 2 583 639 T3

Ser Val Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp

325

330

335

5

Ile Asp Tyr Lys Asp Asp Asp Asp Lys

340

345

10

<210> 141

<211> 4

<212> PRT

<213> Secuencia Artificial

15

<220>

<223> Descripción de Secuencia Artificial: Péptido
sintético

20

<400> 141

Lys Asp Glu Leu

1

25

<210> 142

<211> 6

<212> PRT

<213> Desconocido

30

<220>

<223> Descripción de Desconocido: Secuencia de acil-ACP tioesterasa grasa de planta superior

<400> 142

Leu Asp Met Asn Gln His

1 5

5

<210> 143

<211> 6

<212> PRT

10

<213> Desconocido

<220>

<223> Descripción de Desconocido: Secuencia de acil-ACP tioesterasa grasa de alga

15

<400> 143

Met Asp Met Asn Gly His

1 5

20

REIVINDICACIONES

1. Una célula del género *Prototheca* que comprende un gen exógeno donde dicho gen exógeno está en unión operativa con un promotor y codifica una sacarosa invertasa.
- 5 2. La célula de conformidad con la reivindicación 1, en donde la secuencia de nucleótidos del promotor es endógena a una especie del género *Prototheca*.
3. La célula de conformidad con cualquier reivindicación precedente, en donde 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, por ejemplo en donde la secuencia de nucleótidos del promotor comprende un segmento de 50 o más nucleótidos de una secuencia de nucleótidos de las sec. con núm. ident.: 91-102.
- 10 4. La célula de conformidad con cualquier reivindicación precedente, en donde la célula tiene una secuencia ARNr 23S con al menos 75% de identidad de nucleótido para una secuencia de nucleótidos de SEQ ID NOs: 11-19.
5. La célula de conformidad con cualquier reivindicación precedente, en donde la célula es *Prototheca moriformis*, *Prototheca krugani*, *Prototheca stagnora*, o *Prototheca zopfii*.
6. La célula de la reivindicación 5, en donde la célula es *Prototheca moriformis*.
- 15 7. El uso de las células de conformidad con cualquiera de las reivindicaciones 1-6 para producir un lípido, opcionalmente en donde el lípido se utiliza además para la producción de un aceite, combustible o producto oleo químico.

Figura 1

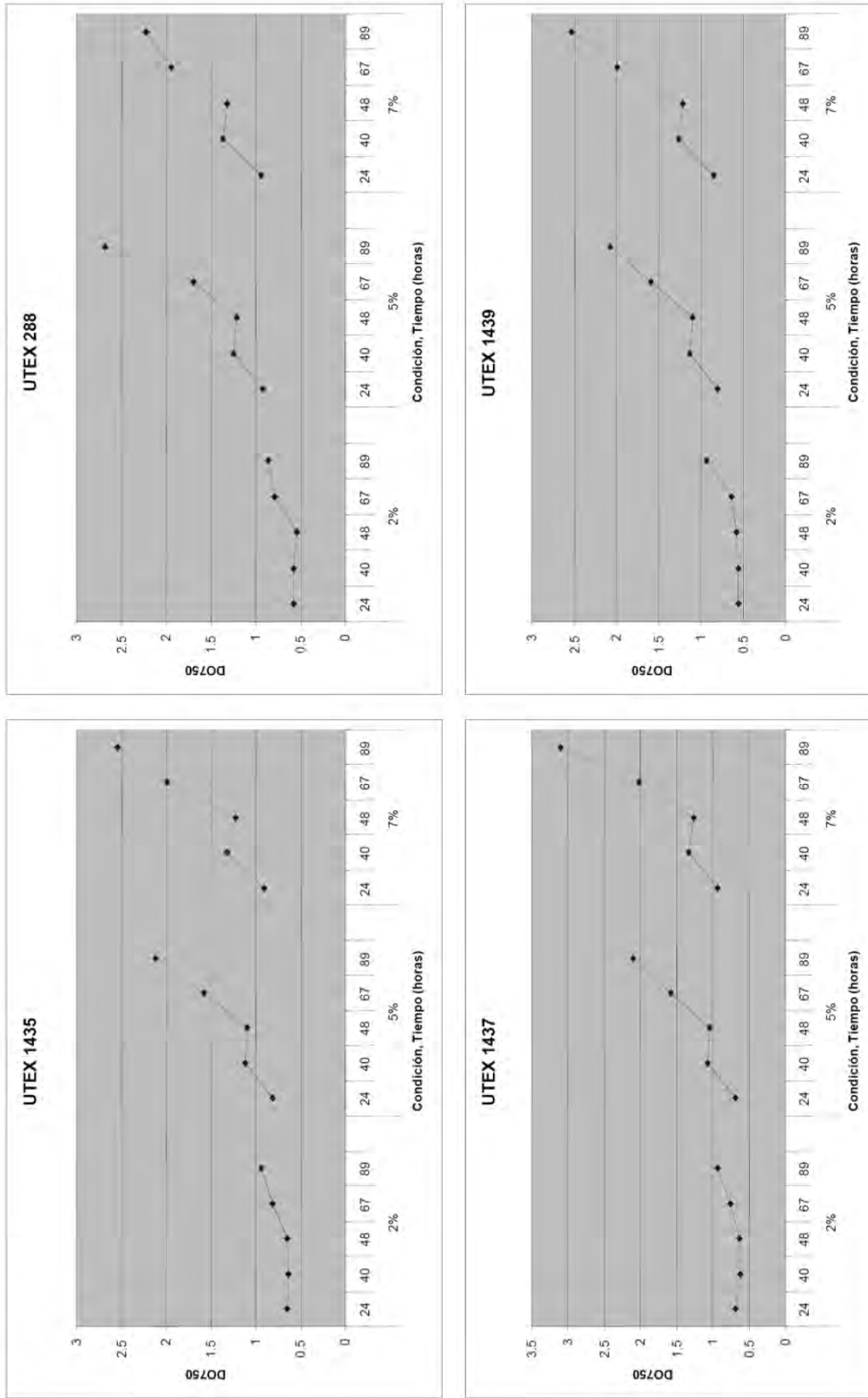


Figura 2

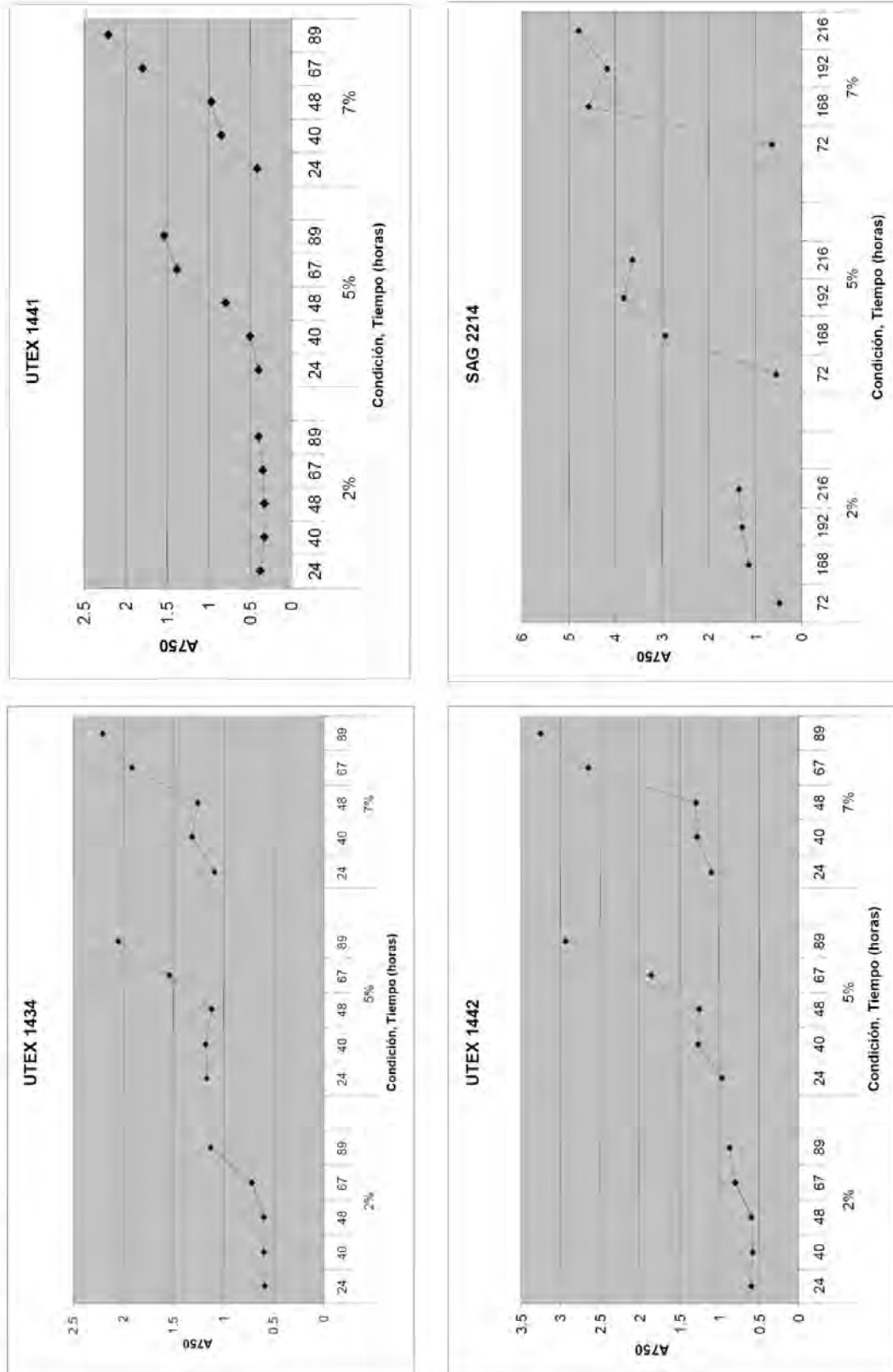


Figura 3

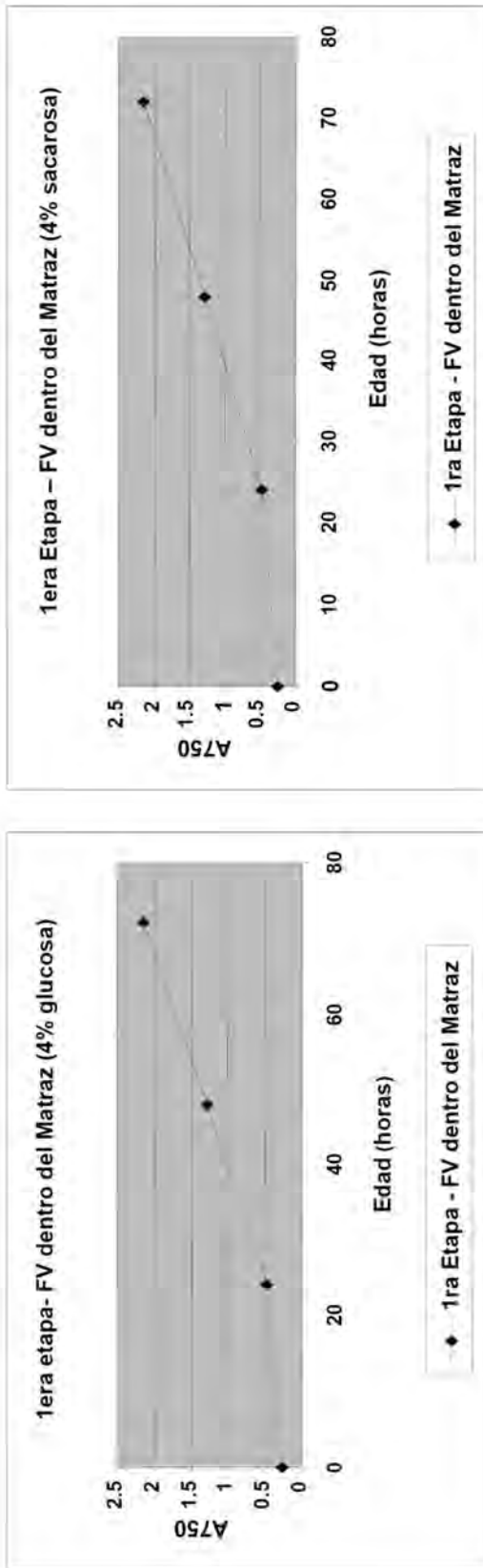


Figura 4

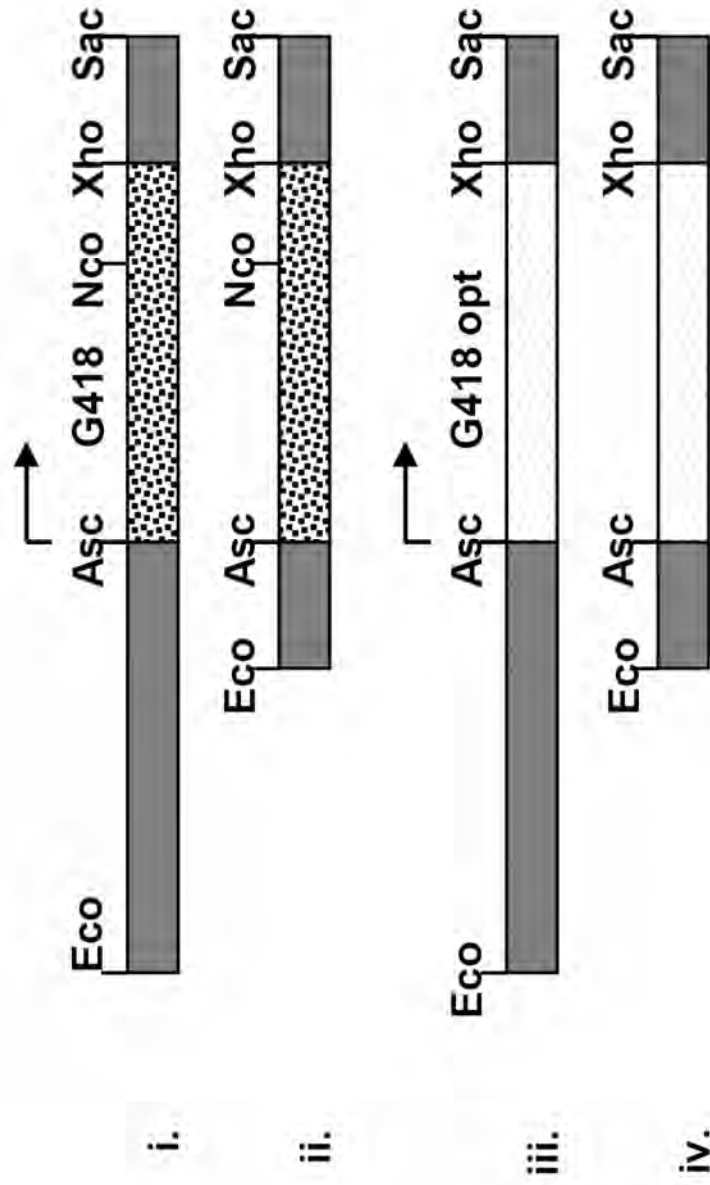


Figura 5

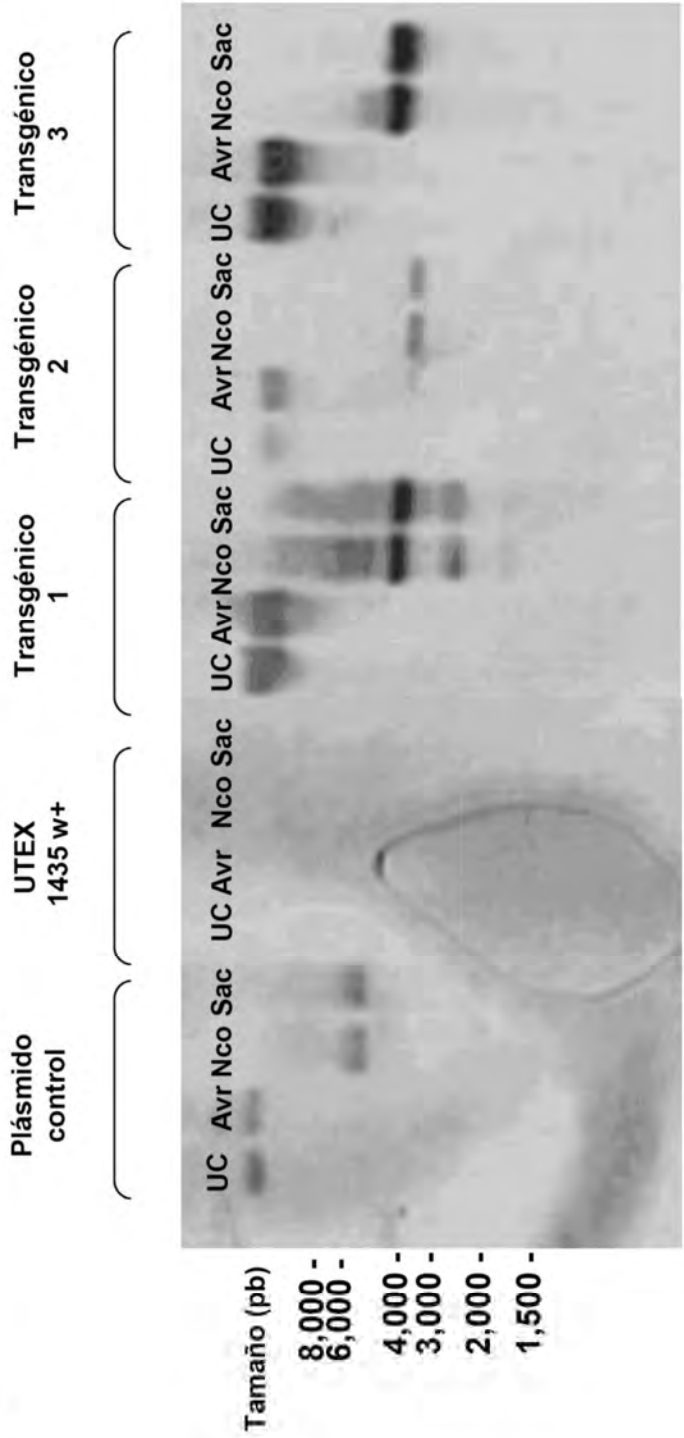


Figura 6

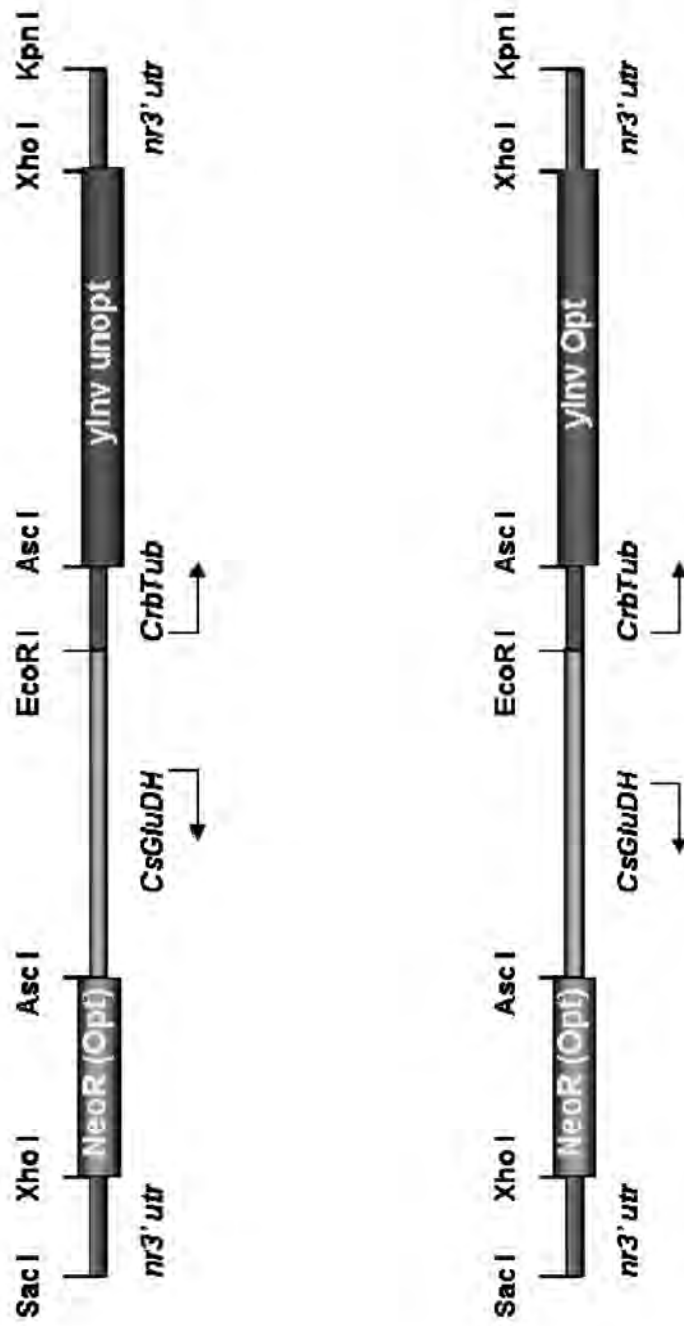


Figura 7a

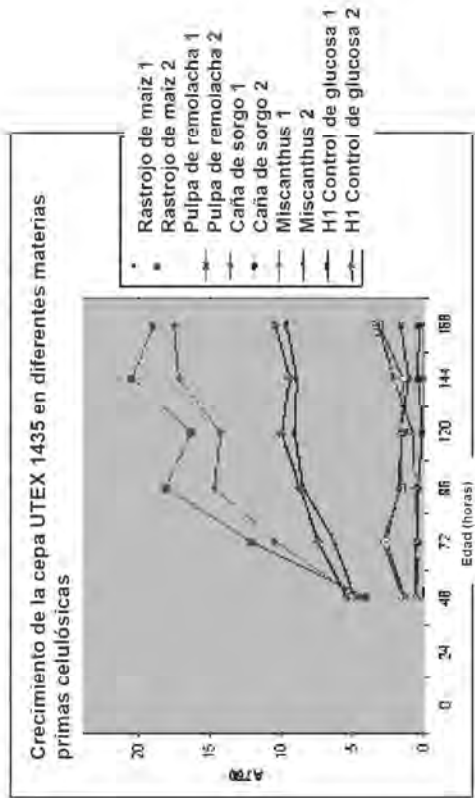


Figura 7c

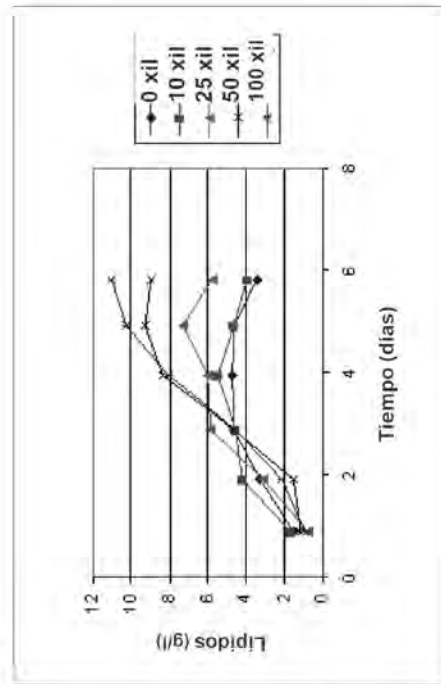


Figura 7b

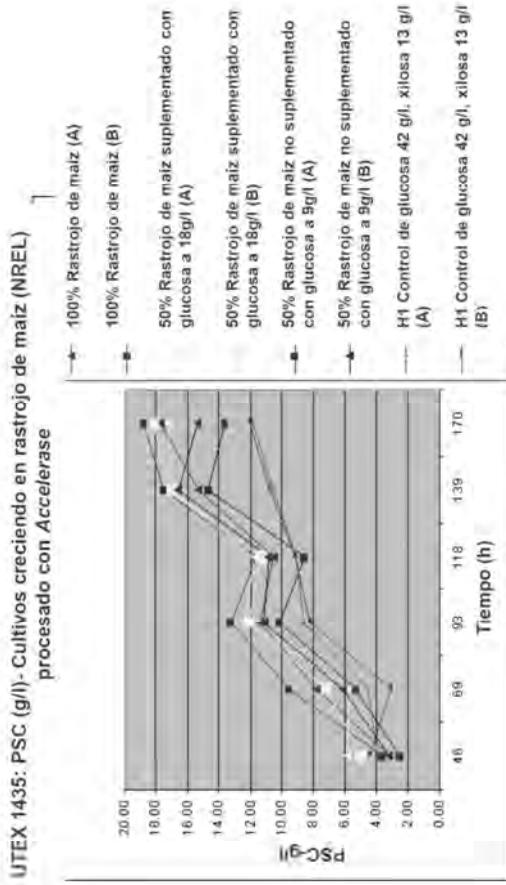


Figura 7d

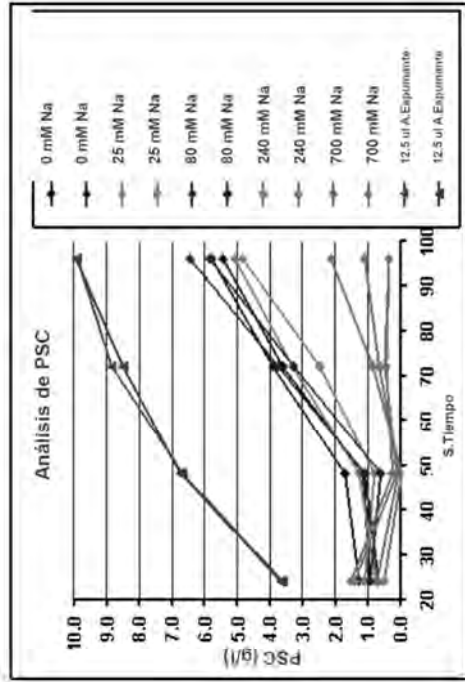


Figura 8

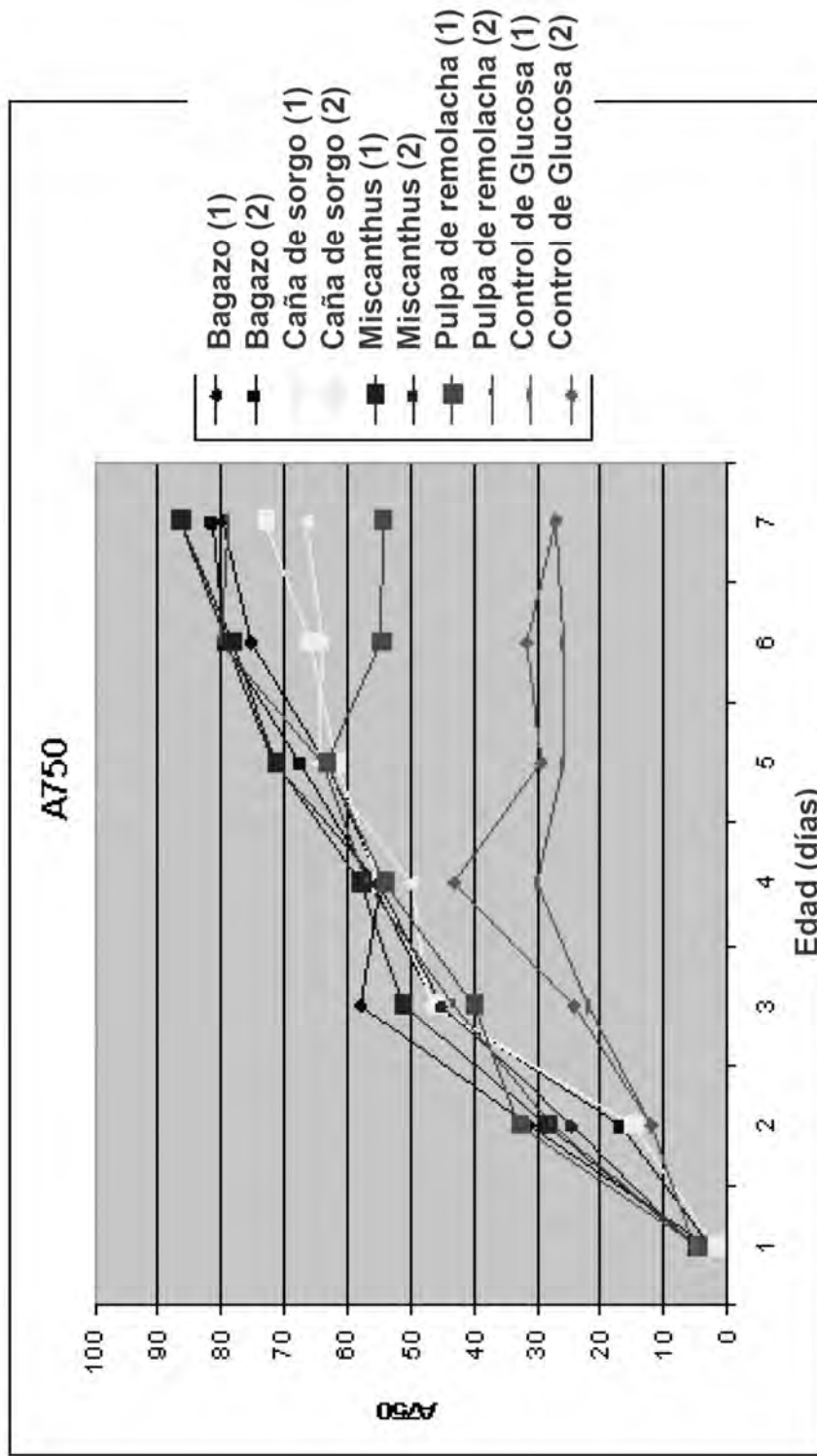


Figura 9

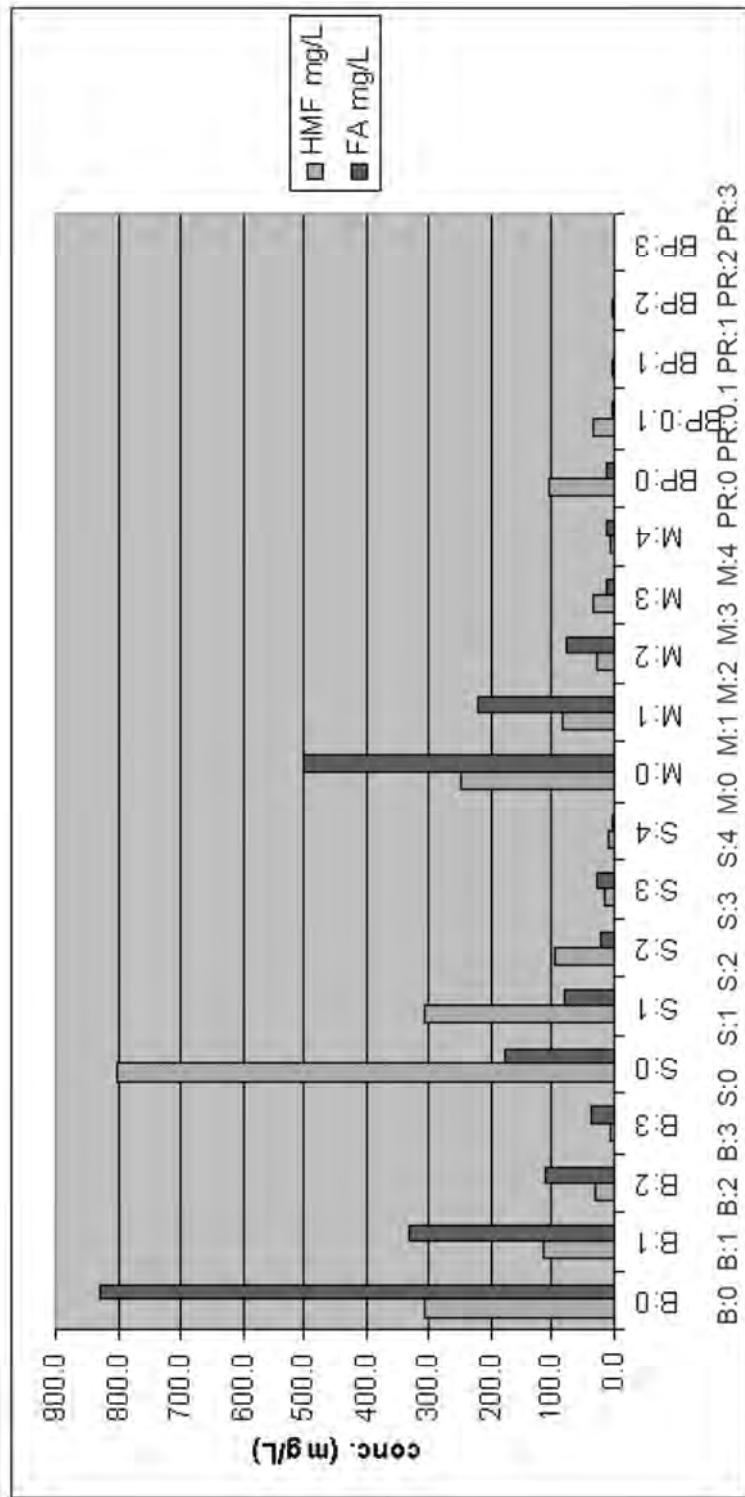


Figura 10

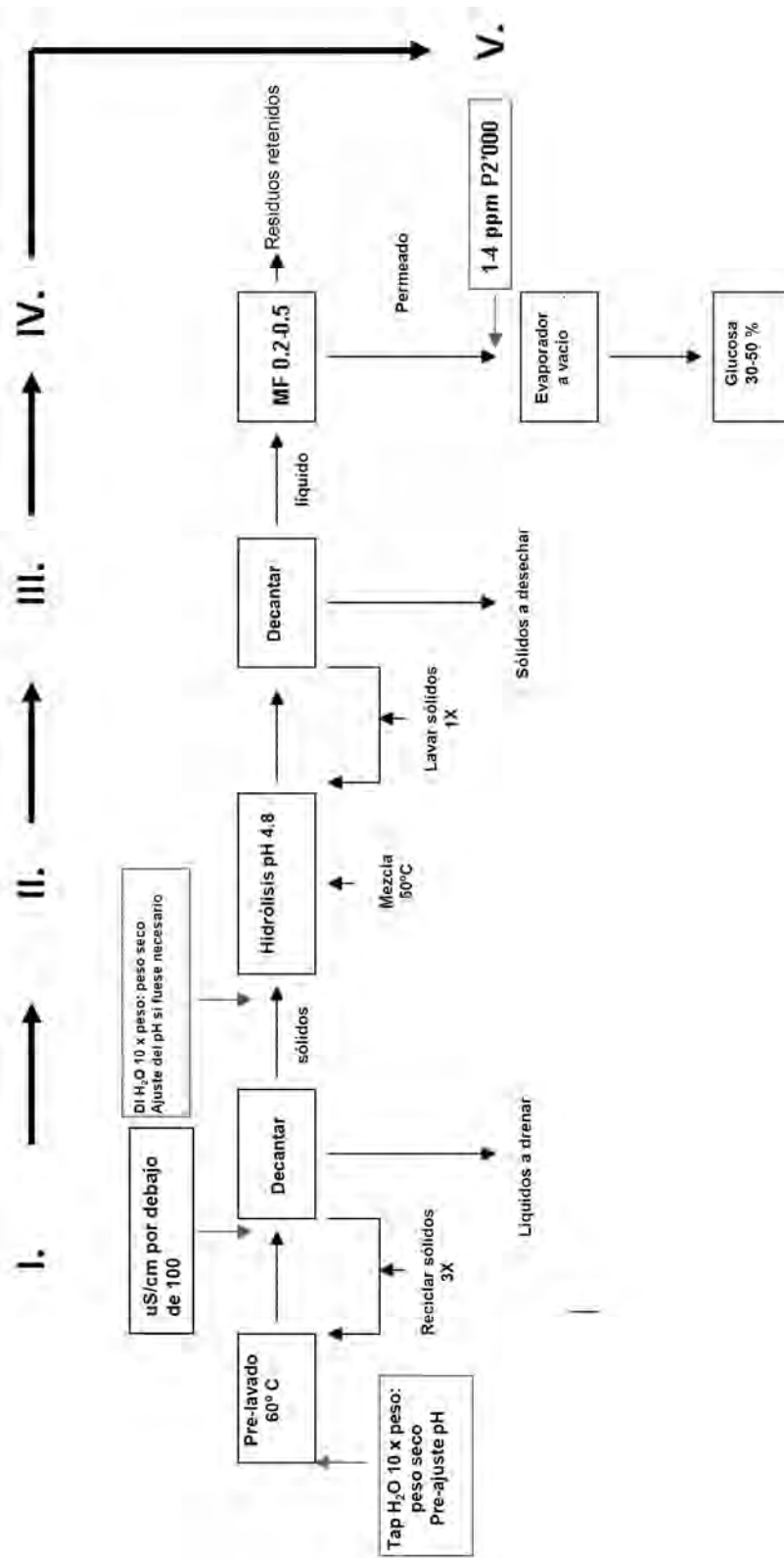


Figura 11

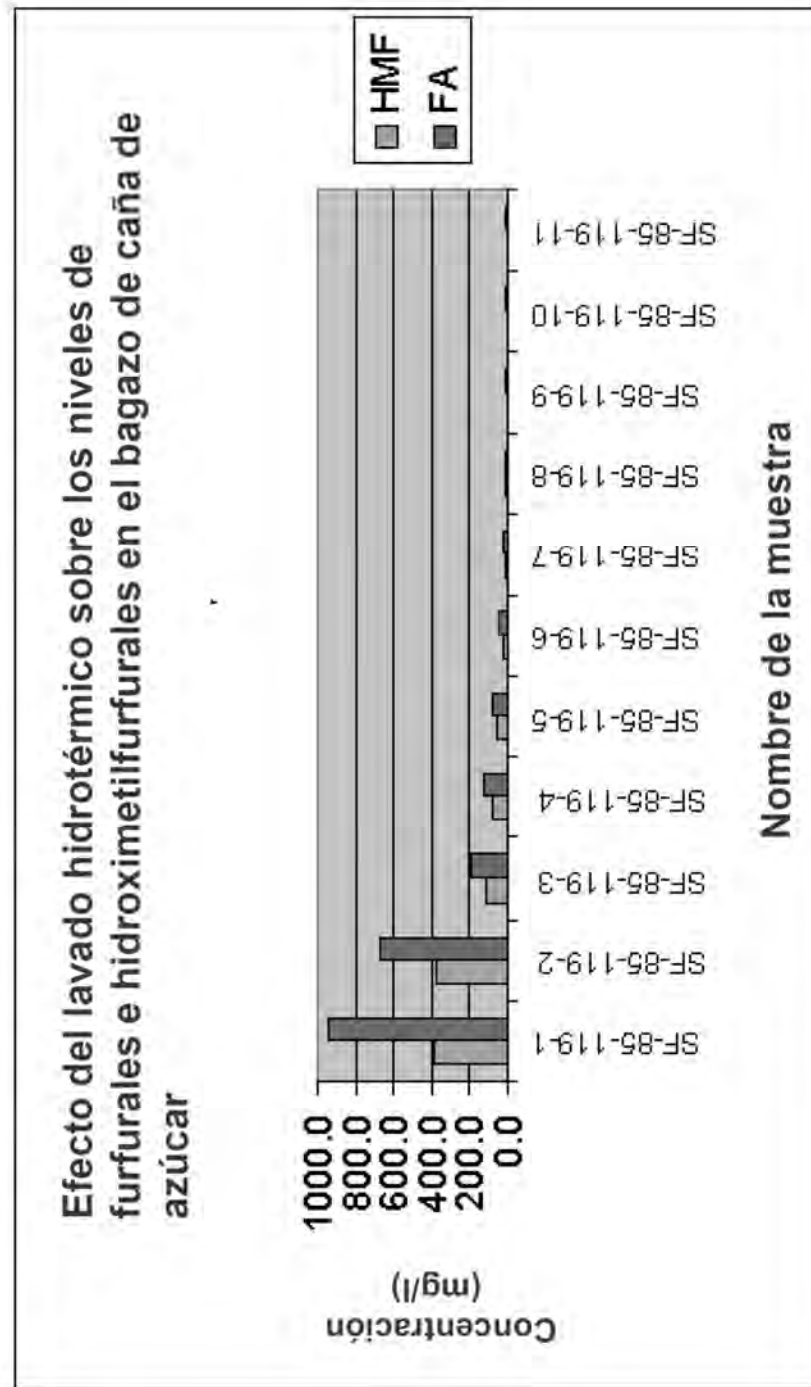


Figura 12

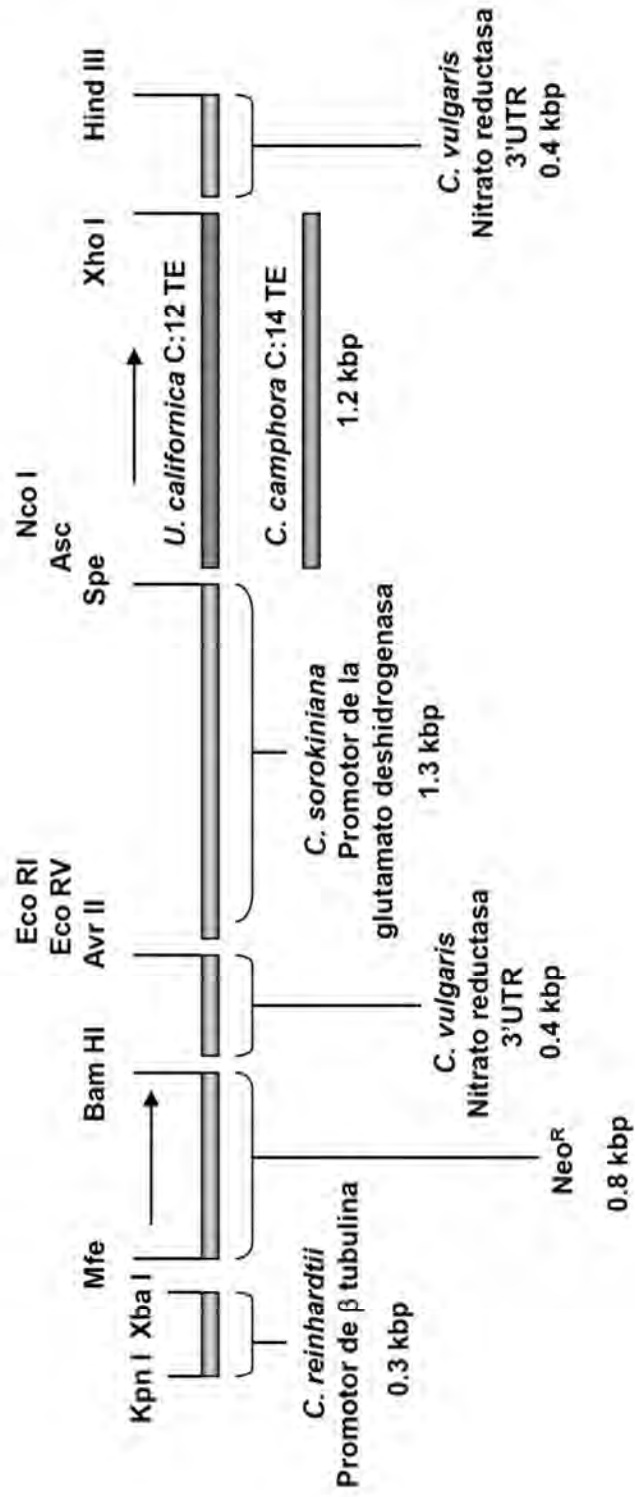


Figura 13

