(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 795 842**

(51) Int. CI.:

C12P 7/64 (2006.01)
C12N 9/10 (2006.01)
C12N 9/16 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **24.02.2012 PCT/US2012/026566**

(87) Fecha y número de publicación internacional: **29.08.2013 WO13126076**

(96) Fecha de presentación y número de la solicitud europea: **24.02.2012 E 12869541 (8)**

(97) Fecha y número de publicación de la concesión europea: **22.04.2020 EP 2814972**

(54) Título: **Producción mejorada de ácidos grasos y derivados de ácidos grasos por microorganismos recombinantes**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
24.11.2020

(73) Titular/es:

EXXONMOBIL RESEARCH AND ENGINEERING COMPANY (100.0%)
1545 Route 22 East P.O. Box 900
Annandale, NJ 08801-0900, US

(72) Inventor/es:

SESHADRI, REKHA;
KIT, JENNIE, LAU;
BAUMAN, NICHOLAS y
BROWN, ROBERT, CHRISTOPHER

(74) Agente/Representante:

ISERN JARA, Jorge

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 mejorada de ácidos grasos y derivados de ácidos grasos por microorganismos recombinantes

REFERENCIA A UN LISTADO DE SECUENCIAS

Esta solicitud contiene referencias a secuencias de aminoácidos y/o a secuencias de ácidos nucleicos que se han presentado simultáneamente con el presente documento como archivo de texto de listado de secuencias "60930641_1.txt", con un tamaño de archivo de 126 kilobytes (kb), creado el 24 de febrero de 2012.

Campo de la invención

La presente invención se refiere a los campos de bioingeniería, bioquímica metabólica y biología molecular. En particular, la invención se refiere a la producción de productos de ácidos grasos en un microorganismo recombinante o célula hospedadora genomodificada para reducir o minimizar la inhibición por retroalimentación de enzimas de la síntesis de ácidos grasos y a métodos de producción de productos de ácidos grasos utilizando dichos microorganismos recombinantes o células hospedadoras.

Antecedentes de la invención

La producción de fuentes renovables para una variedad de combustibles y productos químicos es de gran importancia en un mundo con una creciente demanda de dichos productos. Mientras que el petróleo es un producto de plantas en descomposición y otras materias que se ha incubado balo la superficie terrestre durante millones de años, algunos esfuerzos actuales se centran en el uso directo de plantas y otros organismos para generar, por ejemplo, lípidos, entre los que se puede incluir ácidos grasos y sus derivados, para uso en las industrias de combustible y química. Los ácidos grasos están compuestos por cadenas de alquilo largas, similar al petróleo, y son un metabolito primario utilizado por las células para las funciones de almacenamiento químico y energético. La mejora de la adaptabilidad, controlabilidad y rentabilidad de la producción de ácidos grasos y derivados de ácidos grasos (en conjunto "productos de ácidos grasos") sería beneficiosa para el desarrollo de energía renovable y fuentes químicas.

La síntesis de ácidos grasos en una célula es un ciclo repetitivo en donde la proteína transportadora de acilo y malonilo (malonil-ACP, por las siglas del inglés *malonyl-acyl carrier protein*) se condensa con un sustrato de carbono para formar una cadena creciente de acil-ACP. La producción de ácidos grasos por cianobacterias genomodificadas se ha descrito en la publicación de solicitud de Estados Unidos N.º 2009/0298143, que desvela la introducción de un gen no originario que codifica una acil-ACP tioesterasa grasa para liberar ácidos grasos de la ACP. Sin embargo, como sucede con muchos procesos celulares, la inhibición por retroalimentación mediante productos intermedios y/o diversos productos finales puede limitar la producción de ácidos grasos. Por ejemplo, la inhibición por retroalimentación de enzimas clave de la síntesis de ácidos grasos, tales como ACCasa, FabH y FabI, está mediada por acil-ACP de cadena media a larga. Véase Handke et al., Metabolic Eng. 13:28-37 (2011); Figura 1 (línea discontinua).

Sumario de la invención

La presente invención describe modificaciones genéticas adicionales para mejorar los rendimientos de productos de ácido graso en una cianobacteria recombinante genomodificada para que exprese un gen no originario que codifique una tioesterasa. Dichas modificaciones incluyen la expresión de un gen adicional no originario codificante de la ácido lisofosfatídico aciltransferasa (LPAAT, por las siglas del inglés *lysophosphatidic acid acyltransferase*), que utiliza sustratos de acil-ACP o acil-CoA para acilar el grupo hidroxilo sn-2 del ácido lisofosfatídico para formar ácido fosfatídico, un precursor de los lípidos de membrana. La LPAAT tiene, preferentemente, una preferencia de sustrato de longitud de cadena acilo diferente de la preferencia de sustrato de longitud de cadena acilo de la tioesterasa producida por la expresión de un gen no originario por el microorganismo recombinante. Modificaciones adicionales del microorganismo hospedador pueden incluir opcionalmente la atenuación de un gen de acil-ACP sintetasa o de un gen de acil-CoA sintetasa, que participa en el reciclaje y la degradación, respectivamente, de ácidos grasos, y/o la sobreexpresión de uno o más polipéptidos que tienen actividad lipolítica (por ejemplo, lipasas) que pueden liberar ácidos grasos de los lípidos (véase, por ejemplo, la Figura 1).

En un aspecto, la invención proporciona una cianobacteria recombinante transformada con una secuencia de ácido nucleico no originario que codifica una acil-ACP tioesterasa procedente de una planta superior unida operativamente a un promotor heterólogo y una secuencia de ácido nucleico no originario que codifica una ácido lisofosfatídico aciltransferasa (LPAAT) derivada de una cianobacteria unida operativamente a un promotor heterólogo, en donde la cianobacteria produce al menos un ácido graso libre o al menos un derivado de ácido graso, y en donde la LPAAT tiene una identidad de al menos 85 % con la SEQ ID NO: 2, y la tioesterasa tiene una identidad de al menos 85 % con la SEQ ID NO: 44. Un producto de ácido graso puede ser, por ejemplo, un ácido grado libre, un aldehído graso, un alcohol graso, un éster de ácido graso, un éster de cera, un alcano o un alqueno.

La cianobacteria recombinante, que incluye un gen de tioesterasa no originario, incluye además una molécula de ácido nucleico no originario que incluye una secuencia que codifica una LPAAT, en donde la secuencia que codifica la LPAAT puede ser homóloga (originaria de la misma especie) o heteróloga (originaria de una especie diferente) con respecto a la cianobacteria hospedadora. Por ejemplo, un gen de LPAAT no originario de la misma especie o de una especie diferente, unido operativamente a un promotor heterólogo, puede introducirse en la cianobacteria hospedadora, o de manera alternativa, se puede introducir un gen de LPAAT homólogo o heterólogo en la cianobacteria hospedadora de tal manera que el gen de LPAAT se recombine en el genoma hospedador y se una operativamente a un promotor que sea endógeno al hospedador. Además, de manera alternativa, una cianobacteria recombinante, que incluye un gen de LPAAT no originario, puede comprender un gen de LPAAT endógeno que reside en el genoma de la cianobacteria hospedadora unido operativamente a un promotor heterólogo introducido en la cianobacteria e integrado en el genoma, por ejemplo, por recombinación homóloga. Una molécula de ácido nucleico no originario que incluye un gen de LPAAT en una cianobacteria proporcionada en el presente documento procede de una cianobacteria.

La cianobacteria recombinante comprende al menos un gen de LPAAT no originario y al menos un gen de tioesterasa no originario que puede escindir un sustrato de acil tioéster, en el que la LPAAT y la tioesterasa tienen, preferentemente, diferentes especificidades de sustrato de longitud de cadena de acilo. Por ejemplo, la cianobacteria recombinante puede incluir una molécula de ácido nucleico no originario que codifique una LPAAT que tenga una preferencia de sustrato por uno o más sustratos de acilo de cadena larga (por ejemplo, acil-ACP o acil-CoA C16 y/o C18) y un gen no originario que codifique un tioesterasa que tenga una preferencia de sustrato por uno o más sustratos de acilo de cadena media (por ejemplo, sustratos de acil-ACP de C8, C10, C12 y/o C14). De manera alternativa, la cianobacteria recombinante puede incluir una molécula de ácido nucleico no originario que codifique una LPAAT que tenga una preferencia de sustrato por uno o más sustratos de acilo de cadena media (por ejemplo, sustratos de acil-ACP de C8, C10, C12 y/o C14) y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por uno o más sustratos de acilo de cadena larga (por ejemplo, acil-ACP o acil-CoA de C16 y/o C18). En otro ejemplo, la cianobacteria recombinante puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo C16 (por ejemplo, acil-ACP de C16 y/o acil-CoA de C16) y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por uno o más de un sustrato de acilo de C8, C10, C12 o C14 (por ejemplo, acil-ACP de C8, C10, C12 y/o C14). De manera alternativa, la cianobacteria recombinante puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por sustratos de acilo de C18 (por ejemplo, acil-ACP de C18 y/o acil-CoA de C18) y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por uno o más de un sustrato de acilo de C8, C10, C12, C14 o C16 (por ejemplo, acil-ACP de C8, C10, C12, C14 y/o C16). En otros ejemplos no limitantes, una cianobacteria recombinante puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C18 y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por un sustrato de acilo de C16; una cianobacteria recombinante puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C18 y un gen no originario que codifique un gen de tioesterasa que tenga una preferencia de sustrato por sustratos de acilo de C14 y C16; o una cianobacteria recombinante puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C18 y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por sustratos de acilo de C12, C14 y C16. En ejemplos adicionales, una cianobacteria puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C16 y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por un sustrato de acilo de C14; una cianobacteria puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C16 y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por un sustrato de acilo de C12, y una cianobacteria puede incluir un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato por un sustrato de acilo de C16 y un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato por sustratos de acilo de C12 y C14.

Una cianobacteria recombinante incluye una molécula de ácido nucleico no originario que es al menos 85 %, por ejemplo, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o aproximadamente 100 % idéntica a la SEQ ID NO: 2.

El promotor heterólogo puede ser un promotor regulable, y opcionalmente puede ser un promotor inducible. Por ejemplo, el promotor puede inducirse con luz, temperatura, o con adición o eliminación de un compuesto a los medios de crecimiento, tal como, por ejemplo, un azúcar o análogo de azúcar, una hormona, una sal, un metal, *etc.* En ejemplos particulares, una secuencia de ácido nucleico que codifica un gen de tioesterasa que expresa una cianobacteria recombinante, está unida operativamente a un promotor heterólogo que puede ser, por ejemplo, regulable, y opcionalmente puede ser inducible. Además, una secuencia de ácido nucleico que codifica LPAAT y una secuencia de ácido nucleico que codifica tioesterasa pueden unirse operativamente a diferentes copias del mismo promotor, o a diferentes promotores regulados por la misma o diferentes condiciones o el mismo o diferentes compuestos. De manera alternativa, una secuencia de ácido nucleico que codifica LPAAT y una secuencia de ácido nucleico que codifica tioesterasa pueden unirse operativamente al mismo promotor, por ejemplo, la secuencia de ácido nucleico que codifica LPAAT y la secuencia que codifica tioesterasa se pueden organizar como un operón.

Además, una cianobacteria recombinante puede comprender opcionalmente además un gen de acil-ACP sintetasa o un gen de acil-CoA sintetasa atenuado o alterado. Por ejemplo, el gen endógeno de acil-ACP sintetasa o el gen de acil-CoA sintetasa puede ser una diana del ARN antisentido, de uno o más micro ARN, de un ARNip (p. ej., generado por una construcción de ARNhc) o de una o más ribozimas, para reducir la expresión del gen de la acil-ACP sintetasa o del gen de la acil-CoA sintetasa. De manera alternativa, el gen endógeno de acil-ACP sintetasa o el gen de acil-CoA sintetasa puede interrumpirse, truncarse o delecionarse, por ejemplo, por mutagénesis insercional, modificación del genoma de meganucleasas y/o recombinación homóloga.

La cianobacteria recombinante que incluye un gen de LPAAT no originario y un gen de tioesterasa no originario, puede incluir opcionalmente además un gen no originario que codifique un polipéptido que tenga actividad lipolítica, que puede ser, por ejemplo, una lipasa, una esterasa, una cutinasa o una amidasa. Adicionalmente, o de manera alternativa, la cianobacteria recombinante puede tener un gen atenuado que codifique una acil-ACP sintetasa, una acil-CoA sintetasa o una acil-CoA oxidasa.

Además adicionalmente o de manera alternativa, una cianobacteria recombinante que incluye un gen de LPAAT no originario y un gen de tioesterasa no originario puede incluir adicionalmente opcionalmente además uno o más genes no originarios que codifiquen enzimas para producir derivados de ácidos grasos tales como aldehídos grasos, alcoholes grasos, ésteres de ácidos grasos, ésteres de cera, alcanos y alquenos. Son ejemplos no limitativos de dichas enzimas las acil-ACP reductasas formadoras de aldehído, las acil-CoA reductasas formadoras de aldehído, las ácido carboxílico reductasas, las ácido graso reductasas formadoras de aldehído, las acil-ACP reductasas formadoras de alcohol, las acil-CoA reductasas formadoras de alcohol, las aciltransferasas, las sintasas de cera, las aldehído graso descarbonilasas y ácido graso descarboxilasas.

La cianobacteria recombinante que incluye un gen de LPAAT no originario y un gen de tioesterasa no originario, puede producir al menos 10 % más de un producto de ácido graso que una cianobacteria idéntica en todos los aspectos, excepto que carece de una molécula de ácido nucleico no originario que comprende una secuencia de ácido nucleico que codifica una LPAAT. Por ejemplo, una cianobacteria recombinante puede producir al menos 20 %, al menos 30 %, al menos 40 %, al menos 50 %, al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 100 %, al menos 150 %, al menos 200 %, al menos 400 %, al menos 600 %, al menos 800 % o al menos 1 000 % más que la cantidad del producto de ácido graso producido por una cianobacteria de otra manera idéntica que no incluye la molécula de ácido nucleico no originario que comprende una secuencia de ácido nucleico que codifica una LPAAT cultivada en las mismas condiciones.

La cianobacteria recombinante puede ser una cianobacteria, por ejemplo, de los siguientes géneros *Agmenellum, Anabaena, Anabaenopsis, Anacystis, Aphanizomenon, Arthrospira, Asterocapsa, Borzia, Calothrix, Chamaesiphon, Chlorogloeopsis, Chroococcidiopsis, Chroococcus, Crinalium, Cyanobacterium, Cyanobium, Cyanocystis, Cyanospira, Cyanothece, Cylindrospermopsis, Cylindrospermum, Dactylococcopsis, Dermocarpella, Fischerella, Fremyella, Geitleria, Geitlerinema, Gloeobacter, Gloeocapsa, Gloeothece, Halospirulina, Iyengariella, Leptolyngbya, Limnothrix, Lyngbya, Microcoleus, Microcystis, Myxosarcina, Nodularia, Nostoc, Nostochopsis, Oscillatoria, Phormidium, Planktothrix, Pleurocapsa, Prochlorococcus, Prochloron, Prochlorothrix, Pseudanabaena, Rivularia, Schizothrix, Scytonema, Spirulina, Stanieria, Starria, Stigonema, Symploca, Synechococcus, Synechocystis, Thermosynechococcus, Tolypothrix, Trichodesmium, Tychonema* o *Xenococcus.*

En el presente documento también se proporcionan métodos para producir un producto de ácido graso, en donde el método incluye cultivar cualquiera de las cianobacterias recombinantes de la invención en condiciones en las que se produzca al menos un producto de ácido graso. El al menos un producto de ácido graso es uno o más de un ácido graso libre, un aldehído graso, un alcohol graso, un éster de ácido graso, un éster de cera, o un hidrocarburo, y en donde el producto de ácido graso se secreta en el medio. En algunos ejemplos, la cianobacteria recombinante produce al menos un ácido graso libre de C_8 a C_{24}, tal como al menos un ácido graso libre de C_{12} a C_{18}, tal como, por ejemplo, ácido graso libre de C_{12} a C_{16}. De manera alternativa o además, el ácido graso o derivado de ácido graso producido usando los métodos proporcionados en el presente documento, comprende al menos un alcohol graso o aldehído graso de C_8 a C_{24}, tal como al menos un alcohol graso o aldehído graso de C_{12} a C_{18}, tal como un alcohol graso o aldehído graso de C_{12} a C_{16}; o al menos un éster de cera que tiene una cadena A de C_8 a C_{24} o de C_{12} a C_{18} o, por ejemplo, una cadena A de C_{12} a C_{16} y una cadena B de C_8 a C_{24} o de C_{12} a C_{18} o, por ejemplo, una cadena B de C_{12} a $C_{16;}$ o al menos un alcano o alqueno de C_{11} a C_{23}, tal como un alcano o alqueno de C_{11} a C_{15}. El método puede incluir además aislar un producto de ácido graso de la cianobacteria, del medio de cultivo, o de una combinación de los mismos. Además, la cianobacteria recombinante puede secretar al menos una parte del producto de ácido graso en el medio de cultivo, y opcionalmente toda o una parte del producto secretado puede aislarse del medio de cultivo.

En diversos métodos, la cantidad de producto de ácido graso que produce la cianobacteria recombinante puede ser al menos 10 % mayor, al menos 20 % mayor, al menos 30 % mayor, al menos 40 % mayor, al menos 50 % mayor, al menos 60 % mayor, al menos 70 % mayor, al menos 80 % mayor, al menos 90 % mayor o al menos 100 % mayor, que la cantidad que produce una cianobacteria de otra manera idéntica que carece de la secuencia de ácido nucleico no originario que codifica una LPAAT que se cultiva en las mismas condiciones.

Además se desvela un producto de ácido graso producido utilizando los métodos de la invención. El producto de ácido graso puede ser, por ejemplo, un ácido graso, un alcohol graso, un aldehído graso, un alcano, un alqueno, un éster de ácido graso o un éster de cera. El producto de ácido graso puede tener al menos una cadena de carbono que tenga una longitud de C_7 a C_{24} y en algunas realizaciones preferidas, el producto de ácido graso incluye uno o más ácidos grasos libres, aldehídos grasos, alcoholes grasos, alcanos o alquenos que tienen longitudes de cadena de C_{12} a C_{18}, por ejemplo, longitudes de cadena de C_{12} a C16, o uno o más ésteres de ácidos grasos que tienen una o tanto una cadena A como una cadena B de C_8 a C_{24} o de C_{12} a C_{18}, por ejemplo, de C_{12} a C_{16}. Por ejemplo, un éster de cera producido usando los métodos proporcionados en el presente documento puede tener una cadena A y B de C_8 a C_{24} o de C_{12} a C_{18}, o en algunos ejemplos, de C_{12} a C_{16}.

Breve descripción de los dibujos

La Figura 1 es una diagrama que muestra modificaciones genéticas propuestas para aliviar la inhibición por retroalimentación de la biosíntesis de ácidos grasos al reducir el conjunto de ACP-C18. Las abreviaturas son FFA (*free fatty acid*): ácido grasos libre, AAS: acil-ACP sintasa, Cc1FatB1: acil-ACP tioesterasa de *Cuphea carthagenensis* con una preferencia de sustrato de C12*,* C14 y C16 y LPAAT Sll1752 C18:0 de *Synechocystis* sp. 6803 con una preferencia de sustrato de acilo de C18.

La Figura 2 es una representación esquemática de un ejemplo de un vector de integración que incluye una secuencia de ácido nucleico que codifica una LPAAT (sll1752).

La Figura 3 es una representación esquemática de un ejemplo de un vector de integración que incluye una secuencia de ácido nucleico que codifica una acil-ACP tioesterasa (Cc1FatB1).

La Figura 4 es una representación esquemática de un ejemplo de un vector para alterar un gen que codifica una acil-ACP sintetasa (slr1609).

La Figura 5 es un gráfico que muestra la producción total no normalizada de ácidos grasos libres (µg/ml) de cepas de *Synechocystis* sp. 6803 que comprenden (de izquierda a derecha) el gen de acil-ACP tioesterasa de Cc1FatB1 (Control 1A), el gen de sll1752 recombinante que codifica una LPAAT C18:0 (sll1752), un gen de acil-ACP sintetasa atenuado (AASKO), seguido de tres clones distintos (Parche 3, Parche 5, Parche 6) que contienen el gen de acil-ACP tioesterasa de Cc1FatB1 en combinación con el gen de sll1752 inducible (1A/sll1752), y finalmente el gen de LPAAT sll1752 en combinación con el gen de acil-ACP sintetasa atenuado (AASKO).

La Figura 6 es una tabla que proporciona las cantidades (µg/ml) de ácidos grasos libres de longitudes de cadena particulares y el grado de saturación producido por cepas de *Synechocystis* sp. PCC 6803 que comprenden el gen de acil-ACP tioesterasa de Cc1FatB1 (Control 1A), el gen de sll1752 recombinante que codifica una LPAAT C18: 0 (sll1752), un gen de acil-ACP sintetasa atenuado (AASKO), tres clones distintos (Clon 3, Clon 5, Clon 6) que contienen el gen de acil-ACP tioesterasa de Cc1FatB1, así como el gen de sll1752 recombinante (1A/sll1752) y el gen de sll1752 recombinante en combinación con el gen de acil-ACP sintetasa atenuado (sll1752/AASKO).

Descripción detallada de la invención

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el que entiende comúnmente un experto en la materia a la que pertenece la presente invención. En caso de conflicto, regirá la presente solicitud, incluidas las definiciones. A no ser que el contexto requiera otra cosa, los términos en singular incluirán las pluralidades y los términos en plural incluirán el singular.

Aunque en la práctica o el ensayo de la presente invención, pueden usarse métodos y materiales similares o equivalentes a los descritos en el presente documento, a continuación se describen métodos y materiales adecuados. Los materiales, métodos y ejemplos son solamente ilustrativos y no se pretende que sean limitantes. Otras características y ventajas de la invención serán obvias a partir de la siguiente descripción detallada y de las reivindicaciones.

Para facilitar la comprensión de la presente invención, a continuación se define una serie de términos y expresiones.

Como se usa en la presente divulgación y en las reivindicaciones, las formas en singular "un", "uno", "una", "el" y "la", incluyen formas en plural a menos que el contexto dicte claramente otra cosa.

Dondequiera que en el presente documento se describan realizaciones con la expresión "que comprende" también se proporcionan realizaciones análogas descritas de otra manera en términos de "que consiste en" y/o "que consiste esencialmente en".

La expresión "y/o" como se utiliza en una frase tal como "A y/o B" del presente documento pretende incluir "A y B", "A o B", "A" y "B".

El término "gen" se usa de manera generalizada para referirse a cualquier segmento de molécula de ácido nucleico (normalmente ADN, aunque opcionalmente ARN) que codifica una proteína o ARN expresado. Por tanto, los genes

incluyen secuencias que codifican ARN expresado (que puede incluir secuencias codificantes de polipéptidos). Los genes pueden comprender además las secuencias reguladoras necesarias para su expresión. Los genes pueden obtenerse de una variedad de fuentes, incluida la clonación de una fuente de interés o la síntesis a partir de información de secuencia conocida o prevista, y puede incluir secuencias diseñadas para que tengan parámetros deseados.

La expresión "ácido nucleico" o "molécula de ácido nucleico" se refiere a, por ejemplo, ADN o ARN (por ejemplo, ARNm). Las moléculas de ácido nucleico pueden ser bicatenarias o monocatenarias, el ARN o ADN monocatenario puede ser la cadena codificante (sentido) o la cadena no codificante (antisentido).

Una molécula de ácido nucleico puede "proceder de" una fuente indicada, que incluye el aislamiento (total o parcial) de un segmento de ácido nucleico de una fuente indicada o la purificación de un polipéptido de una fuente indicada. Una molécula de ácido nucleico también puede proceder de una fuente indicada, por ejemplo, mediante clonación directa, amplificación por PCR o síntesis artificial a partir de la fuente de polinucleótidos indicada o basada en una secuencia asociada a la fuente de polinucleótidos indicada. Los genes o las moléculas de ácido nucleico procedentes de una fuente o especie particular también incluyen genes o moléculas de ácido nucleico que tienen modificaciones de secuencia con respecto a las moléculas de ácido nucleico fuente. Por ejemplo, un gen o una molécula de ácido nucleico procedente de una fuente (por ejemplo, un gen de referencia particular) puede generar una o más mutaciones con respecto al gen o a la molécula de ácido nucleico fuente que no son intencionadas o que se introducen deliberadamente, y si una o más mutaciones, incluyendo sustituciones, deleciones o inserciones, se introducen deliberadamente, las alteraciones de secuencia se pueden introducir mediante mutación aleatoria o dirigida de células o ácidos nucleicos, por amplificación u otras técnicas de biología molecular, o por síntesis química. Un gen o una molécula de ácido nucleico que procede de un gen o de una molécula de ácido nucleico de referencia que codifica un ARN o polipéptido funcional puede codificar un ARN o polipéptido funcional que tenga una identidad de secuencia de al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98% o al menos 99 % con el ARN o polipéptido funcional de referencia o fuente, o con un fragmento funcional del mismo. Por ejemplo, un gen o una molécula de ácido nucleico que procede de un gen o de una molécula de ácido nucleico de referencia que codifica un ARN o polipéptido funcional puede codificar un ARN o polipéptido funcional que tenga una identidad de secuencia de al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % con el ARN o polipéptido funcional de referencia o fuente, o con un fragmento funcional del mismo.

Como se usa en el presente documento, un ácido nucleico "aislado" o una proteína "aislada" se elimina de su medio natural o del contexto en el que el ácido nucleico o la proteína existe en la naturaleza. Por ejemplo, una proteína o molécula de ácido nucleico aislada se elimina de la célula u organismo con el que está asociado en su entorno originario o natural. Un ácido nucleico aislado o una proteína aislada puede estar, en algunos casos, parcial o sustancialmente purificado(a), pero no se requiere un nivel particular de purificación para el aislamiento. Por tanto, por ejemplo, una molécula de ácido nucleico aislada puede ser una secuencia de ácido nucleico que se ha escindido del cromosoma, genoma o episoma en el que está integrado en la naturaleza.

Una molécula de ácido nucleico o secuencia de nucleótidos o proteína o secuencia polipeptídica "purificada", carece sustancialmente de material celular y de componentes celulares. La molécula de ácido nucleico o proteína purificada puede carecer de productos químicos además de tampón o disolvente, por ejemplo. "Carece sustancialmente" no pretende significar que otros componentes, además de las nuevas moléculas de ácido nucleico, sean indetectables. En algunas circunstancias, "carece sustancialmente" puede significar que la molécula de ácido nucleico o la secuencia de nucleótidos está carece de al menos 95 % (p/p) de material y componentes celulares.

Las expresiones "origen natural" y "tipo silvestre" se refieren a una forma que se encuentra en la naturaleza. Por ejemplo, una molécula de ácido nucleico, una secuencia de nucleótidos o una proteína de origen natural o de tipo silvestre, pueden estar presentes y aisladas de una fuente natural, y no se modifican intencionalmente mediante manipulación humana.

Como se usa en el presente documento, "atenuado" significa reducido en cantidad, grado, intensidad o fuerza. La expresión génica atenuada puede referirse a una cantidad y/o tasa de transcripción significativamente reducida del gen en cuestión, o de la traducción, plegamiento o ensamblaje de la proteína codificada. Como ejemplos no limitantes, un gen atenuado puede ser un gen mutado o alterado (por ejemplo, un gen alterado por deleción parcial o total, o por mutación insercional) o que tiene una expresión disminuida debido a la alteración de secuencias reguladoras de genes.

La expresión "molécula de ácido nucleico exógeno" o "gen exógeno" se refiere a una molécula de ácido nucleico o a un gen que se ha introducido ("transformado") en una célula. Una célula transformada puede referirse a una célula recombinante, en la pueden introducirse uno o más genes exógenos adicionales. Un descendiente de una célula transformada con una molécula de ácido nucleico también puede referirse a "transformado" si ha heredado la molécula de ácido nucleico exógeno. El gen exógeno puede ser de una especie diferente (y por tanto "heterólogo") o puede ser de la misma especie (y por tanto "homólogo"), en relación con la célula que se transforma. Una molécula de ácido nucleico, un gen o una proteína "endógeno(a)", es una molécula de ácido nucleico, un gen o una proteína

originario(a) tal como aparece en el hospedador o como lo produce de manera natural dicho hospedador.

En el presente documento, el término "originario(a)" se usa para referirse a secuencias de ácido nucleico o de aminoácidos tal como se producen de manera natural en el hospedador. En el presente documento el término "no originario" se usa para referirse a secuencias de ácido nucleico o de aminoácidos que no se producen de manera natural en el hospedador. Una secuencia de ácido nucleico o de aminoácidos que se ha eliminado de una célula hospedadora, sometido a manipulación en el laboratorio e introducido o reintroducido en una célula hospedadora, se considera "no originaria". Los genes sintéticos o parcialmente sintéticos introducidos en una célula hospedadora son "no originarios". Los genes no originarios incluyen además genes endógenos al microorganismo hospedador unidos operativamente a una o más secuencias reguladoras heterólogas que se han recombinado en el genoma del hospedador.

Una molécula de ácido nucleico "recombinante" o "genomodificada" es una molécula de ácido nucleico que se ha alterado a través de manipulación humana. Como ejemplos no limitantes, una molécula de ácido nucleico recombinante que: 1) se ha sintetizado o modificado *in vitro,* por ejemplo, utilizando técnicas químicas o enzimáticas (por ejemplo, mediante el uso de síntesis química de ácidos nucleicos, o mediante el uso de enzimas para la replicación, polimerización, digestión (exonucleolítica o endonucleolítica), ligamiento, transcripción inversa, transcripción, modificación de bases (incluyendo, por ejemplo, metilación), integración o recombinación (incluyendo recombinación homóloga y específica de sitio)) de moléculas de ácido nucleico; 2) incluye secuencias de nucleótidos unidas que no están unidas en la naturaleza, 3) se ha genomodificado usando técnicas de clonación molecular de modo que carece de uno o más nucleótidos con respecto a la secuencia de la molécula de ácido nucleico natural, y/o 4) se ha manipulado usando técnicas de clonación molecular de modo que tenga uno o más cambios o reordenamientos de secuencia con respecto a la secuencia de ácido nucleico de origen natural. Como ejemplos no limitantes, un ADNc es una molécula de ADN recombinante, como lo es cualquier molécula de ácido nucleico que se haya generado *in vitro* mediante una o más reacciones de la polimerasa, o cuyos enlazadores se hayan unido, o que se haya integrado en un vector, tal como un vector de clonación o un vector de expresión.

La expresión "proteína recombinante", como se usa en el presente documento, se refiere a una proteína producida mediante genotecnología.

Cuando se aplica a organismos, los términos recombinante, genomodificado(a) o modificado(a) mediante genotecnología, se refiere a organismos que se han manipulado mediante la introducción de una secuencia de ácido nucleico heterólogo o recombinante en el organismo, e incluye genes desactivados, mutaciones dirigidas y reemplazo de genes, reemplazo, deleción o inserción de promotores, así como la introducción de transgenes en el organismo. Los organismos recombinantes o modificados mediante genotecnología, también pueden ser organismos en los que se han introducido construcciones para "atenuar" genes. Dichas construcciones incluyen, pero sin limitación, construcciones de iARN, microARN, ARNhc, antisentido y ribozimática. También se incluyen organismos cuyos genomas se han alterado por la actividad de meganucleasas o nucleasas de dedos de zinc. La molécula de ácido nucleico heterólogo o recombinante puede integrarse en el genoma del organismo recombinante/modificado mediante genotecnología o, en otros casos, no integrarse en el genoma del organismo recombinante/modificado mediante genotecnología. Como se usa en el presente documento, "microorganismo recombinante" o "célula hospedadora recombinante" incluye descendencia o derivados de los microorganismos recombinantes de la invención. Dado que pueden producirse determinadas mutaciones en generaciones sucesivas debido a mutaciones o a influencias ambientales, dicha descendencia o derivados pueden, de hecho, no ser idénticos a la célula original, pero siguen estando incluidos en el alcance de la expresión tal como se usa en el presente documento.

El término "promotor" se refiere a una secuencia de ácido nucleico capaz de unirse a la ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificante cadena abajo (dirección 3'). Un promotor incluye el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Un promotor puede incluir un sitio de inicio de la transcripción, así como dominios de unión a proteínas (secuencias consenso) responsables de la unión de la ARN polimerasa. Los promotores eucariotas a menudo, pero no siempre, contienen cajas "TATA" y cajas "CAT". Los promotores procariotas pueden contener -10 y -35 secuencias consenso promotoras procariotas. En la técnica se conoce bien una gran cantidad de promotores, incluidos los promotores constitutivos, inducibles y represibles, procedentes de una variedad de fuentes diferentes. Entre las fuentes representativas se incluyen, por ejemplo, tipos de virus, células de mamíferos, de insectos, de plantas, de levaduras y de bacterias, y se dispone fácilmente de promotores adecuados de estas fuentes, o se pueden fabricar sintéticamente, basándose en secuencias en línea disponibles públicamente o, por ejemplo, de depósitos como la ATCC (A*merican Type Culture Collection*, colección americana de cultivos tipo) y otras fuentes comerciales o individuales. Los promotores pueden ser unidireccionales (es decir*,* iniciar la transcripción en una dirección) o bidireccionales (es decir*,* inician la transcripción en ambas direcciones fuera de las cadenas opuestas). Un promotor puede ser un promotor constitutivo, un promotor reprimible, o un promotor inducible. Como ejemplos no limitantes de promotores se incluyen, por ejemplo, el promotor de T7, el promotor de citomegalovirus (CMV), el promotor del SV40 (virus de simio 40) y el promotor del RSV (virus del sarcoma de Rous). Como ejemplos de promotores inducibles se incluyen el promotor de lac, el promotor de pBAD (araA), el promotor de Tet (Patentes de Estados Unidos N.º 5 464 758 y 5 814 618) y el promotor de ecdisona (No et al., Proc. Natl. Acad. Sci. (1996) 93 (8): 3346-3351).

El término "heterólogo", cuando se usa en referencia a un polinucleótido, un gen, un ácido nucleico, un polipéptido o una enzima, se refiere a un polinucleótido, un gen, un ácido nucleico, un polipéptido o una enzima, que no se encuentra de manera natural en el organismo hospedador. Cuando se hace referencia a una secuencia reguladora de genes o a una secuencia auxiliar de ácido nucleico utilizada para mantener o manipular una secuencia de genes (por ejemplo, una región no traducida 5', una región no traducida 3', una secuencia de adición de poli A, una secuencia intrónica, sitios de corte y empalme, sitio de unión a ribosomas, una secuencia de sitio interno al ribosoma, una región de homología del genoma, sitio de recombinación, etc.), "heterólogo" significa que la secuencia reguladora o secuencia auxiliar es de una fuente diferente a la del gen con el que la secuencia reguladora o auxiliar de ácido nucleico se yuxtapone en una construcción, genoma, cromosoma o episoma. Por tanto, en el presente documento, un promotor unido operativamente a un gen al cual no está unido operativamente en su estado natural (es decir, en el genoma de un organismo no modificado mediante genotecnología), se refiere a un "promotor heterólogo". a pesar de que el promotor pueda proceder de la misma especie (o, en algunos casos, del mismo organismo) que la del gen con el que está unido.

Como se usa en el presente documento, el término "proteína" o "polipéptido" pretende abarcar un "polipéptido" en singular así como "polipéptidos" en plural, y se refiere a una molécula compuesta de monómeros (aminoácidos) unidos linealmente por enlaces amida (también conocidos como enlaces peptídicos). El término "polipéptido" se refiere a cualquier cadena o cadenas de dos o más aminoácidos, y no a una longitud específica del producto. Por tanto, péptidos, dipéptidos, tripéptidos, oligopéptidos, "proteína", "cadena de aminoácidos", o cualquier otro término utilizado para referirse a una o más cadenas de dos o más aminoácidos, se incluye en la definición de "polipéptido", y el término "polipéptido" puede utilizarse en lugar de cualquiera de estos términos o indistintamente.

Como se usa en el presente documento, las expresiones "porcentaje de identidad" u "homología", con respecto a secuencias de ácido nucleico o de polipéptidos, se definen como el porcentaje de nucleótidos o de restos de aminoácidos en la secuencia candidata que son idénticos a los polipéptidos conocidos, después de alinear las secuencias para el porcentaje máximo de identidad e introducir huecos, si fuera necesario, para conseguir el máximo porcentaje de homología. Las inserciones o deleciones en los extremos N-terminal o C-terminal no se interpretarán como que afectan a la homología, y las deleciones y/o inserciones internas en la secuencia de polipéptidos de menos de aproximadamente 30, menos de aproximadamente 20 o menos de aproximadamente 10 restos de aminoácidos, no se interpretarán como que afectan a la homología.

La identidad u homología de secuencia a nivel de secuencia de nucleótidos o aminoácidos se puede determinar mediante análisis BLAST (del inglés, *Basic Local Alignment Search Tool*) que usa el algoritmo empleado por los programas blastp, blastn, blastx, tblastn y tblastx (Altschul (1997), Nucleic Acids Res. 25, 3389-3402 y Karlin (1990), Proc. Natl. Acad. Sci. USA 87, 2264-2268), que están diseñados para la búsqueda de similitud de secuencia. La estrategia utilizada por el programa BLAST es considerar primero segmentos similares, con y sin huecos, entre una secuencia de consulta y una secuencia de la base de datos, para evaluar después la significación estadística de todas las coincidencias que se identifican y finalmente extraer solo aquellas coincidencias que satisfacen un umbral de significación preseleccionado. Para un tratamiento de aspectos básicos en búsquedas de similitud de secuencia en bases de datos, véase Altschul (1994), Nature Genetics 6, 119-129. Los parámetros de búsqueda para histogramas, descripciones, alineamientos, previsiones (es decir, el umbral de significación estadística para informar sobre coincidencias frente a secuencias de bases de datos), límite, matriz y filtro (baja complejidad), pueden estar en la configuración predeterminada. La matriz de puntuación por defecto utilizada por blastp, blastx, tblastn y tblastx, es la matriz BLOSUM62 (Henikoff (1992), Proc. Natl. Acad. Sci. USA 89, 10915-10919), recomendada para secuencias de consulta con una longitud mayor de 85 (bases de nucleótidos o de aminoácidos).

Para blastn, diseñado para comparar secuencias de nucleótidos, la matriz de puntuación se establece con las proporciones de M (es decir, la puntuación por recompensa de un par de restos coincidentes) con respecto a N (es decir, la puntuación por penalización de restos no coincidentes), en donde los valores predeterminados de M y N pueden ser +5 y -4, respectivamente. Los cuatro parámetros de blastn pueden ajustarse de la siguiente manera: Q=10 (penalización por creación de huecos); R=10 (penalización por extensión de huecos); wink=1 (genera aciertos de palabras en cada posición winkth a lo largo de la consulta); y gapw=16 (establece el ancho de la ventana dentro de la cual se generan las alineaciones con huecos). La configuración de los parámetros de Blastp equivalente para la comparación de secuencias de aminoácidos puede ser: Q=9; R=2; wink=1; y gapw=32. Una comparación entre secuencias de Bestfit, disponible en el paquete GCG versión 10.0, puede usar parámetros de ADN GAP=50 (penalización por creación de huecos) y LEN=3 (penalización por extensión de huecos), y las configuraciones equivalentes en las comparaciones de proteínas pueden ser GAP=8 y LEN=2.

Por tanto, cuando se hace referencia a las secuencias de polipéptidos o de ácido nucleico de la presente invención, se incluyen identidades de secuencia de al menos 65 %, 70 %, 75 %, 80 % u 85 %, por ejemplo, de al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o aproximadamente 100 % de identidad de secuencia con la secuencia del polipéptido o la secuencia de ácido nucleico de longitud completa, o con fragmentos de las mismas, que comprenden una secuencia consecutiva de al menos 50, al menos 75, al menos 100, al menos 125, al menos 150 o más restos de aminoácidos de la proteína completa; variantes de

dichas secuencias, por ejemplo, en donde se ha insertado al menos un resto de aminoácido N y/o C terminal en, y/o dentro de, la(s) secuencia(s) develada(s) que contiene(n) la inserción y la sustitución. Las variantes contempladas pueden incluir, además o como alternativa, aquellas que contienen mutaciones predeterminadas, por ejemplo, mediante recombinación homóloga o mutagénesis dirigida o por PCR, y los polipéptidos o ácidos nucleicos correspondientes de otras especies, incluyendo, pero sin limitación, los descritos en el presente documento, los alelos u otras variantes de origen natural de la familia de polipéptidos o ácidos nucleicos que contienen una inserción y sustitución; y/o derivados en los que el polipéptido se ha modificado de manera covalente por sustitución, medios químicos, enzimáticos u otros medios apropiados con un residuo distinto de un aminoácido de origen natural que contenga la inserción y sustitución (por ejemplo, un residuo detectable tal como una enzima).

Como se usa en el presente documento, la frase "sustitución de aminoácidos conservativa" o "mutación conservativa" se refiere al reemplazo de un aminoácido por otro aminoácido con una propiedad común. Una forma funcional de definir propiedades comunes entre aminoácidos individuales, es analizar las frecuencias normalizadas de los cambios de aminoácidos entre las proteínas correspondientes de organismos homólogos (Schulz (1979) Principles of Protein Structure, Springer-Verlag). Según dichos análisis, los grupos de aminoácidos pueden definirse cuando los aminoácidos de un grupo se intercambian preferentemente entre sí y, por lo tanto, se asemejan más entre sí en su impacto sobre la estructura general de las proteínas (Schulz (1979) Principles of Protein Structure, Springer-Verlag). Como ejemplos de grupos de aminoácidos definidos de esta manera pueden incluirse: un "grupo con carga/polar" que incluye Glu, Asp, Asn, Gln, Lys, Arg e His; un "grupo aromático o cíclico" que incluye Pro, Phe, Tyr y Trp; y un "grupo alifático" que incluye Gly, Ala, Val, Leu, Ile, Met, Ser, Thr y Cys. Dentro de cada grupo, también pueden identificarse subgrupos. Por ejemplo, el grupo de aminoácidos con carga/polar puede subdividirse en subgrupos que incluyen: el "subgrupo con carga positiva" que comprende Lys, Arg e His; el "subgrupo con carga negativa" que comprende Glu y Asp; y el "subgrupo polar" que comprende Asn y Gln. En otro ejemplo, el grupo aromático o cíclico puede subdividirse en subgrupos que incluyen: el "subgrupo con anillo de nitrógeno" que comprende Pro, His y Trp; y el "subgrupo con fenilo" que comprende Phe y Tyr. En otro ejemplo más, el grupo alifático puede subdividirse en subgrupos que incluyen: el "subgrupo alifático no polar grande" que comprende Val, Leu e Ile; el "subgrupo alifático ligeramente polar" que comprende Met, Ser, Thr y Cys; y el "subgrupo de restos pequeños" que comprende Gly y Ala. Como ejemplos de mutaciones conservativas se incluyen sustituciones de aminoácidos de los aminoácidos de los subgrupos anteriores, tales como, pero sin limitación: Lys por Arg o viceversa, de modo que pueda conservarse una carga positiva; Glu por Asp o viceversa, de modo que pueda conservarse una carga negativa; Ser por Thr o viceversa, de modo que pueda conservarse un -OH libre; y Gln por Asn o viceversa, de modo que pueda conservarse un -NH2 libre.

Como se usa en el presente documento, "expresión" incluye la expresión de un gen al menos al nivel de producción de ARN, y un "producto de expresión" incluye el producto resultante, por ejemplo, un polipéptido o ARN funcional (por ejemplo, un ARN ribosómico, un ARNt, un ARN antisentido, un micro ARN, un ARNhc, una ribozima, etc.), de un gen expresado. La frase "expresión aumentada" incluye una alteración en la expresión génica para facilitar la producción aumentada de ARNm y/o la expresión aumentada de polipéptidos. "Producción aumentada" incluye un aumento en la cantidad de expresión de un polipéptido, en el nivel de la actividad enzimática de un polipéptido, o una combinación de ambos, en comparación con la producción o la actividad enzimática originaria del polipéptido.

El término "segregado" incluye el transporte de polipéptidos o productos de ácidos grasos producidos por los microorganismos recombinantes o por los métodos de la invención en el espacio periplasmático o medio extracelular. "Secreción aumentada" incluye secreción en exceso de la cantidad de secreción de origen natural, por ejemplo, que es al menos 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % o 10 % o al menos 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 %, 200 %, 300 %, 400 %, 500 %, 600 %, 700 %, 800 %, 900 %, 1 000 % o más, en comparación con el nivel de secreción de origen natural.

Las realizaciones de la presente invención proporcionan la "inserción", por ejemplo, la adición, integración, incorporación o introducción, la activación o regulación positiva de determinadas moléculas de ácido nucleico o secuencias de polinucleótidos particulares, con o sin secuencias reguladoras adicionales, dentro de los microorganismos o de las células hospedadoras que afectan a la actividad, tal como la expresión de una enzima, de determinadas moléculas de ácido nucleico o de secuencias de polinucleótidos particulares. En determinadas realizaciones, un microorganismo de interés puede genomodificarse mediante recombinación homóloga dirigida para insertar un gen de interés o un promotor particular que afecte a la expresión de un gen o conjunto de genes particular.

Las realizaciones de la presente invención proporcionan microorganismos recombinantes en los que las moléculas de ácido nucleico o las secuencias de polinucleótidos particulares están parcial, sustancial o completamente delecionadas, silenciadas, inactivadas o reguladas negativamente para afectar a la actividad para la que codifican, tal como la expresión de una enzima. Los genes pueden estar parcial, sustancial o completamente delecionados, silenciados, inactivados o regulados negativamente por la inserción de secuencias de ácido nucleico que alteran la función y/o expresión del gen (por ejemplo, transducción de P1 u otros métodos conocidos en la técnica). Los términos "eliminar", "eliminación", y "desactivación", pueden utilizarse indistintamente con los términos "deleción", "deleción parcial", "deleción sustancial", o "deleción completa". En determinadas realizaciones, un microorganismo de interés puede genomodificarse mediante recombinación homóloga dirigida para desactivar un gen de interés

particular. Incluso en otras realizaciones, para silenciar, inactivar o regular negativamente, parcial, sustancial o completamente, un gen de interés particular, se puede usar iARN o ADN antisentido (ADNas).

Puede entenderse que estas inserciones, deleciones u otras modificaciones de determinadas moléculas de ácido nucleico o secuencias de polinucleótidos particulares, abarcan una o más "modificaciones genéticas" o "transformaciones" de tal manera que puede entenderse que las cepas resultantes de los microorganismos o las células hospedadoras estén "modificadas genéticamente" o "transformadas".

Como se usa en el presente documento, "regulado(a) positivamente" o "regulación positiva" incluye un aumento en la expresión de un gen o una molécula de ácido nucleico de interés o de la actividad de una enzima, por ejemplo, un aumento en la expresión génica o la actividad enzimática en comparación con la expresión o actividad en un gen o enzima de otra manera idéntico que no se ha regulado positivamente.

Como se usa en el presente documento, "regulado(a) negativamente" o "regulación negativa" incluye una disminución en la expresión de un gen o una molécula de ácido nucleico de interés o de la actividad de una enzima, por ejemplo, una disminución en la expresión génica o la actividad enzimática en comparación con la expresión o actividad en un gen o enzima de otra manera idéntico que no se ha regulado negativamente.

El término "Pfam" se refiere a una amplia colección de dominios y familias de proteínas mantenida por el Consorcio Pfam y disponible en varios sitios web mundiales patrocinados, incluyendo: pfam.sanger.ac.uk/ (Welcome Trust, Sanger Institute); pfam.sbc.su.se/ (Centro de Bioinformática de Estocolmo); pfam.janelia.org/ (Janelia Farm, Instituto Médico Howard Hughes); pfam.jouy.inra.fr/ (Institut national de la Recherche Agronomique); y pfam.ccbb.re.kr. La última versión de Pfam es Pfam 26.0 (noviembre de 2011) basada en la base de datos de proteínas UniProt versión 15.6, una mezcla de Swiss-Prot versión 57.6 y TrEMBL versión 40.6. Los dominios y familias Pfam se identifican usando alineamientos múltiples de secuencias y modelos ocultos de Markov (HMM, *Hidden Markov Models*). Las asignaciones de la familia o dominio Pfam-A, son asignaciones de alta calidad generadas por una alineación de semillas curada utilizando miembros representativos de una familia de proteínas y perfilando modelos ocultos de Markov basados en la alineación de semillas. (A menos que se especifique lo contrario, las coincidencias de una proteína consultada con un dominio o familia Pfam son coincidencias Pfam-A). Después, todas las secuencias identificadas que pertenecen a la familia se utilizan para generar automáticamente una alineación completa con la familia (Sonnhammer (1998) Nucleic Acids Research 26, 320-322; Bateman (2000) Nucleic Acids Research 26, 263-266; Bateman (2004) Nucleic Acids Research 32, Database Issue, D138-D141; Finn (2006) Nucleic Acids Research Database Issue 34, D247-251; Finn (2010) Nucleic Acids Research Database Issue 38, D211-222). Al acceder a la base de datos Pfam, por ejemplo, utilizando cualquiera de los sitios web citados anteriormente, secuencias de proteínas pueden consultarse frente a los HMM utilizando el programa informático de búsqueda de homología HMMER (por ejemplo, HMMER2, HMMER3, o una versión superior, hmmer.janelia.org/). Las coincidencias significativas que identifican una proteína consultada como perteneciente a una familia Pfam (o que tiene un dominio Pfam particular) son aquellas en las que la puntuación en bits es mayor que o igual al umbral de recopilación del dominio Pfam. También pueden usarse valores de expectativa (valores e) como criterio de inclusión de una proteína consultada en una Pfam o para determinar si una proteína consultada tiene un dominio Pfam particular, donde valores bajos de e (mucho menores que 1,0, por ejemplo, menores que 0,1, o menores que o iguales a 0,01) representan bajas probabilidades de que una coincidencia se deba al azar.

Una enzima, tal como una LPAAT o una tioesterasa, que actúa en sustratos de acil tioéster (por ejemplo, que utiliza sustratos de acil-ACP y/o acil-CoA), a menudo tiene actividad en sustratos de diferentes longitudes de cadena, pero puede tener mayor actividad en uno o más sustratos que en otros. Por ejemplo, "preferencia de sustrato" se refiere al sustrato o a los sustratos en los que una enzima es más activa. Las diferentes acil-ACP tioesterasas pueden tener diferentes grados de especificidad de longitud de cadena, a veces denominada "preferencia" de la enzima para escindir una longitud particular de ácido graso de la ACP, y las tioesterasas suelen ser activas en la escisión de un ácido graso de una longitud de cadena particular, mientras que tienen menor actividad en la escisión de uno o más ácidos grasos de otra longitud de cadena. De manera similar, una LPAAT puede tener una preferencia de sustrato por uno o más sustratos de acilo, tales como sustratos de acil-ACP o acil-CoA de longitudes de cadena de carbono particulares.

Como se usa en el presente documento, un ácido graso o acil-ACP de "longitud de cadena media" es un ácido graso o acil-ACP que tiene una longitud de cadena de 8-14 carbonos.

Como se usa en el presente documento, un ácido graso o acil-ACP de "cadena larga" es un ácido graso o acil-ACP que tiene una longitud de cadena de más de 14 carbonos.

Como se usa en el presente documento, la expresión "producto de ácido graso" incluye ácidos grasos libres; mono-, di- o triglicéridos; aldehídos grasos; alcoholes grasos; ésteres de ácidos grasos (incluidos, pero sin limitación, ésteres de cera); e hidrocarburos (incluidos, pero sin limitación, alcanos y alquenos).

Rutas metabólicas para la producción de ácidos grasos

La ruta de la biosíntesis de ácidos grasos está muy conservada en procariotas y en los cloroplastos de algas eucariotas y plantas superiores. La ruta de la biosíntesis de ácidos grasos comienza desde el metabolito central acetil-CoA. La biosíntesis de ácidos grasos se inicia mediante la conversión de acetil-CoA en malonil-CoA, catalizada por acetil-CoA carboxilasa (ACCasa). Después, la malonil-CoA se convierte en malonil-ACP, catalizada por malonil-CoA-ACP transacilasa (FabD en *E. coli).* Finalmente, la malonil-ACP se convierte en acil-ACP, catalizada por el complejo enzimático de la ácido graso sintasa (FAS). El complejo de la ácido graso sintasa inicia el ciclo de alargamiento condensando primero la malonil-ACP con la acetil-ACP, catalizado por una beta-cetoacil-ACP sintasa III (por ejemplo, FabH de *E. coli).* La β-cetoacil-ACP (3-cetoacil-ACP) formada por la reacción de FabH se reduce a un β-hidroxiacil-ACP (3-hidroxiacil-ACP) por la 3-cetoacil-ACP reductasa (por ejemplo, FabG). Después, la β-hidroxiacil-ACP se activa con una β-hidroxiacil-ACP deshidratasa (por ejemplo, FabA, FabZ) para formar trans-2-enoil-ACP, que a su vez se reduce con la enoil-ACP reductasa (por ejemplo, Fab I, Fab K, FabL) para formar el producto acil-ACP alargado con 2 carbonos. Los ciclos posteriores se inician mediante una condensación catalizada por la beta-cetoacil-ACP sintasa I o II (por ejemplo, FabB o FabF) de malonil-ACP con acil-ACP. Los ciclos de condensación, reducción, deshidratación y reducción, se repiten con cada ciclo, añadiendo dos carbonos de la malonil-ACP, hasta que la cadena de acilo se transfiera a otra molécula (por ejemplo, glicerol 3-fosfato) mediante una transacilasa o se separe de la ACP mediante una tioesterasa, tal como FatA o FatB en cloroplastos, para formar ácidos grasos libres.

A diferencia de los cloroplastos de las plantas, las cianobacterias no producen ácidos grasos libres, y a diferencia de E. *coli* y otras bacterias heterótrofas, las cianobacterias no producen acil-CoA. Después del alargamiento de los ácidos grasos con la cadena de acilo unida de manera covalente a la proteína transportadora de acilo, las acil transferasas de las cianobacterias transfieren la cadena de acilo a un esqueleto de glicerol para producir lípidos de membrana.

Para producir derivados de ácidos grasos tales como alcoholes grasos, aldehídos grasos, ésteres de cera, alcanos o alquenos en microorganismos, normalmente es necesario introducir uno o más genes codificantes de una o más enzimas para convertir los productos intermedios de acil-tioéster (por ejemplo, acil-CoA o acil-ACP) en el producto final deseado (por ejemplo, un alcohol, aldehído, alcano, alqueno o éster de cera). Por ejemplo, si los aldehídos grasos y/o los alcanos son el producto final deseado, para reducir la acil-CoA a aldehídos grasos; puede introducirse un gen que codifique una aldehído reductasa grasa formadora de aldehído (por ejemplo, acil-CoA reductasa formadora de aldehído, 1.2.1.42 o 1.2.1.50; véase también la patente de Estados Unidos N.º 6.143.538); adicionalmente o de manera alternativa, puede introducirse un gen de ácido carboxílico reductasa (véanse, por ejemplo, los documentos WO 2010/135624 y WO 2010/042664) para reducir los ácidos grasos libres a aldehídos grasos. Además, puede introducirse un gen que codifique una alcohol graso oxidasa (por ejemplo, 1.1.3.20) o una alcohol graso deshidrogenasa (por ejemplo, 1.1.1.164) para convertir los alcoholes grasos a aldehídos grasos. Los aldehídos grasos pueden procesarse adicionalmente a alcanos con la introducción de un gen codificante de una aldehído graso descarbonilasa (por ejemplo, 4.1.99.5). Si los alcoholes grasos, los alquenos y/o los ésteres de cera son el producto final deseado, puede introducirse un gen que codifique una acil grasa reductasa formadora de alcohol (por ejemplo, acil-CoA reductasa formadora de alcohol, 1.2.1.50) en la célula hospedadora. Además, puede introducirse un gen de aldehído graso reductasa para reducir los aldehídos grasos a alcoholes grasos. Los alcoholes grasos pueden procesarse adicionalmente a alquenos con la introducción de uno o más genes codificantes de una alcohol graso deshidratasa. Los ésteres de ácidos grasos, incluidos los ésteres de cera, pueden formarse introduciendo genes codificantes de polipéptidos que catalicen la condensación de un alcohol con un acil graso tioéster, tal como aciltransferasas y cera sintasas.

En algunas realizaciones de la invención descritas anteriormente, la conversión de acil-ACP en alcohol graso puede producirse mediante la síntesis de un aldehído graso, en donde una acil reductasa (por ejemplo, una acil-CoA reductasa formadora de aldehído o una acil-ACP reductasa formadora de aldehído) expresada en la célula hospedadora reduce primero la acil-ACP a un aldehído graso. Por ejemplo, en determinadas realizaciones, la célula hospedadora se puede genomodificar para que sobreexprese una reductasa endógena formadora de aldehído graso (por ejemplo, insertando elementos de control transcripcional promotor y/o potenciador cerca del gen de la reductasa formadora de aldehído graso). En otras realizaciones, la célula hospedadora puede genomodificarse para que exprese una reductasa exógena formadora de aldehído graso.

Ácido Lisofosfatídico Aciltransferasas

Como se usa en el presente documento, la expresión "ácido lisofosfatídico aciltransferasa" o "LPAAT", o "1-acil-sn-glicerol-3-fosfato aciltransferasa" o "AGPAT" (codificada por *plsC* en *E. coli* y otros procariotas), incluye aquellas enzimas capaces de convertir un ácido lisofosfatídico en un ácido fosfatídico transfiriendo una cadena de acilo desde un sustrato de acilo tal como acil-CoA o acil-ACP a la posición sn2 del ácido lisofosfatídico. La LPAAT incluye aquellas enzimas que corresponden al número de la Comisión de Enzimas 2.3.1.51.

La LPAAT y la enzima glicerol-3-fosfato aciltransferasa (glicerolfosfato aciltransferasa; GPAT; codificada por *plsB* en *E. coli)* que cataliza la conversión de glicerol-3-fosfato en ácido lisofosfatídico, comparte firmas canónicas de aciltransferasa tales como HXXXXD o NHXXXXD y XXXXXXG, donde X es un aminoácido arbitrario, separado por 60 a 83 restos de aminoácidos (Okazaki, K., et al. (2006) Plant Physiol. 141:546-56). En muchos organismos se han

identificado ácidos nucleicos que codifican las GPAT y las LPAAT que tienen estas secuencias de firma, incluso en *Arabidopsis thaliana* (Zheng et al. (2003) Plant Cell 15:1872-87; Kim y Huwang (2004) Plant Physiol. 134:1206-16; Yu et al. (2003) Plant Cell Physiol. 15:1872-87) y en *Synechocystis* sp. PCC6803 (Weier et al. (2005) Biochem. Biophys. Res. Commun. 334:1127-34; Okazaki, K., et al. (2006) Plant Physiol. 141:546-56). Por ejemplo, una LPAAT localizada en el retículo endoplásmático (microsómico), que está codificada por el gen *LAT2* de *Limnanthes douglasii* o por el gen *LPAT2* de *A. thaliana,* tiene 68 o 69 restos de aminoácidos entre las firmas canónicas de aciltransferasa y puede utilizar sustratos de acilo de C18, tales como CoA C18:1, para la síntesis de lípidos de membrana (Brown et al. (1995) Plant Mol. Biol. 29:267-78; Kim et al. (2005) Plant Cell 17:1073-89). Otras LPAAT ubicadas en el retículo endoplásmático específicas de semilla, pueden tener 63 restos de aminoácidos o aproximadamente 63 restos de aminoácidos entre las firmas canónicas de aciltransferasa y pueden utilizar sustratos de acilo inusuales, por ejemplo, sustratos de acilo C12:0 o C22:1 que, en diversas especies de plantas, se incorporan en lípidos de almacenamiento de semillas (Brown et al. (1995) Plant Mol. Biol. 29:267-78; Knutzon et al. (1995) Plant Physiol. 109:999-1006; Lassner et al. (1995) Plant Physiol. 109:1389-94).

Las LPAAT útiles en los microorganismos y métodos proporcionados en este documento, pueden tener uno o más dominios conservados encontrados en sll1752 LPAAT de cianobacterias (SEQ ID NO: 2), según se caracteriza en la base de datos de dominios conservados mantenida por el Centro Nacional de Información Biotecnológica (disponible en ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), tales como, pero sin limitación: cd07992 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a AAK14816 desconocido), cd07989 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a AGPAT), cd06551 (Lisofosfolípido Aciltransferasas (LPLAT) de la biosíntesis de glicerofosfolípidos), cd07993 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a GPAT), cd07988 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: ABO13168 desconocido), cd07987 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a MGAT), COG0204: PlsC (1-acil-sn-glicerol-3-fosfato aciltransferasa), TIGR00530: AGP aciltrn (1-acil-sn-glicerol-3-fosfato aciltransferasas), cd07991: LPLAT similar a LPCAT1 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a LPCAT1), cd07990 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: similar a LCLAT1) y cd07986 (Lisofosfolípido Aciltransferasas (LPLAT) de la Biosíntesis de Glicerofosfolípidos: ACT 14924 desconocido).

Dado que las LPAAT son esenciales para la producción de ácido fosfatídico en todos los organismos celulares, excepto en las arqueas (*Archaea*), las LPAAT pueden identificarse a partir de una gran variedad de especies, utilizando, por ejemplo, las firmas canónicas de la secuencia de aciltransferasa descritas anteriormente y/o medios bioinformáticos para identificar dominios conservados característicos de las LPAAT, tales como, pero sin limitación, los proporcionados anteriormente, así como por inclusión en Pfam PF01553 ("Aciltransferasa"), que tiene un límite de recopilación de 21,1. La actividad de LPAAT y la especificidad del sustrato pueden determinarse utilizando, por ejemplo, un ensayo de aciltransferasa como se describe en Okazaki, K. et al. (2006) Plant Physiol. 141:546-56, en el documento WO2007141257 o en la patente de los Estados Unidos 5.563.058, en el que cualquiera de los métodos puede adaptarse para utilizar acil-ACP o acil-CoA como sustrato, o utilizando cualquiera de los métodos desvelados en el presente documento.

Para aumentar los niveles de ácidos grasos libres producidos por las células o los microorganismos hospedadores recombinantes, un gen que codifica una LPAAT con una preferencia de sustrato de longitud de cadena de acilo particular, puede expresarse en una cianobacteria recombinante que expresa un gen de tioesterasa no originario como se proporciona en el presente documento. Se puede seleccionar un gen LPAAT en función de la preferencia del sustrato de la enzima codificada en relación con la preferencia del sustrato de la tioesterasa producida por el microorganismo o la célula hospedadora recombinante. En algunos aspectos de la invención, se puede seleccionar una LPAAT que tenga una preferencia de sustrato de longitud de cadena de acilo diferente a la de la tioesterasa producida por la expresión de un gen no originario, que puede ser, por ejemplo, una acil-ACP tioesterasa, una acil-CoA tioesterasa o una 4-hidroxibenzoil tioesterasa. En algunos aspectos de la invención, se selecciona una LPAAT que tenga una preferencia de longitud de cadena de sustrato de acilo complementaria con respecto a la preferencia de sustrato de la tioesterasa producida por la cianobacteria recombinante, donde los sustratos de acilo preferidos por la LPAAT no son sustratos preferidos de la tioesterasa, y los sustratos de acilo preferidos por la tioesterasa no son sustratos preferidos de la LPAAT. Por ejemplo, en determinados aspectos, se selecciona una LPAAT que tenga una preferencia de sustrato de acilo de longitud de cadena media, por ejemplo, una preferencia por usar acil-ACP o acil-CoA C12 y/o C14, y en otros aspectos, se selecciona una LPAAT que tenga una preferencia de sustrato de acil-ACP de longitud de cadena larga, por ejemplo, una preferencia por usar un sustrato de acil-ACP C16 y/o C18. En determinados aspectos, se selecciona una LPAAT que tenga una preferencia de sustrato de acil-ACP de longitud de cadena media y se selecciona una acil-ACP tioesterasa que tenga una preferencia de sustrato de acil-ACP o acil-CoA de longitud de cadena larga. En determinados aspectos, se selecciona una LPAAT que tenga una preferencia de sustrato de acilo de longitud de cadena larga y se selecciona una acil-ACP tioesterasa que tenga una preferencia de sustrato de acil-ACP de longitud de cadena media. El sustrato de acil-ACP preferido puede tener una cadena de acilo saturada o puede tener una cadena de acilo insaturada.

En algunas realizaciones preferidas de la invención, la LPAAT codificada por una secuencia de ácido nucleico no originario en la cianobacteria recombinante, tiene una preferencia de sustrato por moléculas de acil-ACP o acil-CoA

que tienen longitudes de cadena de acilo más largas que los sustratos de acilo preferidos por la tioesterasa producida por la expresión de un gen de tioesterasa no originario en la cianobacteria recombinante. Sin limitar la invención a ningún mecanismo particular, la LPAAT puede tener una preferencia de sustrato por un sustrato de acilo que tenga una cadena de acilo más larga que el sustrato de acilo preferido por la tioesterasa producida por la cianobacteria recombinante, de tal manera que durante la biosíntesis de ácidos grasos, los productos intermedios de acil-ACP son escindidos por la tioesterasa o, si escapan de la escisión a la longitud de cadena de sustrato preferida, sus cadenas de acilo se alargan aún más en el ciclo de biosíntesis de ácidos grasos de modo que alcanzan una longitud de cadena preferida por la LPAAT. La LPAAT puede actuar sobre el sustrato acil-ACP, transfiriendo la cadena de acilo a un glicerolípido y de ese modo eliminando los productos intermedios de acil-ACP que se han alargado más allá de la preferencia de sustrato de la tioesterasa. Dichos productos intermedios podrían acumularse de otra manera en la célula y, por lo tanto, regular negativamente la síntesis de ácidos grasos. Por ejemplo, el microorganismo recombinante puede expresar una acil-ACP tioesterasa, acil-CoA tioesterasa o 4-hidroxibenzoil tioesterasa, que tenga una preferencia de sustrato de acilo de cadena media (por ejemplo, por sustratos de acilo C8, C10, C12 y/o C14 tales como acil-ACP y/o acil-CoA) y una LPAAT que tenga una preferencia de sustrato de acilo de cadena larga (por ejemplo, por sustratos de acilo C16, C18, C20, C22 o C24 tales como acil-ACP y/o acil-CoA). En algunos ejemplos, la cianobacteria recombinante expresa un gen no originario que codifica una tioesterasa que tiene preferencia por un sustrato de acilo C12 y/o C14 y expresa un gen no originario que codifica una LPAAT que tiene preferencia por uno o más sustratos de acilo C16 y/o C18. En otro ejemplo, tanto la tioesterasa como la LPAAT codificadas por genes no originarios en la célula recombinante, pueden tener preferencias de sustrato de acilo de cadena larga, donde la LPAAT prefiere uno o más sustratos que tengan una longitud de cadena de acilo más larga que la longitud de cadena de acilo del sustrato (o sustratos) preferido por la tioesterasa. Por ejemplo, una cianobacteria recombinante puede expresar un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato de acilo C16 y un gen no originario que codifique una LPAAT que tenga una preferencia de sustrato de acilo C18. En ejemplos adicionales, la cianobacteria recombinante expresa un gen no originario que codifica una tioesterasa que tiene preferencia por un sustrato de acilo C12, C14, y/o C16 y expresa un gen no originario que codifica una LPAAT que tiene preferencia por un sustrato de acilo C18. Por ejemplo, la cianobacteria recombinante puede expresar un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato de acilo C12 o una preferencia de sustrato de acilo C14. Por ejemplo, la cianobacteria recombinante puede expresar una LPAAT que tenga una preferencia de sustrato de acilo C18 y puede expresar un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato de acilo C16, o la cianobacteria recombinante puede expresar un gen no originario que codifique una tioesterasa que tenga una preferencia de sustrato de acilo C14 y C16 o una preferencia por sustratos de acilo C12, C14 y C16.

Las LPAAT de procariotas comúnmente tienen una preferencia de sustrato de acilo C16, aunque la invención considera que las LPAAT de procariotas tienen preferencias de sustrato por cualquier longitud de cadena de acilo, en particular, pero no exclusivamente, longitudes de cadena C16 y C18.

Se ha demostrado que la LPAAT codificada por el gen *sll1752* de *Synechocystis* sp. PCC 6803 (número de registro de GenBank/EMBL NP 440396 o BAA17076; SEQ ID NO: 2), cuyos ortólogos están ampliamente conservados en genomas de cianobacterias publicados, tiene preferencia por sustratos de acilo C18 0 y C18:1 (Okazaki, K., et al. (2006) Plant Physiol. 141:546-56). La LPAAT codificada por el gen *sll1752* de *Synechocystis* sp. PCC 6803, tiene una distancia de 64 restos de aminoácidos entre las firmas canónicas de aciltransferasa, que es similar a la de las LPAAT específicas de semillas de plantas superiores.

Las LPAAT de cianobacterias consideradas como ortólogos del polipéptido codificado por sll 1752 (Registro de Genbank NP_440396; identificador génico 16329668; SEQ ID NO: 2) de *Synechocystis* incluyen, pero sin limitación, la fosfolípido/glicerol aciltransferasa de *Cyanothece* sp. PCC 8801 (Registro de Genbank YP_002372580; identificador génico 218247209; SEQ ID NO: 3), la fosfolípido/glicerol aciltransferasa de *Crocosphaera watsonii* WH 8501 (Registro de Genbank ZP_00513647; identificador génico 67920127; SEQ ID NO: 4), la proteína hipotética CY0110_27159 de *Cyanothece* sp. CCY0110 (Registro de Genbank ZP_01731155; identificador génico 126660033; SEQ ID NO: 5), el producto proteico sin nombre de *Microcystis aeruginosa* PCC 7806 (Registro de Genbank CAO89042; identificador génico 159026728; SEQ ID NO: 6), la proteína hipotética all4871 de *Nostoc* sp. PCC 7120 (Registro de Genbank NP_488911; identificador génico 17232363; SEQ ID NO: 7), el producto proteico sin nombre de *Anabaena variabilis* ATCC 29413 (Registro de Genbank YP_322656; identificador génico 75908360; SEQ ID NO: 8), la fosfolípido/glicerol aciltransferasa de *'Nostoc azollae'* 0708 (Registro de Genbank YP_003721758; identificador génico 298491581; SEQ ID NO: 9), la fosfolípido/glicerol aciltransferasa de *Raphidiopsis brookii* D9 (Registro de Genbank ZP_06305519; identificador génico 282897518; SEQ ID NO: 10), la proteína de dominio aciltransferasa de *Microcoleus chthonoplastes* PCC 7420 (Registro de Genbank ZP_05026385; identificador génico 254412612; SEQ ID NO: 11), la fosfolípido/glicerol aciltransferasa de *Cylindrospermopsis raciborskii* CS-505 (Registro de Genbank ZP_06306989; identificador génico 282899007; SEQ ID NO: 12), la proteína hipotética conservada de *Oscillatoria* sp. PCC 6506 (Registro de Genbank ZP_07110610; identificador génico 300865867; SEQ ID NO: 13), la fosfolípido/glicerol aciltransferasa de *Nodularia spumigena* CCY9414 (Registro de Genbank ZP_01628592; identificador génico 119509444; SEQ ID NO: 14), y la fosfolípido/glicerol aciltransferasa de *Microcoleus vaginatus* FGP-2 (Registro de Genbank ZP_08495621; identificador génico 334121556; SEQ ID NO:15). Estos y otros ortólogos de sll1752 (SEQ ID NO: 2) de *Synechocystis* se tienen en cuenta para su uso en los microorganismos y métodos proporcionados en el presente documento. Las cianobacterias recombinantes que incluyen una secuencia

de ácido nucleico no originario que codifica una LPAAT y que producen al menos un ácido graso libre o al menos un derivado de ácido graso, pueden incluir secuencias no originarias que codifican las LPAAT que tienen una identidad de al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos el 99 % o 100 % con la SEQ ID NO:2.

Tioesterasas

Como se usa en el presente documento, el término "tioesterasa" pretende incluir hidrolasas capaces de actuar sobre un enlace tioéster. Dichas enzimas pueden corresponder, por ejemplo, a los números de la Comisión de Enzimas 3.1.2.2, 3.1.2.14, 3.1.2.18, 3.1.2.19, 3.2.1.20, 3.1.2.22, 3.1.2.23 o 3.1.2.27. Una tioesterasa expresada en un microorganismo recombinante o en una célula hospedadora de la invención, es una acil-ACP tioesterasa. Por ejemplo, una cianobacteria o célula hospedadora recombinante puede transformarse con un gen que codifique una acil-ACP tioesterasa exógena, tal como un gen que codifique un polipéptido que cuando se consulte en la base de datos Pfam, proporcione una coincidencia con Pfam PF01643 que tenga una puntuación en bits menor que o igual a 20,3 (el límite de recopilación de PF01643). El gen de acil-ACP-tioesterasa codifica una acil-ACP-tioesterasa de una especie de planta superior. Los genes que codifican acil-ACP tioesterasas procedentes de plantas superiores pueden incluir, sin limitación, genes que codifican acil-ACP tioesterasas de especies del género *Cuphea* (por ejemplo, *Cuphea carthagenensis, Cuphea wrightii* (por ejemplo, AAC49784.1 GI:1336008), *Cuphea lanceolata* (por ejemplo, CAA54060, GI495227), *Cuphea palustris,* (por ejemplo, AAC49783.1 GI:1336006; AAC49179.1 GI:1215718); *Cuphea hookeriana* (por ejemplo, AAC72882.1 GI:3859830; AAC49269.1 GI:1292906; AAC72881.1 GI:3859828; AAC72883.1 GI:3859832), *Cuphea calophylla* (por ejemplo, ABB71580.1 GI:81361963) o genes de varias especies de *Cuphea* desveladas en la publicación de solicitud de patente de los Estados Unidos US 2011/0020883, o genes de otras especies de plantas superiores. Por ejemplo, una cianobacteria utilizada en los métodos y cultivos desvelados en el presente documento, puede incluir un gen que codifique una acil-ACP tioesterasa de especies tales como, aunque si limitación, *Arabidopsis* (XP_002885681.1 GI:297835598; NP_172327.1 GI:15223236); *Arachis hypogaea* (por ejemplo, AB038556.1 GI:133754634); especies de *Brassica* (por ejemplo, CAA52069.1 GI: 435011), *Camelia oleifera* (por ejemplo, ACQ57189.1 GI:229358082); *Cinnamonum camphorum* (por ejemplo, AAC49151.1 GI:1143156); *Cocos nucifera; Glycine max* (por ejemplo, ABD91726.1 GI:90192131); *Garcinia mangostana* (por ejemplo, AAB51525.1 GI:1930081); *Gossypium hirsutum* (por ejemplo, AAD01982.1 GI:4104242); *Helianthus annuus* (por ejemplo, AAQ08226 GI:33325244); *Jatropha curcas* (por ejemplo, ABU96744.1 GI:156900676); *Macadamia tetraphylla* (por ejemplo, ADA79524.1 GI:282160399); *Elaeis oleifera* (por ejemplo, AAM09524.1 GI:20067070); *Elaeis guineensis* (por ejemplo, AAD42220 GI:30962820); *Oryza sativa* (por ejemplo, BAA83582.1 GI:5803272); *Populus tomentosa* (por ejemplo, ABC47311 GI:83778888); *Umbellularia californica* (por ejemplo, AAC49001.1 GI:595955); *Ulmus americana* (por ejemplo, AAB71731.1 GI:2459533); y *Zea mays* (ACG41291.1 GI:195643646), o cualquiera de las desveladas en las patentes de los Estados Unidos N.º 5.455.167; 5.654.495 y 5.455.167 y en las publicaciones de solicitud de patente de Estados Unidos N.º 2009/0298143 y 2011/0020883. También se incluyen acil-ACP tioesterasas de musgos (*Bryophyta*), tal como, por ejemplo, *Physcomitrella patens,* (por ejemplo, XP 001770108 GI:168035219). Estos ejemplos no son limitantes con respecto a los tipos o ejemplos específicos de genes de acil-ACP tioesterasa que pueden usarse.

Los genes que codifican acil-ACP tioesterasas de plantas y briófitas, pueden incluir secuencias que codifican péptidos de tránsito para el transporte a los plástidos que opcionalmente se pueden delecionar para la expresión de tioesterasas truncadas en el extremo N en cianobacterias hospedadoras. Los genes que codifican tioesterasas pueden incluir opcionalmente además, truncamientos adicionales, tales como, aunque sin limitación, truncamientos en el extremo N-terminal que incluyen una parte, tal como hasta diez, hasta veinte, o hasta treinta aminoácidos, de una secuencia de acil-ACP tioesterasa madura.

Normalmente, las acil-ACP tioesterasas pueden ser activas hasta cierto punto en sustratos de acil-ACP que tienen una pluralidad de diferentes longitudes de cadena de acilo, pero puede tener mayor actividad (por ejemplo, tener una preferencia de sustrato por) sobre uno o más sustratos de acil-ACP que tienen longitudes de cadena acilo particulares que en otros sustratos de longitud de cadena. En algunos ejemplos, una enzima que tiene una preferencia de sustrato por sustratos de longitud de cadena acilo particulares, produce al menos el doble de producto del sustrato o sustratos preferidos que de un sustrato no preferido y, por ejemplo, puede producir al menos tres veces, al menos cuatro veces o al menos cinco veces más producto de un sustrato preferido que de un sustrato no preferido. Por ejemplo, una acil-ACP tioestesasa puede tener una preferencia de sustrato por uno o más sustratos de acil-ACP que tienen longitudes de cadena de acilo de 8, 10, 12, 14, 16, 18, 20, 22 y/o 24 carbonos. Además o como alternativa, la acil-ACP tioesterasa puede hidrolizar uno o más sustratos de acil-ACP que tengan una longitud de cadena de acilo de 8 a 18 carbonos, por ejemplo, de 12 a 18 carbonos. Por ejemplo, la acil-ACP tioesterasa puede hidrolizar uno o más sustratos de acil-ACP que tengan una longitud de cadena de acilo de 12 a 16 carbonos. Además o como alternativa, una acil-ACP tioesterasas de la presente invención puede, en algunas realizaciones, tener nivel de actividad más alto en un sustrato de acil-ACP que tenga una longitud de cadena de acilo de 12, 14 y/o 16 carbonos.

Microorganismos y células hospedadoras

Las algas, para uso en la invención, incluyen, sin limitación, microalgas, tales como, pero sin limitación, especies de

los géneros *Achnanthes, Amphiprora, Amphora, Ankistrodesmus, Asteromonas, Boekelovia, Borodinella, Botryococcus, Bracteococcus, Chaetoceros, Carteria, Chlamydomonas, Chlorococcum, Chlorogonium, Chlorella, Chroomonas, Chrysosphaera, Cricosphaera, Crypthecodinium, Cryptomonas, Cyclotella, Dunaliella, Ellipsoidon, Emiliania, Eremosphaera, Ernodesmius, Euglena, Franceia, Fragilaria, Gloeothamnion, Haematococcus, Halocafeteria, Hymenomonas, Isochrysis, Lepocinclis, Micractinium, Monoraphidium, Nannochloris, Nannochloropsis, Navicula, Neochloris, Nephrochloris, Nephroselmis, Nitzschia, Ochromonas, Oedogonium, Oocystis, Ostreococcus, Pavlova, Parachlorella, Pascheria, Phaeodactylum, Phagus, Picochlorum, Platymonas, Pleurochrysis, Pleurococcus, Prototheca, Pseudochlorella, Pseudoneochloris, Pyramimonas, Pyrobotrys, Scenedesmus, Schizochytrium, Skeletonema, Spyrogyra, Stichococcus, Tetraselmis, Viridiella* o *Volvox.*

Las especies de cianobacterias que pueden utilizarse para la producción de productos de ácidos grasos incluyen, sin limitación, especies de los géneros *Agmenellum, Anabaena, Anabaenopsis, Anacystis, Aphanizomenon, Arthrospira, Asterocapsa, Borzia, Calothrix, Chamaesiphon, Chroococcus, Chlorogloeopsis, Chroococcidiopsis, Chroococcus, Crinalium, Cyanobacterium, Cyanobium, Cyanocystis, Cyanospira, Cyanothece, Cylindrospermopsis, Cylindrospermum, Dactylococcopsis, Dermocarpella, Fischerella, Fremyella, Geitleria, Geitlerinema, Gloeobacter, Gloeocapsa, Gloeothece, Flalospirulina, lyengariella, Leptolyngbya, Limnothrix, Lyngbya, Microcoleus, Microcystis, Myxosarcina, Nodularia, Nostoc, Nostochopsis, Oscillatoria, Phormidium, Planktothrix, Pleurocapsa, Prochlorococcus, Prochloron, Prochlorothrix, Pseudanabaena, Rivularia, Schizothrix, Scytonema, Spirulina, Stanieria, Starria, Stigonema, Symploca, Synechococcus, Synechocystis, Thermosynechococcus, Tolypothrix, Trichodesmium, Tychonema* y *Xenococcus.* Por ejemplo, el microorganismo fotosintético recombinante puede ser una especie de los géneros*Cyanobium, Cyanothece* o *Cyanobacterium*, o además, de manera alternativa, el microorganismo fotosintético recombinante puede ser una especie de los géneros *Gloeobacter, Lyngbya o Leptolyngba.* De manera alternativa, el microorganismo fotosintético recombinante puede ser una especie de los géneros *Synechococcus, Synechocystis* o *Thermosynechococcus.* Se conocen diversas especies de cianobacterias y se han manipulado utilizando técnicas de biología molecular, incluyendo entre ellas las cianobacterias unicelulares *Synechocystis* sp. PCC6803 y *Synechococcus elongates* PCC7942, cuyos genomas se han secuenciado por completo.

La invención proporciona una cianobacteria recombinante que comprende una molécula de ácido nucleico no natural que incluye una secuencia de ácido nucleico que codifica una LPAAT y una secuencia de ácido nucleico no natural que codifica una tioesterasa. La secuencia de ácido nucleico que codifica la LPAAT y la secuencia de ácido nucleico que codifica la tioesterasa pueden estar presentes en la misma o en diferentes moléculas de ácido nucleico. La cianobacteria recombinante puede producir un producto de ácido graso, tal como un ácido graso, alcohol graso, aldehído graso, éster de ácido graso, éster de cera o hidrocarburo (tal como un alcano o alqueno). La célula hospedadora recombinante puede comprender, por ejemplo, cualquiera de las secuencias de ácido nucleico que codifican una LPAAT descrita en el presente documento y puede comprender cualquiera de las secuencias de ácido nucleico que codifican una tioesterasa descrita en el presente documento. La célula hospedadora recombinante puede comprender, por ejemplo, cualquiera de los vectores descritos en el presente documento.

La célula hospedadora recombinante puede comprender un gen no originario que codifique una LPAAT con una identidad de secuencia de al menos 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % con la SEQ ID NO: 2.

Como ejemplos ilustrativos de cianobacterias recombinantes que expresan un gen de LPAAT no natural y un gen de tioesterasa no natural, se incluyen, cianobacterias recombinantes que expresan un gen no originario que codifica una LPAAT que tiene una preferencia de sustrato de C18, tal como una LPAAT que tiene una identidad de al menos 85 % con la SEQ ID NO: 2 (por ejemplo, una identidad de al menos 90 % o de al menos 95 % con la SEQ ID NO: 2) y una acil-ACP tioesterasa de planta superior que tiene una preferencia de sustrato por uno o más de un sustrato de acilo C12, C14 o C16, por ejemplo, la tioesterasa FatB1 de *Umbellularia californica* que prefiere C12 (AAC49001.1 GI:595955), la tioesterasa de *Cinnamomum camphorum* que prefiere C14 (Q39473.1 GI:8469216), la acil-ACP tioesterasa de *Elaeis oleifera* truncada en el extremo N-terminal que prefiere C16 (SEQ ID NO: 56), la acil-ACP tioesterasa de *Elaeis guineensis* que prefiere C16 (SEQ ID NO: 57) o una versión truncada de la misma, la tioesterasa FatB2 de *Cuphea palustris* que prefiere C14 y C16 (AAC49180.1 GI:1215720), la tioesterasa Cc1FatB1 de *Cuphea carthagenensis,* o una variante de la misma (por ejemplo, el polipéptido truncado de SEQ ID NO: 44), o acil-ACP tioesterasas de otras especies del género *Cuphea* (véase, por ejemplo, la publicación de solicitud de patente de los Estados Unidos US 2011/0020883), que incluye, por ejemplo, la acil-ACP tioesterasa Cd1FatB1 que prefiere C14, C16 truncada en el extremo N terminal de SEC ID Nº: 54 o una variante de la misma, la tioesterasa Cp1FatB1 C14 o que prefiere C16 truncada en el extremo N terminal de SEQ ID NO: 55 o una variante de la misma, o la tioesterasa C. *hookeriana* FatB1 que prefiere C16 (Q39513.1 GI:8469217). También se consideran específicamente las tioesterasas que tienen una identidad de al menos 85 %, por ejemplo, una identidad de al menos 90 % o de al menos 95 % con las acil-ACP tioesterasas de *Umbelularia, Cinnamomum, Cuphea* y *Elaeis* mencionadas anteriormente, en donde las acil-ACP tioesterasas tienen especificidades de sustrato C12, C14 y/o C16.

La célula hospedadora recombinante que expresa un gen de LPAAT en combinación con un gen de tioesterasa, produce una mayor cantidad de un producto de ácido graso que una célula hospedadora de control que no expresa

el gen de LPAAT. En algunas realizaciones, la cantidad de producto de ácido graso que produce un cultivo de la célula hospedadora recombinante que expresa un gen no originario que codifica una LPAAT y un gen no originario que codifica una tioesterasa es al menos 10 %, 15 %, 20 %, 25 %, 30 %, 40 %, 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, 90 %, 95 %, 100 %, 125 %, 150 %, 175 %, 200 %, 225 %, 250 %, 275 %, 300 %, 325 %, 350 %, 375 %, 400 %, 425 %, 450 %, 475 %, 500 %, 525 %, 550 %, 575 %, 600 %, 625 %, 650 %, 675 %, 700 %, 725 %, 750 %, 775 %, 800 %, 825 %, 850 %, 875 %, 900 %, 925 %, 950 %, 975 % o 1 000 % mayor que la cantidad de éster de cera producida por una célula hospedadora de control que no expresa el gen de LPAAT no originario.

En algunas realizaciones, la célula hospedadora recombinante que expresa una LPAAT y una tioesterasa produce una mayor cantidad de un producto de ácido graso que una célula hospedadora de control que no expresa la LPAAT y la tioesterasa. En algunas realizaciones, la cantidad de un producto de ácido graso producido por un cultivo de la célula hospedadora recombinante que expresa una LPAAT y una tioesterasa es al menos 10 %, 15 %, 20 %, 25 %, 30 %, 40 %, 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, 90 %, 95 %, 100 %, 125 %, 150 %, 175 %, 200 %, 225 %, 250 %, 275 %, 300 %, 325 %, 350 %, 375 %, 400 %, 425 %, 450 %, 475 %, 500 %, 525 %, 550 %, 575 %, 600 %, 625 %, 650 %, 675 %, 700 %, 725 %, 750 %, 775 %, 800 %, 825 %, 850 %, 875 %, 900 %, 925 %, 950 %, 975 % o 1 000 % mayor que la cantidad de producto de ácido graso que produce una célula hospedadora de control que no expresa la LPAAT y la tioesterasa.

Modificaciones adicionales de la célula hospedadora para la producción de ácido graso libre (FFA)

Se pueden realizar modificaciones adicionales en las cianobacterias recombinantes o en las células hospedadoras de la invención, para afectar a la producción de ácidos grasos libres. Por ejemplo, la cianobacteria recombinante puede incluir uno o más genes no originarios adicionales cuya expresión aumente la producción de un producto de ácido graso. El gen adicional puede estar codificado por una molécula de ácido nucleico que sea igual que la molécula de ácido nucleico que codifique la LPAAT y/o que la molécula de ácido nucleico que codifique la tioesterasa, o el gen adicional puede estar codificado por una molécula de ácido nucleico distinta. Cuando dos o más genes están codificados por la misma molécula de ácido nucleico (por ejemplo, en el mismo vector de expresión), opcionalmente, la expresión de cada gen puede regularse independientemente por un mismo o diferente promotor y/o potenciador. En determinadas realizaciones, el gen adicional puede aumentar la velocidad y/o el nivel de producción de derivados de ácidos grasos o ácidos grasos libres.

Por ejemplo, la célula hospedadora recombinante que expresa una LPAAT y una tioesterasa, puede expresar además al menos un gen recombinante o exógeno adicional que codifique un polipéptido que tenga actividad lipolítica. Por ejemplo, una célula hospedadora recombinante de la invención puede incluir uno o más genes que codifiquen uno o más polipéptidos que tengan actividad lipolítica, donde el polipéptido o los polipéptidos que tienen actividad lipolítica son capaces de producir ácidos grasos libres a partir de lípidos de membrana o lípidos de almacenamiento, por ejemplo, fosfolípidos, triacilgliceroles, diacilgliceroles, monoacilgliceroles, o similares, o combinaciones de los mismos. Los polipéptidos que tienen actividad lipolítica pueden ser, por ejemplo, lipasas, esterasas, cutinasas o amidasas.

Sin limitar la invención a ningún mecanismo particular, se considera que la sobreexpresión de un gen de LPAAT puede dar como resultado una producción significativamente mayor de lípidos de membrana o de almacenamiento. El exceso de lípidos de membrana o de almacenamiento puede activarse por polipéptidos con actividad lipolítica (tal como, por ejemplo, una o más lipasas o amidasas) para liberar ácidos grasos adicionales (que opcionalmente pueden convertirse en derivados de ácidos grasos), por lo tanto aumentando adicionalmente la cantidad de productos de ácidos grasos que produce el microorganismo recombinante (véase, por ejemplo, el modelo proporcionado en la Figura 2).

El uso de genes que codifican polipéptidos que tienen actividad lipolítica en microorganismos para la producción de ácidos grasos libres, se desvela en la solicitud de patente de Estados Unidos N.º 13/324,653 otorgada conjuntamente titulada "Production of Free Fatty Acids and Fatty Acid Derivatives by Recombinant Microorganisms Expressing Polypeptides Having Lipolytic Activity", presentada el 13 de diciembre de 2011. El polipéptido que tiene actividad lipolítica puede ser, por ejemplo, una lipasa, por ejemplo, que libera un ácido graso de un glicerolípido (incluido un monoglicérido, un diglicérido, un triglicérido, un fosfolípido, un galactolípido, *etc.)* o puede ser una amidasa. Por ejemplo, el microorganismo recombinante puede incluir un gen no originario que codifique una lipasa, como, pero sin limitación, una lipasa que sea un miembro de un Pfam perteneciente al clan AB Hidrolasa Pfam (CL0028). Por ejemplo, un gen no originario que codifica un polipéptido que tiene actividad lipolítica puede codificar una lipasa que incluya un dominio LipA identificado como dominio de proteína conservado COG1075, o que esté incluido en la familia de proteínas Pfam PF01674 (Lipasa 2); una molécula de ácido nucleico no originario que codifique una lipasa que incluya un dominio de Lipasa 3 identificado como dominio de proteína conservado COG3675, o que esté incluido en la familia de proteínas Pfam PF01764 (Lipasa 3); una molécula de ácido nucleico no originario que codifique una lipasa que esté incluida en la familia de proteínas Pfam PF07819 (PGAP1); o una molécula de ácido nucleico no originario que codifique un polipéptido que esté incluido en cualquiera de las familias de proteínas Pfam PF03583, Pfam PF00151 (lipasa), Pfam PF00561 (Ab hidrolasa 1), Pfam PF02230 (Ab hidrolasa 2), Pfam PF07859 (Ab hidrolasa 3), Pfam PF08386 (Ab hidrolasa 4), Pfam PF12695 (Ab hidrolasa 5), Pfam PF12697 (Ab hidrolasa 6), Pfam PF12715 (Ab hidrolasa 7), Pfam PF04083 (Ab hidro lipasa). En algunas realizaciones, un gen

exógeno de lipasa introducido en un microorganismo puede codificar una proteína con una secuencia de aminoácidos que tenga un parámetro de valor E de 0.01 o menor cuando se consulta utilizando el Perfil Pfam de HMM (modelos ocultos de Markov) para cualquiera de Pfam PF01674, Pfam PF 01764, Pfam PF07819, Pfam PF03583 y/o Pfam PF00151. Además, el microorganismo recombinante puede incluir un gen no originario que codifique una amidasa que tenga actividad lipolítica, tal como, pero sin limitación, una amidasa que recluta a Pfam PF01425 (Amidasa) con una puntación en bits mayor que el límite de recopilación de 20,1 y puede catalizar la liberación de ácidos grasos procedente de lípidos.

Además o como alternativa, se contemplan cianobacterias recombinantes que están genomodificadas para que incluyan secuencias reguladoras de genes que induzcan o aumenten la expresión de un gen endógeno de lipasa. Por ejemplo, una cianobacteria puede genomodificarse de tal manera que se inserte un promotor heterólogo cadena arriba de una región codificante de un gen endógeno de lipasa. El promotor heterólogo puede reemplazar a un promotor endógeno y/o puede insertarse cadena arriba o cadena abajo del promotor endógeno que regula la expresión del gen endógeno de lipasa, por ejemplo, usando recombinación homóloga o recombinación específica de sitio. El promotor heterólogo puede ser un promotor constitutivo o un promotor inducible que aumente la expresión del gen endógeno de lipasa.

Se pueden realizar modificaciones adicionales a las cianobacterias recombinantes de la invención para afectar a la producción de ácidos grasos libres. Por ejemplo, una cianobacteria recombinante o una célula hospedadora de la invención, puede incluir una o más moléculas de ácido nucleico exógeno que codifiquen un polipéptido que participe en la síntesis de un ácido graso, incluyendo, pero sin limitación, una acetil-CoA carboxilasa, una malonil CoA: ACP transacilasa o una beta-cetoacil-ACP sintasa.

Adicionalmente, o de manera alternativa, la cianobacteria recombinante puede comprender una modificación de una molécula de ácido nucleico endógena que codifique, por ejemplo, una acil-CoA sintetasa, acil-ACP sintetasa, acil CoA deshidrogenasa, glicerol-3-fosfato deshidrogenasa, acetaldehído CoA deshidrogenasa, piruvato deshidrogenasa, acetato cinasa, y similares, y combinaciones de las mismas. En determinadas realizaciones, la modificación regula negativamente el ácido nucleico endógeno e incluye deleción, silenciamiento o inactivación parcial, sustancial o completa, del ácido nucleico o sus elementos reguladores.

En algunos ejemplos, la célula hospedadora puede tener expresión atenuada de un gen endógeno que codifique una acil-ACP sintetasa que participe en el reciclaje de ácidos grasos en lípidos. El gen endógeno de acil-ACP sintetasa puede estar, por ejemplo, regulado negativamente por deleción o mutación del promotor, o la proteína que codifica el gen puede delecionarse o alterarse internamente, por ejemplo, por mutagénesis insercional. De manera alternativa, se puede delecionar todo el gen de la acil-CoA sintetasa, por ejemplo, por recombinación homóloga u otras técnicas de modificación del genoma. En otras alternativas adicionales, para atenuar la expresión del gen endógeno de la acil-ACP sintetasa, en la célula hospedadora pueden introducirse construcciones que disminuyan la expresión (*knockdown*) de genes, tales como, aunque sin limitación, construcciones ribozimáticas, antisentido, o de iARN.

Adicionalmente, o de manera alternativa, la cianobacteria recombinante puede modificarse de modo que uno o más genes que codifican las enzimas de la ruta de oxidación beta se hayan inactivado y/o regulado negativamente, y/o de modo que las propias enzimas que están operativas en dichas rutas de oxidación beta puedan inhibirse. Esto podría impedir la degradación de los ácidos grasos liberados por las acil-ACP, mejorando así el rendimiento de los productos de ácidos grasos. En los casos en los que los productos deseados sean ácidos grasos de cadena media, se puede realizar la inactivación y/o regulación negativa de genes que codifican las enzimas acil-CoA sintetasa y/o acil-CoA oxidasa que utilizan preferentemente estas longitudes de cadena como sustratos. Las mutaciones en los genes que codifican enzimas acil-CoA sintetasa y/o acil-CoA oxidasa específicas de cadena media, de manera que la actividad de las enzimas se vea disminuida, pueden ser adicional o alternativamente eficaces para aumentar el rendimiento de los productos de ácidos grasos producidos y/o liberados. En los genes pueden introducirse mutaciones utlizando métodos recombinantes o no recombinantes.

Los genes pueden ser específicamente objeto de una alteración, deleción, generación de secuencias antisentido, generación de ribozimas, iARN, modificación del genoma de meganucleasas y/u otras estrategias recombinantes. La inactivación de los genes, puede realizarse además o como alternativa, mediante técnicas de mutación al azar, tales como exposición a UV y/o a mutágenos químicos, y los genes y/o las enzimas resultantes pueden explorarse con respecto a mutantes con la actividad deseada. Las propias proteínas pueden inhibirse mediante la generación intracelular de anticuerpos apropiados, la generación intracelular de inhibidores peptídicos, o similares, o alguna de estas combinaciones.

En realizaciones adicionales, una cianobacteria recombinante de la invención puede modificarse de modo que uno o más genes codificantes de enzimas de la ruta de la biosíntesis de almacenamiento de hidratos de carbono y/o polihidroxialcanoato (PHA) estén inactivados o regulados negativamente, y/o de modo que las propias enzimas que están operativas en dichas rutas estén inhibidas. Como ejemplos se incluyen, pero sin limitación, enzimas implicadas en la síntesis de glicógeno, almidón o crisolaminarina, incluyendo glucanosintasas y/o enzimas ramificantes. Otros ejemplos incluyen enzimas implicadas en la biosíntesis de PHA tales como acetoacetil-CoA sintasa y PHA sintasa.

Modificaciones para la producción de derivados de ácidos grasos

Las cianobacterias recombinantes de la invención pueden incluir modificaciones adicionales para la producción de derivados de ácidos grasos tales como, por ejemplo, aldehídos grasos, alcoholes grasos, ésteres de ácidos grasos, ésteres de cera e hidrocarburos, incluyendo alcanos y alquenos. En algunas realizaciones, la cianobacteria recombinante comprende una o más moléculas de ácido nucleico que codifican una acil-CoA reductasa, una ácido carboxílico reductasa y/o una acil-ACP reductasa exógena y puede producir un alcohol graso.

Por ejemplo, para la producción de aldehídos grasos, que opcionalmente se pueden convertir en productos tales como alcoholes grasos, ésteres de cera o alcanos, una cianobacteria transgénica, como se proporciona en el presente documento, puede incluir uno o más genes exógenos que codifiquen una reductasa formadora de aldehído, tales como, por ejemplo, una acil-CoA reductasa formadora de aldehído, una acil-ACP reductasa formadora de aldehído, o una ácido carboxílico reductasa. Los genes, o partes de los genes, que se enumeran en el GenBank y en otras bases de datos genéticas y que se espera que codifiquen proteínas que sean homólogas a las acil-CoA reductasas conocidas que producen aldehídos grasos, denominadas en el presente documento "acil-CoA reductasas grasas generadoras de aldehído", pueden introducirse en varios microorganismos para analizar la producción de aldehídos grasos o alcoholes grasos específicos producidos a partir de ellos. Como ejemplos no limitantes de acil-CoA reductasas generadoras de aldehído graso se incluyen el gen de Acr1 de *Acinetobacter baylyi* (Registro U77680, GI:1684885), el gen de AcrM-1 de *Acinetobacter* sp. M-1 (Registro YP 001086217, GI:18857900) y los genes de luxC y luxE de diversas bacterias fotoluminiscentes, por ejemplo, de especies de los géneros *Altermonas, Photobacterium, Shewanella, Vibrio* o *Xenorhabdus.* Las enzimas codificadas por estos y otros genes identificados, por ejemplo, por homología de secuencias o dominios de proteínas, pueden analizarse para determinar sus sustratos y productos usando ensayos conocidos en la técnica.

Como ejemplos no limitantes de ácido carboxílico reductasas que pueden utilizarse en la invención se incluyen el gen del CAR (receptor de antígeno quimérico) de *Nocardia* (Registro AY495697; GI:40796034) y sus homólogos, algunos de los cuales se desvelan en el documento US2010/0105963.

En algunos ejemplos, la célula hospedadora puede incluir un gen no originario que codifica una acil-ACP reductasa formadora de aldehído tal como, pero sin limitación, cualquiera de los desvelados en los documentos WO 2009/140696 y WO 2011/066137. Por ejemplo, la célula hospedadora recombinante puede comprender una acil-ACP reductasa formadora de aldehído que tenga una identidad de secuencia de al menos 50 %, 60 %, 70 %, 80 %, 90% o 95 % con una reductasa formadora de aldehído, por ejemplo, como se desvela en los documentos WO 2009/140696 o WO 2011/066137, tal como, por ejemplo, cualquiera de las reductasas que tengan los números de registro AAM82647; AAM82647; BAD78241; ABA22149; BAB76983; ZP_03763674; ACL42791; ZP_01628095; ZP_01619574; YP_001865324; YP_721978; NP_682102; YP_001518341; YP_002371106; ZP_05027136; ZP_03273554; NP_442146; ZP_01728620; ZP_05039135; YP_001802846; NP_926091; YP_001660322; ZP_00516920; CAO90781; ZP_01085337; YP_001227841; ABD96327; NP_897828; YP_001224378; ABD96480; ZP_01123215; ABB92249; ZP_01079773; YP_377636; NP_874926; NP_895058; ABD96274; ABD96442; ZP_01469469; ZP_05045052; YP_001014416; YP_001010913; YP_381056; YP_001550421; NP_892651; YP_001090783; ZP_01472595; YP_293055; ZP_05138243; YP_731192; YP_001483815; YP_001008982; YP_473896; YP_478638; o YP_397030. En algunas realizaciones, la célula hospedadora recombinante incluye un gen exógeno que codifica una acil-ACP reductasa formadora de aldehído, donde la acil-ACP reductasa formadora de aldehído puede ser de una especie de cianobacteria y puede ser de la misma especie que la del microorganismo hospedador, o puede ser de una especie diferente. De manera alternativa, un hospedador de cianobacteria puede genomodificarse para que sobreexprese un gen endógeno de acil-ACP reductasa.

Para la producción de alcoholes grasos, una cianobacteria recombinante, como se proporciona en el presente documento, puede incluir un gen exógeno que codifique una acil reductasa formadora de alcohol tal como bfar de *Bombyx mori;* jjfar de *Simmondsia chinensis,* una acil-CoA reductasa de *Titicum aestivum,* mfar1 de *Mus musculus,* mfar2 de *Mus musculus,* hfar de *H. sapiens,* FARXIII de *Ostrinia scapulalis,* MS2 de *Z. mays,* la supuesta acil-coA reductasa grasa de *Oryza sativa* (registro de GenBank BAC84377) o MS2, FAR4, FAR6 o CER4 de *Arabidopsis thaliana.* Una acil-CoA reductasa grasa formadora de alcohol también puede ser una enzima procariota, tal como, por ejemplo, las que tienen los números de registro de Genbank AAC45217 *(*acil-CoA reductasa grasa de *Acinetobacter baylyi*), YP_047869 *(Acinetobacter* sp. ADP1 acil-CoA reductasa grasa), BAB85476 *(*M-1 acil coenzima A reductasa de *Acinetobacter* sp.), YP_001086217 *(acil coenzima A reductasa de Acinetobacter baumannii* ATCC 17978), YP_580344 SDR deshidrogenasa/reductasa de cadena corta *[Psychrobacter cryohalolentis* K5], YP_001280274 *(SDR deshidrogenasa/reductasa de cadena corta de Psychrobacter* sp. PRwf-1), la acil reductasa de *Marinobacter algicola* DG893 (Registro ZP_01892457), la acil deshidrogenasa de cadena corta de *Marinobacter aquaeolei* Maqu_2507 (YP_959769) *Marinobacter aquaeolei* VT8 Maqu_2220 (YP_959486), *Hahella chejuensis* Hch_05075 (YP_436183), *Marinobacter adhaerens* HP15_810 (ADP96574) o una acil reductasa de especies del género *Oceanobacter* (e.g., RED65_09894, Registro EAT13695). Las reductasas formadoras de alcohol pueden incluir las que pueden utilizar acil-ACP como sustrato, como se desvela en la solicitud de patente de Estados Unidos en trámite, otorgada conjuntamente N.º 61/539,640, titulada "Fatty Alcohol-Forming Acyl-ACP Reductases", presentada el 27 de septiembre de 2011.

De manera alternativa o además, la cianobacteria recombinante comprende una o más moléculas de ácido nucleico

que codifican una acil-CoA reductasa exógena, una ácido carboxílico reductasa y/o una acil-ACP reductasa y una cera sintasa exógena y puede producir un éster de cera. Los ésteres de cera incluyen una cadena A y una cadena B unidas a través de un enlace éster, uno o ambos de los cuales pueden proceder de un ácido graso generado por las cianobacterias recombinantes de la invención. Los ésteres de cera producidos por las cianobacterias recombinantes de la invención incluyen, por ejemplo, longitudes de cadena A de 8 a 24 carbonos, por ejemplo, de 12 a 18 carbonos o de 12 a 16 carbonos, y/o longitudes de cadena B de 8 a 24 carbonos, por ejemplo, de 12 a 18 carbonos, o de 12 a 16 carbonos. Por ejemplo, los ésteres de cera pueden tener longitudes de cadena A + B que incluyan, pero sin limitación, de 16 a 48 carbonos, de 24 a 36 carbonos o de 24 a 32 carbonos.

Las sintasas de cera incluyen polipéptidos que tienen el número de clasificación de enzimas EC (*enzyme classification*) 2.3.1.75, así como cualquier otro péptido capaz de catalizar la conversión de un acil-tioéster en ésteres grasos, por ejemplo, algunas aciltransferasas, incluyendo algunas DGAT. Algunos péptidos de cera sintasa también pueden catalizar otras reacciones, por ejemplo, para producir ésteres grasos, algunos péptidos de cera sintasa aceptarán las acil-CoA de cadena corta y los alcoholes de cadena corta. En la patente de Estados Unidos N.º 7.118.896, se proporcionan métodos para identificar la actividad cera sintasa. Los ejemplos no limitantes de cera sintasas que una molécula de ácido nucleico exógeno introducida en una cianobacteria recombinante como se describe en el presente documento puede codificar, incluyen, la ester de cera sintasa /acil-CoA:diacilglicerol aciltransferasa bifuncional de *Simmondsia chinensis* (AAD38041), la cera sintasa de *Acinetobacter* sp. cepa ADP 1 (CAG67733), *Pseudomonas aeruginosa* (AAG06717), *Arabidopsis thaliana* (Q93ZR6), *Alcanivorax* (EDX90960), *Rhodococcus opacus* (YP_002782647), *Homo sapiens* (Q6E213), *Mus musculus* (Q6E1M8) o *Petunia x hybrida* (AAZ08051).

En algunas realizaciones, las cianobacterias recombinantes de la invención comprenden al menos una molécula de ácido nucleico que codifica una ácido graso descarboxilasa exógena o una aldehído graso descarboxilasa exógena, o adicionalmente al menos una molécula de ácido nucleico exógeno que codifica una acil-CoA reductasa exógena, ácido carboxílico reductasa, o acil-ACP reductasa, y pueden producir un alcano y/o alqueno. Los alcanos y alquenos producidos por los microorganismos recombinantes o las células hospedadoras de la invención pueden, por ejemplo, tener longitudes de cadena de 7, 9, 11, 13, 15, 17, 19, 21 y/o 23 carbonos, incluyendo, por ejemplo, longitudes de cadena de 7, 9, 11, 13, 15 y/o 17 carbonos, o longitudes de cadena de 7, 9, 11, 13 y/o 15 carbonos, o longitudes de cadena de 11, 13 y/o 15 carbonos.

Además, las cianobacterias recombinantes de la invención que producen un alcohol graso, aldehído graso, éster de ácido graso, éster de cera o hidrocarburos, incluyendo un alcano o un alqueno, pueden incluir, opcionalmente, una molécula de ácido nucleico que codifique una acil-CoA sintetasa exógena, o puede genomodificarse para que tenga una expresión regulada positivamente de un gen endógeno de acil-CoA sintetasa.

Además, opcionalmente, la célula hospedadora recombinante puede genomodificarse para que exprese un transportador transmembrana exógeno para facilitar la secreción de uno o más productos de ácidos grasos. Por ejemplo, la célula hospedadora recombinante puede incluir un gen no originario que codifique un transportador de casete de unión a ATP (ABC, siglas del inglés *ATP-binding cassette*) o una bomba de RND. En algunas realizaciones, la secuencia del transportador es al menos 80 % idéntica a la de una proteína transportadora codificada por los genes de *Arabidopsis*, CER5, WBC11, AtMRPS, AmiS2 y AtPGP1 o por genes transportadores de ácidos grasos (FATP) de especies de los géneros *Saccharomyces, Drosophila*, de micobacterias o de mamíferos.

También se desvelan genes que codifican variantes de estas y otras enzimas de origen natural que participan en la síntesis de productos de ácidos grasos que tienen una identidad de al menos 65 % con las proteínas mencionadas o de origen natural, en donde la actividad de la enzima no está sustancialmente reducida con respecto a la de la enzima de tipo silvestre o mencionada anteriormente.

Las células hospedadoras recombinantes descritas anteriormente pueden utilizarse en cualquiera de los métodos de producción de un producto de ácido graso como se describe en el presente documento.

Moléculas de ácido nucleico

Las moléculas de ácido nucleico y los polipéptidos descritos en el presente documento pueden utilizarse en cualquiera de los métodos de la invención, y pueden incluirse en cualquiera de las cianobacterias recombinantes de la invención. Se proporcionan moléculas de ácido nucleico que comprenden secuencias que codifican las LPAAT para su uso en cianobacterias hospedadoras y métodos para producir productos de ácidos grasos, incluidos los ácidos grasos libres, aldehídos grasos, alcoholes grasos, ésteres de ácidos grasos, ésteres de cera, alcanos y/o alquenos. Una molécula de ácido nucleico como se desvela en el presente documento puede aislarse y/o purificarse.

En algunas realizaciones, la expresión en una cianobacteria hospedadora de una molécula de ácido nucleico recombinante o secuencia que codifica una LPAAT como se describe en el presente documento, da como resultado un nivel de producción más alto de un ácido graso o un derivado de ácido graso por la cianobacteria hospedadora en comparación con el nivel de producción en una cianobacteria de control, donde la cianobacteria de control se cultiva en las mismas condiciones y es sustancialmente idéntica, en todos los aspectos, a la cianobacteria que expresa la

molécula o secuencia de ácido nucleico no originario que codifica la LPAAT, con la excepción de que la cianobacteria de control no expresa una molécula de ácido nucleico no originario que codifique una LPAAT.

En realizaciones adicionales, la expresión en una cianobacteria hospedadora de una secuencia de ácido nucleico no originario que codifica una LPAAT y una secuencia de ácido nucleico no originario que codifica una tioesterasa como se describe en el presente documento, da como resultado un nivel de producción más alto de un producto de ácido graso por la cianobacteria hospedadora en comparación con el nivel de producción en una cianobacteria de control, donde la cianobacteria de control se cultiva en las mismas condiciones y es sustancialmente idéntica, en todos los aspectos, a la cianobacteria que expresa la(s) molécula(s) de ácido nucleico o secuencias recombinantes, con la excepción de que la cianobacteria de control no expresa una tioesterasa y no expresa secuencias de ácido nucleico no originario que codifique una LPAAT.

En algunas realizaciones, la invención se refiere a moléculas de ácido nucleico que codifican mutantes de deleción de una LPAAT donde uno o más aminoácidos se han delecionado de la proteína. Por ejemplo, el polipéptido codificado puede carecer de al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70 u 80 aminoácidos del extremo N y/o C y puede tener una secuencia de aminoácidos al menos 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntica a la secuencia de aminoácidos correspondiente de SEQ ID NO: 2. En algunos ejemplos, las secuencias delecionadas pueden incluir secuencias de direccionamiento, por ejemplo, al menos una parte de un péptido de tránsito al cloroplasto, al menos una parte de una secuencia de direccionamiento mitocondrial, al menos una parte de una secuencia de direccionamiento al retículo endoplasmático, *etc.*

Además, la invención se refiere a moléculas de ácido nucleico que codifican variantes de secuencias de aminoácidos de LPAAT de origen natural, tales como, pero sin limitación, variantes de la secuencia de LPAAT de SEQ ID NO: 2. Las variantes pueden ser de origen natural o no natural, tales como las inducidas por diversos mutágenos y procesos mutagénicos. En algunas realizaciones, una molécula de ácido nucleico codifica una variante de una LPAAT en la que al menos un resto de aminoácido N y/o C terminal se ha insertado en, y/o dentro de, la secuencia de referencia. En algunas realizaciones, al menos un resto de aminoácido N y/o C terminal se ha delecionado de, y/o dentro de, la secuencia de referencia. En algunas divulgaciones, las moléculas de ácido nucleico pueden codificar variantes que pueden ser secuencias que contienen mutaciones predeterminadas, por ejemplo, por recombinación homóloga o mutagénesis dirigida o por PCR; proteínas correspondientes de otras especies; alelos u otras variantes de origen natural; y/o derivados en los que la proteína se ha modificado de manera covalente, a través de medios químicos, enzimáticos u otros medios apropiados, con un residuo distinto de un aminoácido natural.

Una sustitución, inserción o deleción, puede afectar negativamente a la proteína cuando la secuencia alterada inhibe sustancialmente una función biológica asociada a la proteína. En determinados aspectos de la divulgación, una variante de una LPAAT puede tener una actividad que se reduce en no más de aproximadamente 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 30 %, 40 % o 50 %, en comparación con la actividad de la LPAAT de la que procede la variante (por ejemplo, SEQ ID NO: 2). En algunos aspectos de la divulgación, la cantidad de un producto de ácido graso producido por una célula hospedadora que expresa la LPAAT variante no es inferior a aproximadamente 99 %, 98 %, 97 %, 96 %, 95 %, 94 %, 93 %, 92 %, 91 %, 90 %, 85 %, 80 % o 75 % de la cantidad o el producto de ácido graso producido por una célula hospedadora que expresa la LPAAT de la que procede la variante (por ejemplo, SEQ ID NO: 2).

La invención también se refiere a fragmentos y variantes de una LPAAT que tienen una actividad aumentada en comparación con la de los polipéptidos de referencia. En determinados aspectos, el fragmento o la variante de LPAAT puede tener una actividad que está aumentada en al menos un 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 %, 200 %, 300 %, 400 %, 500 %, 600 %, 700 %, 800 %, 900 % o 1000 % en comparación con la actividad de la LPAAT de la que procede la variante. En determinados aspectos, la cantidad de ácido graso o derivado de ácido graso producida por una célula hospedadora que expresa el fragmento o la variante es al menos un 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 %, 200 %, 300 %, 400 %, 500 %, 600 %, 700 %, 800 %, 900 % o 1 000 % de la cantidad de producto de ácido graso que produce una célula hospedadora que expresa la LPAAT de la que procede el fragmento o la variante.

Cualquiera de las moléculas de ácido nucleico desveladas, puede comprender, opcionalmente además, una secuencia de ácido nucleico adicional de al menos 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 700, 800, 900, 1 000 o 1 500 nucleótidos de un organismo fotosintético.

Otras modificaciones

La invención también se refiere a variantes adicionales de las secuencias de nucleótidos de la invención. En algunas realizaciones, las variantes de secuencia de nucleótidos codifican fragmentos o variantes de los polipéptidos como se describe en el presente documento. En algunas realizaciones, las variantes de la secuencia de nucleótidos son de origen natural. En otras realizaciones, las variantes de la secuencia de nucleótidos son de origen no natural, tales como las inducidas por diversos mutágenos y procesos mutagénicos. En determinadas realizaciones, las variantes

de la secuencia de nucleótidos son una combinación de origen natural y no natural. Se puede modificar una secuencia de ácido nucleico dada, por ejemplo, según métodos estándar de mutagénesis o evolución artificial o de intercambio de dominios, para producir secuencias modificadas. Por ejemplo, en Stemmer (1994) Nature 370, 389-391 y Stemmer (1994) Proc. Natl. Acad. Sci. Estados Unidos 91, 10747-10751, se describen métodos de evolución acelerada. La alteración química o enzimática de los ácidos nucleicos y polipéptidos expresados puede realizarse por métodos estándar. Por ejemplo, una secuencia puede modificarse mediante la adición de grupos fosfato, grupos metilo, lípidos, azúcares, péptidos o compuestos orgánicos o inorgánicos, mediante la inclusión de nucleótidos o aminoácidos modificados, o similares.

Para la expresión óptima de una proteína recombinante, en determinados casos, puede ser beneficioso emplear secuencias codificantes que producen ARNm con codones utilizados preferentemente por la célula hospedadora para ser transformados ("optimización de codones"). Por tanto, para mejorar la expresión de transgenes, el uso de codones del transgén puede coincidir con el sesgo de codones específico del organismo en el que se desea que se exprese el transgén. Se cree que los mecanismos exactos subyacentes a este efecto son muchos, pero puede incluir el equilibrio adecuado de los grupos de ARNt aminoacilados disponibles con las proteínas que se sintetizan en la célula, junto con una traducción más eficiente del ARN mensajero (ARNm) transgénico cuando se satisface esta necesidad. En algunas realizaciones, solo una parte de los codones se cambia para reflejar un uso de codones preferido de una cianobacteria hospedadora. En determinadas realizaciones, uno o más codones se cambian por codones que no son necesariamente el codón más preferido del microorganismo hospedador que codifica un aminoácido particular. Se dispone de información adicional sobre la optimización de codones, por ejemplo, en la base de datos de uso de codones de GenBank. Las secuencias codificantes pueden ser secuencias con codones optimizados para la producción óptima de un producto deseado en el organismo hospedador seleccionado para la expresión. En determinadas realizaciones, la secuencia de ácido nucleico no originaria que codifica una LPAAT y/o la secuencia de ácido nucleico no originaria que codifica una tioesterasa es un secuencia con codones optimizados para la expresión en una cianobacteria.

En algunas realizaciones, las moléculas de ácido nucleico de la divulgación codifican proteínas de fusión que comprenden una LPAAT. Por ejemplo, los ácidos nucleicos de la invención pueden comprender secuencias de polinucleótidos que codifiquen glutatión-S-transferasa (GST) o una parte de la misma, tiorredoxina o una parte de la misma, proteína de unión a maltosa o una parte de la misma, poli-histidina (por ejemplo, His_6), poli-HN, poli-lisina, una secuencia de etiqueta de hemaglutinina, HSV-Tag y/o al menos una parte de HIV-Tat fusionada a la secuencia codificante de LPAAT.

En algunas realizaciones, las moléculas de ácido nucleico de la divulgación comprenden secuencias no codificantes adicionales tales como secuencias 3' y 5' no codificantes (que incluyen, por ejemplo, secuencias reguladoras) que pueden ser homólogas o heterólogas al gen LPAAT.

Construcciones de ácido nucleico

La invención también se refiere a construcciones que comprenden una secuencia de ácido nucleico que codifica una LPAAT y/o una tioesterasa que puede incluir además una o más secuencias que regulan o median la transcripción, traducción o integración de secuencias de nucleótidos en un genoma hospedador. Por ejemplo, la invención se refiere a construcciones de expresión que comprenden uno o más "elementos de control de expresión" o secuencias que regulan la transcripción de expresión de un gen unido operativamente, o la traducción del ARN transcrito. Por ejemplo, un elemento de control de la expresión puede ser un promotor que puede estar unido operativamente al gen de interés (por ejemplo, un gen de LPAAT) en una construcción de expresión o "casete de expresión". En algunas realizaciones de lo anterior, el promotor es regulable, por ejemplo, inducible.

En aspectos de la divulgación en los que la construcción de ácido nucleico no contiene un promotor en unión operativa con la secuencia de ácido nucleico que codifica el gen de interés (por ejemplo, un gen de LPAAT), la secuencia de ácido nucleico puede transformarse en las células de tal manera que vuelva a estar en unión operativa con un promotor endógeno, por ejemplo, mediante recombinación homóloga, integración específica de sitio y/o integración vectorial. En algunos aspectos, las secuencias genómicas del hospedador incluidas en una construcción de ácido nucleico para mediar en la recombinación homóloga en el genoma del hospedador, pueden incluir secuencias reguladoras de genes, por ejemplo, una secuencia promotora, que puede regular la expresión de un gen de LPAAT de la construcción de ácido nucleico. En dichos aspectos, el transgén o los transgenes de la construcción pueden unirse operativamente a un promotor que es endógeno al microorganismo hospedador. En algunas realizaciones, el promotor o los promotores endógenos son regulables, por ejemplo, inducibles.

Un promotor unido operativamente a una secuencia de ácido nucleico que codifica una LPAAT puede ser un promotor que sea heterólogo con respecto al gen de LPAAT. En algunas realizaciones de la divulgación anterior, el promotor puede ser un promotor inducible, es decir, un promotor que media en la transcripción de un gen unido operativamente en respuesta a un estímulo particular. Dichos promotores pueden ser ventajosos, por ejemplo, para minimizar cualquier efecto nocivo sobre el crecimiento de la célula hospedadora y/o para maximizar la producción del producto de ácido graso. Un promotor inducible puede ser sensible, por ejemplo, a la luz o a la oscuridad o a la alta o baja temperatura, y/o puede ser sensible a compuestos específicos. El promotor inducible puede ser, un

promotor de ara, un promotor de lac, promotor de trp, un promotor de tet (por ejemplo, patente de Estados Unidos N.º 5.851.796), un promotor híbrido que incluya una parte de un promotor de trp, lac, o tet, un promotor sensible a hormonas (p. ej., un promotor sensible a la ecdisona, tal como se describe en la patente de Estados Unidos N.º 6.379.945), un promotor de metalotionina (por ejemplo, patente de Estados Unidos N.º 6.410.828), un promotor relacionado con la patogénesis (PR) que puede ser sensible a un producto químico tal como, por ejemplo, ácido salicílico, etileno, tiamina y/o BTH (patente de Estados Unidos N.º 5.689.044), o similares, o alguna combinación de los mismos. Un promotor inducible también puede responder a la luz o la oscuridad (patente de Estados Unidos N.º 5.750.385, patente de Estados Unidos N.º 5.639.952), a metales (Eukaryotic Cell 2: 995-1002 (2003)) o a la temperatura (patente de Estados Unidos N.º 5.447.858; Abe et al. Plant Cell Physiol. 49: 625-632 (2008); Shroda et al. Plant J. 21: 121-131 (2000)). La lista anterior es ejemplar y no limitante. La secuencia promotora puede ser de cualquier organismo, siempre que sea funcional en el organismo hospedador. En determinadas realizaciones, los promotores inducibles se forman fusionando una o más partes o dominios de un promotor inducible conocido, con al menos una parte de un promotor diferente que puede intervenir en la célula hospedadora, por ejemplo, para conferir inducibilidad a un promotor que intervenga en la especie hospedadora.

Para la transformación de cianobacterias, se puede utilizar una variedad de promotores que intervengan en las cianobacterias, incluyendo, pero sin limitación, los promotores de ara lac y trc, así como derivados que también son inducibles mediante la adición de isopropil β-D-1-tiogalactopiranósido (IPTG) tal como el promotor de *trcY* o *trcE*. Otros promotores que pueden encontrar uso en la invención incluyen promotores que están naturalmente asociados a genes de resistencia a antibióticos transmitidos por cromosomas bacterianos o transposones (por ejemplo, neomicina fosfotransferasa, cloranfenicol acetiltransferasa, espectinomicina adeniltransferasa, o similares, o combinaciones de los mismos), promotores asociados a diversos genes heterólogos de bacterias y de cianobacterias originarios, promotores de virus y fagos, promotores sintéticos, o similares, o combinaciones de los mismos. Por ejemplo, el promotor, o promotores, pueden seleccionarse de promotores procariotas de una variedad de especies, incluidas especies de eubacterias y cianobacterias, tales como, por ejemplo, un promotor de araC o pBAD, un promotor de rha, un promotor de Pm, un promotor de xylS, un promotor de nir, un promotor de nar, un promotor de pho, un promotor de tet, un promotor de cys, un promotor de metalotioneína, un promotor de ftf, un promotor de gln, un promotor de choque térmico, un promotor inducible por frío o un promotor vírico. Los promotores anteriores son ejemplares y no limitativos. Como promotores aislados de cianobacterias que pueden utilizarse pueden incluirse, pero sin limitación, los siguientes: nrs (inducible por níquel), secA (secreción; controlado por el estado redox de la célula), rbc (operón de Rubisco), psaAB (proteínas del centro de reacción de PS I; reguladas por luz), psbA (proteína D1 de PSII; inducible por luz), y similares, y combinaciones de los mismos. En algunas realizaciones, los promotores están regulados por compuestos de nitrógeno, tales como, por ejemplo, promotores de nar, ntc, nir o nrt. En algunos aspectos de la divulgación, los promotores están regulados por fosfato (por ejemplo, promotores de pho o pst) o metales, por ejemplo, el promotor de nrs (Liu y Curtis (2009) Proc Natl Acad Sciences USA 106: 21550-21554) o el promotor de petE (Buikema y Haselkorn (2001) Proc Natl Acad Sciences USA 98: 2729-2734). Los promotores inducibles, como los utilizados en las construcciones de la presente divulgación, puede utilizar una o más partes o dominios de los promotores mencionados anteriormente y/u otros promotores inducibles fusionados a al menos una parte de un promotor diferente que pueda intervenir en el organismo hospedador, por ejemplo, para conferir inducibilidad a un promotor que intervenga en la especie hospedadora.

Asimismo, para la construcción del vector de expresión puede utilizarse una gran variedad de terminadores transcripcionales. Como ejemplos de posibles terminadores pueden incluirse, pero sin limitación, psbA, psaAB, rbc, secA, proteína de recubrimiento de T7, y similares, y combinaciones de los mismos.

En algunos aspectos de la divulgación, una molécula de ácido nucleico aislada o recombinante de la invención puede comprender tanto una secuencia de ácido nucleico que codifica una LPAAT como una secuencia de ácido nucleico que codifica una tioesterasa. Las secuencias de ácido nucleico que codifican la LPAAT y la tioesterasa pueden ser, por ejemplo, cualquiera de las secuencias de ácido nucleico descritas en el presente documento.

En determinados aspectos de la divulgación, la secuencia de ácido nucleico que codifica una LPAAT y la secuencia de ácido nucleico que codifica una tioesterasa, pueden unirse operativamente al mismo promotor y/o potenciador. Por ejemplo, en aspectos particulares de la divulgación, los dos genes (que codifican una LPAAT y una tioesterasa) pueden organizarse como un operón, en el que, por ejemplo, después de una secuencia promotora viene, en la dirección 5' a 3', una secuencia codificante de tioesterasa y después una secuencia codificante de LPAAT. En una configuración alternativa del operón, después de una secuencia promotora viene, en la dirección 5' a 3', una secuencia codificante de LPAAT y después una secuencia codificante de tioesterasa. En algunos aspectos de la divulgación, una molécula de ácido nucleico aislada puede incluir dos o más genes dispuestos en tándem, donde la molécula de ácido nucleico aislada no incluye una secuencia promotora que intervenga en el microorganismo hospedador previsto cadena arriba de los dos o más genes. En estos aspectos, el operón sin promotor puede diseñarse para su integración (por ejemplo, recombinación homóloga) en un sitio del genoma del hospedador que puede incluir una secuencia promotora, de modo que el operón sintético puede regularse transcripcionalmente gracias a un promotor en el genoma del microorganismo hospedador. Además, el operón puede diseñarse para la integración (por ejemplo, recombinación homóloga) en un sitio del genoma del hospedador que puede incluir una secuencia potenciadora, de modo que el operón introducido puede regularse transcripcionalmente gracias a un potenciador en el genoma del microorganismo hospedador. En cualquiera de los operones desvelados anteriormente

que incluyen genes de LPAAT y de tioesterasa, se pueden incluir una o más secuencias reguladoras adicionales en la molécula de ácido nucleico aislada, por ejemplo, se puede incluir una secuencia para mejorar la traducción cadena arriba de cualquiera de las secuencias codificantes de genes, y/o se puede incluir opcionalmente un terminador transcripcional en o cerca del extremo 3' del operón sintético.

Además de un gen de LPAAT y de un gen de tioesterasa, en un operón sintético, como se proporciona en el presente documento, pueden incluirse uno o más genes adicionales, donde el uno o más genes adicionales pueden incluir, por ejemplo, uno o más genes que codifican enzimas o proteínas de la ruta de síntesis de ácidos grasos y/o uno o más genes que codifican enzimas o proteínas que pueden mejorar la síntesis de productos de ácidos grasos, uno o más genes que pueden mejorar la fotosíntesis o la fijación de carbono, y/o uno o más genes indicadores o marcadores de selección.

La secuencia de ácido nucleico que codifica una LPAAT y la secuencia de ácido nucleico que codifica una tioesterasa pueden proporcionarse en la misma construcción de ácido nucleico donde están unidas operativamente a diferentes promotores y/o potenciadores de la transcripción. Los promotores y potenciadores pueden ser, por ejemplo, cualquiera de los promotores y potenciadores transcripcionales descritos en el presente documento.

El vector que comprende una secuencia de ácido nucleico que codifica una LPAAT está diseñado para la transformación en cianobacterias. El vector puede permitir la recombinación homóloga de la secuencia codificante de LPAAT con el genoma de la cianobacteria.

Una molécula de ácido nucleico aislada de la presente divulgación puede incluir las secuencias desveladas en el presente documento que codifican una LPAAT u otro polipéptido en un vector, tal como, pero sin limitación, un vector de expresión. Un vector puede ser un ácido nucleico que se haya generado mediante intervención humana, incluyendo medios recombinantes y/o por síntesis química directa, y puede incluir, por ejemplo, uno o más de: 1) un origen de replicación para la propagación de las secuencias de ácido nucleico en uno o más hospedadores (que pueden incluir o no el hospedador de producción); 2) uno o más marcadores de selección; 3) uno o más genes indicadores; 4) una o más secuencias de control de expresión, tal como, pero sin limitación, secuencias promotoras, secuencias potenciadoras, secuencias terminadoras, secuencias para mejorar la traducción, *etc.;* y/o 5) una o más secuencias para promover la integración de las secuencias de ácido nucleico en el genoma del hospedador, por ejemplo, una o más secuencias que tienen homología con una o más secuencias de nucleótidos del microorganismo hospedador. Un vector puede ser un vector de expresión que incluya uno o más "elementos de control de expresión" de ácido nucleico especificados que permiten la transcripción y/o traducción de un ácido nucleico particular en una célula hospedadora. El vector puede ser un plásmido, una parte de un plásmido, una construcción vírica, un fragmento de ácido nucleico, o similar, o una combinación de los mismos.

El vector puede ser un vector con un número de copias alto, un vector lanzadera que puede replicarse en más de una especie de célula, un vector de expresión, un vector de integración, o una combinación de los mismos. Normalmente, el vector de expresión puede incluir un ácido nucleico que comprende un gen de interés unido operativamente a un promotor en un "casete de expresión" que también puede incluir, pero sin limitación, un terminador transcripcional, un sitio de unión al ribosoma, un sitio de corte y empalme o una secuencia de reconocimiento de corte y empalme, un intrón, un potenciador, una señal de poliadenilación, un sitio interno de entrada al ribosoma y elementos similares. La presente divulgación puede implicar microorganismos recombinantes transformados con un ácido nucleico aislado que comprende un gen de interés bajo el control de un promotor heterólogo. De manera alternativa, si el vector no contiene un promotor unido operativamente con un ácido nucleico aislado que comprende un gen de interés, el ácido nucleico aislado puede transformarse en los microorganismos o en las células hospedadoras de modo que se una operativamente unido a un promotor endógeno por recombinación homóloga, integración específica de sitio y/o integración vectorial.

La presente invención proporciona adicionalmente cianobacterias recombinantes transformadas con un ácido nucleico aislado que comprende un gen de interés que está unido operativamente a uno o más elementos de control de la expresión. En algunos casos, puede ser ventajoso expresar la proteína a un punto determinado durante el crecimiento del microorganismo recombinante, por ejemplo, para minimizar cualquier efecto nocivo sobre el crecimiento del microorganismo recombinante y/o para maximizar la producción del producto de ácido graso de interés. En dichos casos, uno o más genes exógenos introducidos en la cianobacteria recombinante pueden unirse operativamente a un promotor inducible, que media la transcripción de un gen unido operativamente en respuesta a un estímulo particular.

Los vectores de transformación pueden incluir además o como alternativa, un marcador de selección, tal como, pero sin limitación, un gen de resistencia a fármacos, un gen de resistencia a herbicidas, una enzima metabólica y/o un factor necesario para la supervivencia del hospedador (por ejemplo, un marcador auxotrófico), o similares, o una combinación de los mismos. Las células transformadas pueden seleccionarse en función de la capacidad de crecer en presencia del antibiótico y/u otro marcador de selección en condiciones en las que las células que carecen del casete de resistencia o marcador auxotrófico no podrían crecer. Además, en un vector puede haber un marcador no seleccionable, tal como un gen que codifique una proteína o enzima fluorescente que genere un producto de reacción detectable.

Un vector que comprende un ácido nucleico aislado que comprende un gen de interés, también puede ser un vector de integración que incluya una o más secuencias que promuevan la integración del gen de interés o un casete de expresión génica en el genoma del microorganismo hospedador o de la célula hospedadora. Por ejemplo, un vector de integración utilizado para transformar cianobacterias puede incluir al menos una secuencia de al menos 50, al menos 100, al menos 200, al menos 300, al menos 400, al menos 500 o al menos 600 nucleótidos con homología con una secuencia en el genoma del organismo hospedador para permitir que se produzca la integración del gen de interés o casete de expresión génica en el genoma del microorganismo hospedador o de la célula hospedadora a través de recombinación homóloga. En algunos ejemplos, el gen o casete de expresión génica está flanqueado por secuencias homólogas en una región del cromosoma hospedador para promover la integración del gen de interés o del casete de expresión génica en el cromosoma del hospedador. De manera alternativa o además, un vector de integración puede incluir una o más secuencias que promuevan la recombinación específica de sitio o la integración aleatoria, tales como, pero sin limitación, secuencias reconocidas por recombinasas, integrasas o transposasas. El vector de integración puede incluir además un gen que codifique una recombinasa, integrasa o transposasa.

Transformación de microorganismos y células hospedadoras

Un vector que comprende un ácido nucleico aislado que comprende un gen de interés puede introducirse en las cianobacterias mediante técnicas convencionales de transformación y/o transfección. Los términos "transformación", "transfección", "conjugación", y "transducción", como se usan en el presente contexto, pretenden comprender una multiplicidad de métodos conocidos por los expertos en la técnica para la introducción de ácido nucleico extraño (por ejemplo, ADN exógeno) en una célula hospedadora, incluida la coprecipitación con fosfato de calcio y/o cloruro de calcio, la transfección mediada por DEAE-dextrano, lipofección, competencia natural, transferencia mediada químicamente, electroporación, bombardeo de partículas, o similares, o combinaciones de los mismos. Pueden encontrarse ejemplos de métodos adecuados para la transformación y/o transfección de células hospedadoras, por ejemplo, en Molecular Cloning - A Laboratory Manual (2010), Cold Spring Harbor Laboratory Press.

Las células hospedadoras para su uso en la invención pueden transformarse por cualquier medio factible, incluyendo, sin limitación, el uso de la transformación mediada por pistola de partículas, transformación mediada por láser o electroporación. Las bacterias fotosintéticas pueden transformarse por cualquier método adecuado, incluyendo, como ejemplos no limitantes, absorción de ADN natural (Chung et al. (1998) FEMS Microbiol. Lett. 164: 353-361; Frigaard et al. (2004) Methods Mol. Biol. 274: 325-40; Zang et al. (2007) J. Microbiol. 45: 241-245), conjugación, transducción, transformación con perlas de vidrio (Kindle et al. (1989) J. Cell Biol. 109: 2589-601; Feng et al. (2009) Mol. Biol. Rep. 36: 1433-9; patente de Estados Unidos N.º 5.661.017), transformación con bigotes de carburo de silicio (Dunahay et al. (1997) Methods Mol. Biol. (1997) 62: 503-9), biolística (Dawson et al. (1997) Curr. Microbiol. 35: 356-62; Hallmann et al. (1997) Proc. Natl. Acad. USA 94: 7469-7474; Jakobiak et al. (2004) Protist 155:381-93; Tan et al. (2005) J. Microbiol. 43: 361-365; Steinbrenner et al. (2006) Appl Environ. Microbiol. 72: 7477-7484; Kroth (2007) Methods Mol. Biol. 390: 257-267; patente de Estados Unidos N.º 5.661.017) electroporación (Kjaerulff et al. (1994) Photosynth. Res. 41: 277-283; Iwai et al. (2004) Plant Cell Physiol. 45: 171-5; Ravindran et al. (2006) J. Microbiol. Methods 66: 174-6; Sun et al. (2006) Gene 377: 140-149; Wang et al. (2007) Appl. Microbiol. Biotechnol. 76: 651-657; Chaurasia et al. (2008) J. Microbiol. Methods 73: 133-141; Ludwig et al. (2008) Appl. Microbiol. Biotechnol. 78: 729-35), transformación mediada por láser o incubación con ADN en presencia o después de un tratamiento previo con cualquiera de los dendrímeros de poli(amidoamina) (Pasupathy et al. (2008) Biotechnol. J. 3: 1078-82), polietilenglicol (Ohnuma et al. (2008) Plant Cell Physiol. 49: 117-120), lípidos catiónicos (Muradawa et al. (2008) J. Biosci. Bioeng. 105: 77-80), dextrano, fosfato de calcio o cloruro de calcio (Mendez-Alvarez et al. (1994) J. Bacteriol. 176: 7395-7397), opcionalmente después del tratamiento de las células con enzimas degradadoras de la pared celular (Perrone et al. (1998) Mol. Biol. Cell 9: 3351-3365).

Métodos de producción de productos de ácidos grasos

La invención abarca métodos de producción de un producto de ácido graso cultivando las cianobacterias recombinantes descritas en el presente documento, en condiciones en las que se produce al menos un producto de ácido graso. Los métodos pueden comprender además aislar al menos un producto de ácido graso. En algunas realizaciones, el microorganismo libera o secreta en el medio de crecimiento al menos una parte del ácido graso y/o derivado del ácido graso producido por las cianobacterias recombinantes. En algunas realizaciones, la expresión de un polipéptido codificado por las moléculas de ácido nucleico descritas en el presente documento puede inducirse en la cianobacteria recombinante para producir el producto de ácido graso.

La liberación y la secreción, tal como se utilizan en el presente documento en el contexto de los productos de la invención, se utilizan indistintamente para referirse a un proceso por el cual los mecanismos de transporte activo y/o pasivo permiten que los productos de la invención atraviesen la membrana celular para salir de la célula. Los ejemplos de dichos mecanismos de transporte pueden incluir, pero sin limitación, difusión en función del gradiente, difusión facilitada, transporte activo y combinaciones de los mismos.

El cultivo se refiere a fomentar intencionadamente el crecimiento (por ejemplo, aumentar el tamaño de las células, el contenido celular y/o la actividad celular) y/o la propagación (por ejemplo, aumentar el número de células por

mitosis) de una o más células utilizando condiciones seleccionadas y/o controladas. La combinación tanto del crecimiento como de la propagación puede denominarse proliferación. Los ejemplos no limitantes de condiciones seleccionadas y/o controladas pueden incluir el uso de un medio definido (con características conocidas tales como pH, fuerza iónica y/o fuente de carbono), temperatura, tensión de oxígeno, niveles de dióxido de carbono específicos, crecimiento en un biorreactor, o similar, o combinaciones de los mismos. En algunas realizaciones, la cianobacteria puede crecer en condiciones heterótrofas, utilizando una fuente de carbono reducida, o en condiciones mixotróficas, utilizando tanto luz como una fuente de carbono reducida. Además o como alternativa, la cianobacteria puede cultivarse en condiciones fototróficas. Cuando crece en condiciones fototróficas, la cianobacteria puede usar ventajosamente la luz como fuente de energía. Para la síntesis de biomoléculas por el microorganismo, puede utilizarse una fuente de carbono inorgánica, tal como CO_2 o bicarbonato. "Carbono inorgánico", tal como se usa en el presente documento, incluye compuestos o moléculas que contienen carbono que un organismo no puede utilizar como fuente de energía sostenible. Normalmente, el "carbono inorgánico" puede estar en forma de CO_2 (dióxido de carbono), ácido carbónico, sales de bicarbonato, sales de carbonato, sales de carbonato de hidrógeno, o similares, o combinaciones de los mismos, que no puede oxidarse para obtener energía sostenible ni utilizarse como fuente de energía reductora por parte de los organismos. Si en el medio de cultivo de una cianobacteria que crece en un medio fototrófico se proporciona una molécula o un compuesto de carbono orgánico, generalmente la célula no lo puede absorber y/o metabolizar para obtener energía y/o normalmente no está presente en una cantidad suficiente para proporcionar energía sostenible para el crecimiento del cultivo celular. Sin embargo, las cianobacterias que crecen en un medio heterotrófico pueden utilizar moléculas de carbono orgánico. Por tanto, la presente invención incluye un proceso para convertir una fuente de carbono en un producto de ácido graso que comprende poner en contacto la fuente de carbono con una cianobacteria recombinante de la invención. En algunos aspectos, la fuente de carbono es una fuente de carbono inorgánica y en otros aspectos la fuente de carbono es una fuente de carbono orgánico.

Las cianobacterias que pueden ser útiles de acuerdo con los métodos de la presente invención se pueden encontrar en varios lugares y entornos en todo el mundo. Sin desear vincularse a ninguna teoría, se observa que, quizás como consecuencia de su aislamiento de otras especies y de su divergencia evolutiva, el medio de crecimiento particular para un crecimiento y generación óptimos de lípidos y/u otros constituyentes de hidrocarburos, puede variar. En algunos casos, es posible que determinadas cepas de cianobacterias no sean capaces de crecer en un medio de crecimiento particular debido a la presencia de algún componente inhibidor o a la ausencia de algún requisito nutricional esencial de la cepa de cianobacteria particular.

Generalmente los medios de crecimiento sólidos y líquidos están disponibles en una gran variedad de fuentes, al igual que las instrucciones para la preparación de medios particulares adecuados para una gran variedad de cepas de cianobacterias.

En algunas realizaciones, los medios utilizados para cultivar un organismo que produce ácidos grasos pueden incluir una mayor concentración de un metal (normalmente proporcionado como una sal y/o en forma iónica) tal como, por ejemplo, sodio, potasio, magnesio, calcio, estroncio, bario, berilio, plomo, hierro, níquel, cobalto, estaño, cromo, aluminio, cinc, cobre, o similares, o combinaciones de los mismos (particularmente metales multivalentes, tales como magnesio, calcio y/o hierro), con respecto a una formulación de medio estándar, tal como, por ejemplo, medio BG-11 estándar (Medio 616 ATCC, Tabla 5), o un medio modificado tal como Medio 854 ATCC (BG-11 modificado para contener vitamina B12) o Medio 617 ATCC (BG-11 modificado para cianobacterias marinas, que contiene NaCl adicional y vitamina B12).

Por ejemplo, un medio utilizado para el crecimiento de microorganismos productores de ácidos grasos libres, puede incluir al menos 2 veces, por ejemplo al menos 3 veces, al menos 4 veces, al menos 5 veces, al menos 6 veces, al menos 7 veces, al menos 8 veces, al menos 9 veces, al menos 10 veces, entre 2 veces y 10 veces y/o entre 10 veces y 100 veces la cantidad de metal (por ejemplo, calcio) en comparación con un medio estándar. El medio utilizado para el crecimiento de microorganismos que pueden producir ácidos grasos libres puede incluir, por ejemplo, al menos 0,5 mM, entre aproximadamente 0,5 mM y aproximadamente 1 mM, entre aproximadamente 1 mM y aproximadamente 2 mM, entre aproximadamente 2 mM y aproximadamente 5 mM, entre aproximadamente 5 mM y aproximadamente 10 mM, entre aproximadamente 10 mM y aproximadamente 25 mM, y más de 25 mM de metal (por ejemplo, calcio) en la formulación.

En realizaciones adicionales, utilizando la cantidad en exceso de metal (por ejemplo, calcio) en el medio, al menos una parte del ácido graso o de los ácidos grasos puede quedar retenida como precipitados de jabón, lo que puede provocar la disminución de los efectos tóxicos del ácido graso libre o de los ácidos grasos libres. La adición de metal (por ejemplo, calcio) en el medio, puede aumentar además o como alternativa, la tolerancia de los microorganismos en medios con una concentración relativamente alta de ácidos grasos libres. Además o como alternativa, las cepas productoras de ácidos grasos pueden ser ventajosamente más contundentes con un contenido excesivo de metales (por ejemplo, calcio). Aunque en el presente documento el componente en exceso se describe como un metal, se contempla que el componente pueda describirse más generalmente como una fuente de contraiones de carboxilato, por ejemplo, una fuente de contraiones formadores de jabón, una fuente de iones metálicos (indicada como "metal" en el presente documento), una fuente de contraiones multivalente (es decir, que tiene una valencia de +2 o mayor), una fuente de contraiones divalente, o alguna combinación de las mismas. Otros detalles con respecto a esta fuente de contraiones de metal/carboxilato se describen en la solicitud de patente de Estados Unidos en trámite, otorgada

conjuntamente N.º 13/324 636, titulada "Culturing a Microorganism in a Medium with an Elevated Level of a Carboxylate Counterion Source", presentada el 13 de diciembre de 2011.

Los métodos de cultivo pueden incluir inducir la expresión de un gen particular descrito en el presente documento para la producción de productos de ácidos grasos y/o regular la ruta metabólica en la cianobacteria. Inducir la expresión puede incluir añadir un nutriente o compuesto al cultivo, eliminar uno o más componentes del medio de cultivo, aumentar o disminuir la luz y/o la temperatura, y/u otras manipulaciones que promueven la expresión del gen de interés. Dichas manipulaciones pueden depender en gran medida de la naturaleza del promotor (heterólogo) unido operativamente al gen de interés.

En algunas realizaciones de la presente invención, las cianobacterias recombinantes pueden cultivarse en un biorreactor. Un "biorreactor" se refiere a un recinto o recinto parcial en el que se cultivan células, opcionalmente en suspensión y, cuando es en suspensión, es preferentemente en un líquido acuoso. Los biorreactores pueden ofrecer muchas ventajas para su uso en métodos de crecimiento heterótrofo y propagación. Para producir biomasa para su uso como alimento, las cianobacterias se fermentan preferentemente en grandes cantidades en líquido, tal como, por ejemplo, en cultivos en suspensión. Los biorreactores, tales como fermentadores de acero, pueden dar cabida a volúmenes de cultivo muy grandes (en diversas realizaciones de la invención pueden utilizarse biorreactores de 40 000 litros y de mayor capacidad). Normalmente, los biorreactores también pueden permitir el control de una o más condiciones de cultivo, tales como temperatura, pH, tensión de oxígeno, niveles de dióxido de carbono y similares, así como combinaciones de los mismos. Normalmente, los biorreactores pueden configurarse, por ejemplo, utilizando puertos conectados a conducciones, para permitir que componentes gaseosos, tales como CO_2, aire rico en CO_2, oxígeno y/o nitrógeno, se pongan en contacto (por ejemplo, a través de burbujeo) con un cultivo líquido. Otros parámetros de cultivo, tales como el pH de los medios de cultivo, la identidad y/o la concentración de oligoelementos y/o nutrientes, la identidad y/o concentración de otros componentes de los medios, o similares, o combinaciones de los mismos, normalmente se puede manipular más fácilmente utilizando un biorreactor.

Las cianobacterias pueden cultivarse además o como alternativa en un biorreactor dotado de una fuente de luz artificial, un "fotobiorreactor", y/o puede tener una o más paredes suficientemente transparentes a la luz, incluyendo la luz solar, para permitir, facilitar y/o mantener un crecimiento aceptable de cianobacterias. Para la producción de productos de ácidos grasos, los microorganismos fotosintéticos o las células hospedadoras pueden cultivarse además o como alternativa en matraces de agitación, tubos de ensayo, viales, platos de microtitulación, placas de Petri, o similares, o combinaciones de los mismos.

Además o como alternativa, las cianobacterias recombinantes pueden cultivarse en estanques, canaletas, recipientes de cultivo en el mar, fosas, salinas, canales, o similares, o combinaciones de los mismos. Al igual que con los biorreactores estándar, al cultivo puede suministrarse una fuente de carbono inorgánico (tal como, pero sin limitación, CO_2, bicarbonato, sales de carbonato y similares), incluyendo, pero sin limitación, aire, aire rico en CO_2, gases de combustión, o similares, o combinaciones de los mismos. Cuando se suministran gases de combustión y/o de otras fuentes inorgánicas que puedan contener CO además de CO_2, puede ser necesario tratar previamente dichas fuentes de tal manera que el nivel de CO introducido en el (foto)biorreactor no constituya una dosis peligrosa y/o letal con respecto al crecimiento y/o a la supervivencia de las cianobacterias.

Los métodos incluyen el cultivo de una cianobacteria recombinante que incluya una o más proteínas como se describe en el presente documento para producir al menos un producto de ácido graso, en el que el método da como resultado la producción de al menos 0,1 %, 0,5 %, 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 100 %, 110 %, 120 %, 130 %, 140 %, 150 %, 160 %, 170 %, 180 %, 190 %, 200 %, 300 %, 400 %, 500 %, 600 %, 700 %, 800 %, 900 % o 1 000% más que la cantidad del producto de ácido graso producido por una cianobacteria de otra manera idéntica que no incluye la(s) proteína(s), cultivada en las mismas condiciones. Además o como alternativa, los métodos incluyen producir al menos 100 mg, al menos 110 mg, al menos 120 mg, al menos 130 mg, al menos 140 mg, al menos 150 mg, al menos 160 mg, al menos 170 mg, al menos 180 mg, al menos 190 mg, al menos 200 mg, al menos 210 mg, al menos 220 mg, al menos 230 mg, al menos 240 mg, al menos 250 mg, al menos 260 mg, al menos 270 mg, al menos 280 mg, al menos 290 mg, al menos 300 mg, al menos 310 mg, al menos 320 mg, al menos 330 mg, al menos 340 mg, al menos 350 mg, al menos 360 mg, al menos 370 mg, al menos 380 mg, al menos 390 mg, al menos 400 mg, al menos 450 mg, al menos 500 mg, al menos 550 mg, al menos 600 mg, al menos 650 mg, al menos 700 mg, al menos 750 mg, al menos 800 mg, al menos 850 mg, al menos 900 mg, al menos 950 mg, por litro de cultivo de un producto de ácido graso cultivando las cianobacterias recombinantes descritas en el presente documento. Aunque muchas veces el objetivo puede ser producir y/o recuperar la mayor cantidad posible de producto de ácido graso, en algunos casos, la cantidad del producto de ácido graso producido y/o recuperado por el método descrito en el presente documento puede limitarse a 600 mg o menos por litro de cultivo, por ejemplo a 550 mg o menos por litro de cultivo, o a 500 mg o menos por litro de cultivo.

Los productos de ácidos grasos pueden recuperarse del cultivo por medios de recuperación conocidos por los expertos habituales en la materia, tal como por extracción de cultivo completo, por ejemplo, utilizando disolventes orgánicos. En algunos casos, la recuperación de los productos de ácidos grasos se puede mejorar mediante la homogeneización de las células. Cuando los productos de ácidos grasos se liberan o secretan suficientemente de

los microorganismos al medio de cultivo, el método de recuperación se puede adaptar para recuperar de manera eficaz solo los productos de ácidos grasos liberados, solo los productos de ácidos grasos producidos y almacenados en los microorganismos, o los productos de ácidos grasos tanto producidos como liberados.

Los productos de ácidos grasos secretados/liberados en el medio de cultivo por las cianobacterias recombinantes descritas anteriormente, pueden recuperarse de varias maneras. Puede emplearse, por ejemplo, un método de aislamiento directo, por ejemplo, por división, usando solventes inmiscibles. Además o como alternativa, pueden emplearse adsorbentes particulados. Estos pueden incluir partículas lipófilas y/o resinas de intercambio iónico, dependiendo del diseño del método de recuperación. Pueden circular en el medio separado y después recogerse, y/o el medio puede pasar sobre una columna de lecho fijo, por ejemplo, una columna cromatográfica, que contiene estas partículas. Los productos de ácidos grasos se pueden eluir de los adsorbentes particulados, por ejemplo, usando un disolvente apropiado. En dichas circunstancias, un método de aislamiento puede incluir llevar a cabo la evaporación del disolvente, seguido de un procesamiento adicional de los productos de ácidos grasos aislados, para producir productos químicos y/o combustibles que pueden usarse para una variedad de fines comerciales.

En algunas realizaciones de los métodos descritos en el presente documento, el nivel de un producto de ácido graso, por ejemplo, un ácido graso C8-C24, un ácido graso C12-C24 o un ácido graso C12-C18, tal como, por ejemplo, al menos uno de un ácido graso C12, C14, C16 y/o C18, puede aumentarse en el cultivo con respecto a un cultivo de una cianobacteria de otra manera idéntica, pero sin las modificaciones genéticas de la invención.

Debe entenderse que la divulgación de la presente invención se extiende a métodos, productos y sistemas según los diversos aspectos de la invención que comprenden combinaciones de una o más características analizadas en el presente documento con referencia a determinadas realizaciones de la invención con una o más características adicionales analizadas en el presente documento con referencia a determinadas otras realizaciones de la invención.

Ejemplos

La invención, tal y como se ha descrito anteriormente, puede entenderse fácilmente por referencia a los siguientes ejemplos, que se incluyen con fines ilustrativos de determinados aspectos y realizaciones de la presente invención, y que no pretenden limitar la invención.

Ejemplo 1. Producción de ácidos grasos libres en la inhibición por retroalimentación reducida en cepas de *Synechocystis* sp. PCC6803

Aunque la invención no está ligada a ninguna teoría, los inventores plantearon la hipótesis de que la acil-ACP C18:0, que no es un sustrato preferido de la de acil-ACP tioesterasa de Cc1FatB1 no originaria expresada en la célula, podría acumularse en células productoras de ácidos grasos que expresan la acil-ACP tioesterasa deCc1FatB1 que prefiere C12-C16, posiblemente limitando la producción de ácidos grasos mediante la inhibición por retroalimentación, y que la sobreexpresión de un gen de LPAAT que prefiere C18, podría limitar cualquier efecto inhibidor al reducir el conjunto de acil-ACP C18:0 a través de la condensación, catalizada por LPAAT, de acil-ACP C18:0 con ácido lisofosfatídico para generar ácido fosfatídico. Véase, por ejemplo, la Figura 1. Además de limitar la inhibición por retroalimentación de la síntesis de ácidos grasos, los inventores consideraron que la modificación podría ayudar a estabilizar la membrana celular al promover la biosíntesis de lípidos de la membrana.

1.1. Construcción del vector de LPAAT

Se diseñó una construcción de expresión que incluía un gen codificante de LPAAT con preferencia por las acil-ACP C18:0. El vector pSGI-NB55 (SEQ ID NO: 27; Figura 2) se construyó para sobreexpresar el gen de la Ácido Lisofosfatídico Acil Transferasa (LPAAT) sll1752 (número de registro de Genbank, ID del gen; SEQ ID NO: 1) de *Synechocystis* sp. PCC 6803 utilizando un promotor de trcY (SEQ ID NO: 28). El gen sll1752 se amplificó a partir de ADN genómico de *Synechocystis* utilizando los cebadores NB395 (SEQ ID NO: 29) y NB396 (SEQ ID NO: 30) que contenían una superposición de secuencia de aproximadamente 15 pb con el vector de integración de cianobacterias pSGI-YC63 (SEQ ID NO: 31). El vector pSGI-YC63 contenía un marcador de espectinomicina para la selección, brazos "RS2 ascendente" (SEQ ID NO: 32) y "RS2 descendente" (SEQ ID NO: 33) homólogos para la integración en el sitio RS2 del genoma de *Synechocystis* sp. PCC 6803, el represor lacIQ para el promotor de trcY (SEQ ID NO: 28) para impulsar la expresión de sll1752, y un origen de replicación de pUC para la propagación del vector en *E. coli.* El vector pSGI-YC63 (SEQ ID NO: 31) se amplificó por PCR para producir una molécula de ácido nucleico linealizada utilizando los cebadores NB393 (SEQ ID NO: 34) y NB394 (SEQ ID NO: 35). El producto de PCR purificado para el esqueleto del vector y el gen sll1752 se combinaron en *E. coli* utilizando el kit de clonación por PCR de Quick y se propagó en células Alpha-Select Gold de *E. coli*.

1.2. Construcción del vector de tioesterasa de Cc1FatB1. Un gen de tioesterasa de la planta superior *Cuphea carthagenensis,* Cc1FatB1 (US 2009/0298143; WO 2009/076559), se clonó en un vector diseñado para promover la integración génica en el sitio RS1 del genoma de *Synechocystis* sp. PCC 6803.

La construcción de acil-ACP tioesterasa pSGI-NB158 (SEQ ID NO: 36; Figura 3) contenía un origen de replicación

de P15A para la propagación en *E. coli,* amplificado a partir del vector pACYC184 usando los cebadores NB470 (SEQ ID NO: 37) y NB471 (SEQ ID NO: 38); un marcador de resistencia a kanamicina para la selección, amplificado a partir del vector de integración de cianobacterias SGI-YC28 usando los cebadores NB466 (SEQ ID NO: 39) y NB467 (SEQ ID NO: 40), fragmentos "RS1 ascendente" (SEQ ID NO: 41) y "RS1 descendente" (SEQ ID NO: 42) para la recombinación homóloga en *Synechocystis* 6803, amplificados usando los cebadores NB466 y NB467, respectivamente; el promotor de trcY (SEQ ID NO: 28) y un represor de lacIQ para el gen de tioesterasa de Cc1FatB1 impulsado por promotor de trcY de *Cuphea* (SEQ ID NO: 43) que codifica una acil-ACP tioesterasa truncada en el extremo N (SEQ ID NO: 44) que tiene una preferencia de sustrato por las acil-ACP C12-C16. Para amplificar el casete de tioesterasa trcY-Cc1FatB1 a partir de un vector de expresión anterior, se utilizaron los cebadores NB 462 (SEQ ID NO: 45) y NB 463 (SEQ ID NO: 46). El fragmento IQ de lac se amplificó utilizando los cebadores NB464 (SEQ ID NO: 47) y NB 465 (SEQ ID NO: 48) de otro vector, y el gen de resistencia a kanamicina se amplificó utilizando los cebadores NB466 (SEQ ID NO: 49) y NB467 (SEQ ID NO: 50) del vector pSGI-YC28 (véase la Tabla 1). Los cebadores se diseñaron para que incluyesen secuencias que se incorporasen en los extremos terminales de los productos finales de la PCR de tal manera que los fragmentos amplificados tuviesen una homología terminal de 10-20 pb los con fragmentos de ácido nucleico a clonar adyacentes a los fragmentos. La clonación se realizó utilizando el kit de clonación por PCR de Quick (BPS Biosciences, San Diego). Los clones se transformaron y propagaron en células Alpha-Select Gold competentes *E. coli* (Bioline, Tauton).

Tabla 1. Cebadores utilizados en la construcción de vectores

NB460	GTTTATCACAGTTAAATTATTGCTGAAGCGGAATCCCTGG	SEQ ID NO:58
NB461	CAGCCTCAGGCCATATAACCATCAAAGCCATAGT TGG C	SEQ ID NO:59
NB462	CTTTGATGGTTATATGGCCTGAGGCTGAAATGAGCTGTTG AC	SEQ ID NO:45
NB463	GATCTGAAGCTTGAAAGGTTATTAACTGGCGGGATTACCA TTACTGG	SEQ ID NO:46
NB464	GTAATCCCGCCAGTTAATAACCTTTCAAGCTTCAGATCAA TTCGCGC	SEQ ID NO:47
NB465	ATATTTTTATCTTGTGGTCCGGAATTCGCCCTTCTAAGC	SEQ ID NO:48
NB466	GGGCGAATTCCGGACCACAAGATAAAAATATATCATCAT GAACAATAAAACTGTCTG	SEQ ID NO:49
NB467	GGTGCCATCCATACCGGGCCGCCGTCCCGTCAAG	SEQ ID NO:50
NB468	GGGACGGCGGCCCGGTATGGATGGCACCGATG	SEQ ID NO:60
NB469	CCGTCAGTAGCTGAACATGGGGGACCATTCTCTGGATCAT TG	SEQ ID NO:64
NB470	AGAGAATGGTCCCCCATGTTCAGCTACTGACGGGGTG	SEQ ID NO:37
NB471	GGATTCCGCTTCAGCAATAATTTAACTGTGATAAACTACC GCATTAAAGCT TATCG	SEQ ID NO:38
NB1	CTCGAGCCCCCGTGCTATGACTAGC	SEQ ID NO:52
NB9	CTCGAGCCCGGAACGTTTTTTGTACCCC	SEQ ID NO:53
NB87	CAATTGGCATGCACCCCTATTTGTTTATTTTTCTAAATAC	SEQ ID NO:61
NB88	CAATTGGTTTAAACGGCCTAAGCTTTATGCTTGTAAACCG TTTTGTG	SEQ ID NO:62
NB5	GATCGACGTATAAACTGTCAACATATTTCTGCAAG	SEQ ID NO:63
NB6	CTCGAGTGTGGTGCCGGAGGTGTAG	SEQ ID NO:65
NB393	TTCTAGATATCTGCAGGCCTAAGCTTTATGC	SEQ ID NO:34
NB394	TTTTTTTCCTCCTTAGTGTGAAATTGTTATCCGC	SEQ ID NO:35
NB395	CACTAAGGAGGAAAAAAAATGACCAATTCTCCCCTGGCG	SEQ ID NO:29
NB396	CCTGCAGATATCTAGAATCACGAAGCGGCGATCG	SEQ ID NO:30

1.3. Construcción del vector de desactivación de acil-ACP sintetasa

El vector pSGI-NB19 (SEQ ID NO:51) se utilizó para desactivar el gen de acil-ACP sintetasa slr1609 de *Synechocystis* sp. PCC 6803. Para amplificar el fragmento del gen slr1609 de *Synechocystis* sp. PCC 6803, se utilizaron los cebadores NB1 (SEQ ID NO:52) y NB9 (SEQ ID NO:53). Este fragmento se ligó al vector TOPO PCR2.1 (Life Technologies, Carlsbad, CA) siguiendo el protocolo del fabricante, y la construcción del plásmido se propagó en células TOP10 de *E. coli* (Life Technologies, Carlsbad, CA). Este plásmido se cortó con una enzima de restricción Mfe1 (New England Biolabs, Ipswich, MA). El fragmento GmR se amplificó por PCR a partir de un plásmido pAW4 con los cebadores NB87 y NB88 que contenían sitios de restricción Mfe1 utilizados para cortar los extremos del fragmento de la PCR. Después, esto se ligó en el vector TOPO que contenía el fragmento slr1609 de la etapa anterior para crear el plásmido pSGI-NB19 (SEQ ID NO:51).

1.4. Producción de ácidos grasos libres

Synechocystis sp. PCC 6803 se transformó con la construcción de expresión génica pSGI-NB55 LPAAT (SEQ ID NO:27; Figura 2), la construcción de expresión pSGI-NB158 acil-ACP tioesterasa (SEQ ID NO:36; Figura 3) o la construcción pSGI-NB19 con el gen de la acil-ACP sintetasa atenuado (AAS KO, siglas del inglés *Acyl-ACP Synthetase Knock Out*) (SEQ ID NO:51; Figura 4) o las combinaciones de pSGI-NB55 y pSGI-NB158 o pSGI-NB55 y pSGI-NB19 esencialmente de acuerdo con Zang et al. (2007) J. Microbiology 45: 241-245. Resumiendo, las células crecieron con luz constante hasta alcanzar una densidad óptica de 730 (DO_{730}) de aproximadamente 0,7 a 0,9 (una DO_{730} de ~ 0,25 corresponde a ~ 1 x 10^8 células/ml) y se recogieron por centrifugación a ~ 2 000 g durante ~ 15 minutos a temperatura ambiente (~ 20-25°°C). El sedimento celular se resuspendió en aproximadamente 0,3 veces el volumen de medio de crecimiento BG-11 reciente y se utilizó inmediatamente para la transformación. Aproximadamente, se añadió 1 microgramo de ADN plasmídico (que contenía la construcción pSGI-NB55 (SEQ ID NO: 27) que incluía un gen de LPAAT, o la construcción pSGI-NB158 acil-ACP tioesterasa (SEQ ID NO: 36) o la construcción pSGI-NB19 con AAS desactivado (AAS KO) (SEQ ID NO: 51)) a ~ 0,3 ml de células, se mezcló cuidadosamente y se incubó con iluminación durante aproximadamente 5 horas a ~ 30°°C sin agitación. Las células se extendieron sobre un filtro (membrana de policarbonato Track-Etched de Whatmann Nuclepore, PC ~ 47 mm, ~ 0,2 micrómetros) colocado en placas de agar BG-11 de ~ 50 ml y se dejó recuperar con luz durante aproximadamente 16 a 24 horas, después de esto, para seleccionar transformantes, el filtro se levantó y se colocó en una placa BG-11 nueva que contenía espectinomicina (10 µg/ml) para la selección de la construcción de expresión pSGI-NB55de LPAAT (SEC ID NO: 27), kanamicina (10 µg/ml) para la selección de la construcción de expresión pSGI-NB158 de acil-ACP tioestearasa (SEQ ID NO: 36) o gentamicina (5 µg/ml) para la selección de la construcción pSGI-NB19 con el gen de acil-ACP sintetasa desactivado (SEQ ID NO: 51). Las colonias resultantes se exploraron adicionalmente mediante PCR para detectar la presencia de los genes de tioesterasa, utilizando los cebadores que se utilizaron para generar los fragmentos de genes.

Tabla 2. Medio BG-11

$NaNO_3$	1,5 g	
K_2HPO_4	0,04 g	
$MgSO_4$ * $7H_2O$	0,075 g	
$CaCl_2$ * $2H_2O$	0,036 g	
Ácido cítrico	6,0 mg	
Citrato de amonio férrico	6,0 mg	
EDTA	1,0 mg	
Na_2CO_3	0,02 g	
#Mezcla de Metales Traza A5	1,0 ml	
Agar (si fuese necesario)	(hasta) 10,0 g	
Agua destilada	1,0 l	
#Mezcla de Metales Traza A5	H_3BO_3	2,86 g
	$MnCl_2$ * $4H_2O$	1,81 g
	$ZnSO_4$ * $7H_2O$	0,22 g
	Na_2MoO_4 * $2H_2O$	0,39 g
	$CuSO_4$ * $5H_2O$	0,080 g
	$Co(NO_3)_2$ * $6H_2O$	49,4 mg
	Agua destilada	hasta 1,0 l

Para la producción de ácidos grasos, parches de células de *Synechocystis* transformadas con las construcciones deseadas (en función de la selección por tamaño utilizando cebadores específicos para la exploración por PCR) se rasparon y se hicieron crecer en 25 ml de medio BG-11 (ATCC 616, como se muestra en la Tabla 2; los pesos de los componentes son aproximados; pH final 7,1; se esterilizaron en autoclave a aproximadamente 121°°C durante aproximadamente 15 minutos) con antibióticos apropiados. Los cultivos líquidos se hicieron crecer en matraces de agitación a 30º C, ~ 65 µmol/m^2/s luz, a aproximadamente 215 rpm con un aporte de CO_2 de ~ 5 % hasta la turbidez

(aproximadamente 3 días). La D.O. (730 nm) de cada cultivo, se midió para determinar el volumen de medio BG-11 para resuspender las células para cultivos de producción. Se realizaron resuspensiones para que todos los cultivos tuviesen una D.O. de 1,1 o superior (730 nm). Para garantizar que se alcanzara una D.O. inicial deseada de 1,1 (730 nm) para la inducción, los cultivos se centrifugaron a 2.000 g durante 10 minutos (20-25º C) y se resuspendieron en 9-10 ml de medio BG-11 complementado con 10 mg/ml de BSA. Las lecturas de O.D. se midieron nuevamente después de la resuspensión y se diluyeron hasta alcanzar una D.O. de 1,1 (730 nm), de nuevo en medio BG-11 complementado con 10 mg/ml de BSA. La complementación con BSA redujo la toxicidad de los ácidos grasos que podría causar el blanqueamiento de los cultivos de producción. Se añadieron antibióticos apropiados a los cultivos y el día 1 las células se indujeron con IPTG en viales de vidrio de rosca de tornillo de ~ 4 ml con cinta permeable a los gases para sellar, crecieron a aproximadamente 30º C, ~ 65 µmol/m^2/s luz, agitando a aproximadamente 215 rpm con un aporte de CO_2.de ~ 5 %. El medio BG-11 no proporciona una fuente de carbono reducida que pueda usarse como fuente de energía para el crecimiento de las células. En cambio, las células crecieron de manera fotoautótrofa utilizando CO_2 como sustancialmente la única fuente de carbono, utilizando la luz como fuente de energía e incorporando carbono del CO_2 en biomoléculas, incluyendo ácidos grasos. Para inducir la producción de ácidos grasos libres, a todas las muestras se añadió una concentración final de ~ 1 mM de IPTG. Los viales completos se enviaron para análisis de ácidos grasos mediante CG.

1.5. Análisis de ácidos grasos libres

Los ácidos grasos libres se analizaron por cromatografía de gases con detección de ionización de llama (CG-DIL). Se añadieron aproximadamente 1,0 ml de los cultivos de *Synechocystis* a viales de vidrio de CG de ~ 4 ml con tapones recubiertos con PTFE (National Scientific). a la muestra de cultivo, se añadieron aproximadamente ochenta y cuatro microlitros de un conjunto estándar interno (E.I.) que incluía los ácidos grasos libres C9:0, C13:0 y C17:0, cada uno a aproximadamente 600 µg/ml en hexano, seguido de aproximadamente 83 microlitros de H_2SO_4 a ~ 50 %, aproximadamente 167 microlitros de NaCl 5 M y aproximadamente 1,4 mililitros de hexano. La concentración final de cada E.I. fue de aproximadamente 50 µg/ml. Para preparar el conjunto estándar interno, los ácidos grasos se adquirieron en Fluka o Nu Chek Prep. Después, los cultivos se sometieron a agitación vorticial en un agitador vorticial multitubo a aproximadamente 2 500 rpm durante aproximadamente 30 minutos. Los viales se centrifugaron finalmente durante aproximadamente 3 minutos a aproximadamente 2 000 rpm, para proporcionar una buena separación entre las fases orgánica y acuosa. La capa de hexano se analizó por cromatografía de gases (CG).

Las muestras de ácido graso de *Synechocystis* se analizaron en un CG/DIL de Agilent modelo 7890A que incluía una columna capilar DB-FFAP de J&W Scientific (de 10 m de longitud, 0,10 mm de diámetro interno, 0,10 µm de espesor de película). Para el análisis de muestras de cianobacterias, el horno de CG se programó de la siguiente manera: aproximadamente 120º C durante aproximadamente 0,1 minutos, después se calentó de aproximadamente 40º C/minuto a aproximadamente 240°°C (con 3,5 minutos de espera). La temperatura del inyector se mantuvo a aproximadamente 260º C, y se utilizó una inyección de 0,9 µl dividida a 30:1. Como gas portador se utilizó Helio a un caudal de aproximadamente 0,599 ml/minuto. Los analitos se identificaron comparando los tiempos de retención con patrones inyectados individualmente. El intervalo de calibración para los analitos fue de aproximadamente 2 µg/ml a aproximadamente 200 µg/ml para los ácidos grasos C8:0 - C16:1 y de aproximadamente 0,625 µg/ml a aproximadamente 50 µg/ml para los ácidos grasos C18:0-C18:3.

Como se muestra en la Figura 5, para la cepa de combinación que contenía tioesterasa y LPAAT (sll1752) de Cc1FatB1, se detectó un aumento de 60 % en la producción total de FFA sobre la cepa que contenía tioesterasa Cc1FatB1 (Control 1A). Sin embargo, la sobreexpresión de sll 1752 en solitario, produjo una disminución de más de 3 veces en la producción total de FFA en comparación con el Control 1A. En cepas con un gen de acil-ACP sintetasa atenuado (AASKO, siglas del inglés *attenuated acyl-ACP synthetase gene*) o con una combinación de sll1752 y AASKO, se observó una disminución similar en la producción total de FFA en comparación con el Control 1A.

La figura 6 muestra las longitudes de cadena de los ácidos grasos producidos por las cepas que expresan la acil-ACP tioesterasa Cc1FatB1 con y sin expresión de sll1752 de LPAAT. La preferencia de sustrato para C14 y C16, así como para los sustratos de acilo C12 se ve claramente en el control 1A y se mantiene en la cepa que coexpresó el gen de LPAAT que prefiere C18; sin embargo, el rendimiento de los ácidos grasos libres aumenta significativamente.

LISTADO DE SECUENCIAS

<110> ExxonMobil Research and Engineering Company

<120> PRODUCCIÓN MEJORADA DE ÁCIDOS GRASOS Y DERIVADOS DE ÁCIDOS GRASOS POR MICROORGANISMOS RECOMBINANTES

<130> 2012EM036PCT / 16244-000011/WO/PCT

<160> 57

<170> PatentIn versión 3.5

<210> 1
<211> 723
<212> ADN
<213> Synechocystis PCC6803

<400> 1

atgaccaatt ctcccctggc gatcgaccct gctgtcattg gcccttccag ggtttccccc 60

tggctaatca aattaattta tcccctaggc accaggtttt tacattggta ttttggcccg 120

atcgccatcc atggtcagga gcatttaccc cgcagtggcc cgattattct ggcccctacc 180

caccgttccc gttgggatgc aattttgctt tctttggccg ccggtcgggg ggtaactggt 240

cgagacctcc gttttatggt ggcggtgacg gaagtgcagg gtttacaggg ttggtttatt 300

cgccatttgg ggggatttcc cgttgacgtg aaaaggccgg aaatcagtag tgtgagttat 360

agtgtgcagt tactccaagc aggggaaatg ttggtcattt ttcctgaggg gggcattttt 420

cgggatcaac acactgtcca tcccctcaaa cggggtattg ggcgcattgc catggaagtt 480

tgtaagcaaa accccaacac tgacattaaa gtgatcccag tgacgatcgc ctacagtgac 540

ccctacccag gcaaaggcac ttcagtggaa attaattttg gccagggcat cgccgctagg 600

gactacgacc ccagcaccat caaggctagc tcccaaaaac tgacccgttg tttggcccat 660

agaatgcaga acctttacag ttccgagaat gtctctctgt gcgaggcgat cgccgcttcg 720

tga 723

<210> 2
<211> 240
<212> PRT
<213> Synechocystis PCC6803

<400> 2

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

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

Phe Leu His Trp Tyr Phe Gly Pro Ile Ala Ile His Gly Gln Glu His
35 40 45

```
Leu Pro Arg Ser Gly Pro Ile Ile Leu Ala Pro Thr His Arg Ser Arg
    50                  55                  60

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

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

Gly Trp Phe Ile Arg His Leu Gly Gly Phe Pro Val Asp Val Lys Arg
            100                 105                 110

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

Glu Met Leu Val Ile Phe Pro Glu Gly Gly Ile Phe Arg Asp Gln His
    130                 135                 140

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

Cys Lys Gln Asn Pro Asn Thr Asp Ile Lys Val Ile Pro Val Thr Ile
                165                 170                 175

Ala Tyr Ser Asp Pro Tyr Pro Gly Lys Gly Thr Ser Val Glu Ile Asn
            180                 185                 190

Phe Gly Gln Gly Ile Ala Ala Arg Asp Tyr Asp Pro Ser Thr Ile Lys
        195                 200                 205

Ala Ser Ser Gln Lys Leu Thr Arg Cys Leu Ala His Arg Met Gln Asn
    210                 215                 220

Leu Tyr Ser Ser Glu Asn Val Ser Leu Cys Glu Ala Ile Ala Ala Ser
225                 230                 235                 240
```

<210> 3
<211> 236
<212> PRT
<213> Cyanothece PCC 8801

<400> 3

```
 Met Asn Gln Ser Thr Thr Ser Asn Ser Ile Glu Ser Ser Ile Asn Pro
 1               5                  10                  15

 Trp Leu Ile Arg Leu Val Tyr Pro Leu Gly Cys Ser Leu Ile Met Pro
             20                  25                  30

Leu Phe Phe Gly Arg Ile Thr Ile Ser Gly Gln Glu Asn Ile Pro Thr
        35                  40                  45

Thr Gly Pro Val Ile Val Ala Pro Thr His Arg Ser Arg Trp Asp Ala
    50                  55                  60

Leu Ile Val Pro Gln Ala Val Gly Arg Leu Val Ser Gly Arg Asp Leu
65                  70                  75                  80

Arg Phe Met Val Met Ser Thr Glu Met Thr Gly Leu Gln Gly Trp Leu
                85                  90                  95

Ile Arg Arg Leu Gly Gly Phe Pro Ile Asp Val Lys Arg Pro Gly Leu
            100                 105                 110

Asp Ser Leu Glu His Ser Val Ser Ile Leu Lys Gln Gly Glu Met Leu
        115                 120                 125

Val Ile Phe Pro Glu Gly Gly Ile Phe Arg Asp Asn His Val His Pro
    130                 135                 140

Leu Lys Arg Gly Val Ala Arg Ile Ala Leu Glu Val Val Ser Gln Gln
145                 150                 155                 160

Pro Asn Ser Gly Ile Lys Ile Leu Pro Val Ser Val Gln Tyr Thr Gln
                165                 170                 175

Pro Tyr Pro Ala Trp Gly Thr Asp Val Ile Val Asn Ile Gly Gln Pro
            180                 185                 190

Ile Asp Val Ala Lys Tyr Asn Pro Gln Arg Met Lys Ser Ser Ser Glu
        195                 200                 205

Ala Leu Thr Asn Asp Leu Glu Ala Lys Leu Lys Asp Leu Tyr Gln Gly
    210                 215                 220

Gln Lys Ala Leu Asn Leu Ala Glu Phe Val Gly Ser
225                 230                 235
```

<210> 4
<211> 251
<212> PRT
<213> Crocosphaera watsonii WH 8501

<400> 4

Met Cys Asn Leu His Arg Glu Ile Val Phe Val Ala Phe Cys Leu Lys
1 5 10 15

```
Lys Met Asn Ile His Lys Asn Lys Ser Asp Arg Val Lys Ser Gln Ile
            20                  25                  30

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

Leu Pro Leu Tyr Phe Gly Ser Ile Thr Ile Asn Gly Gln Glu Asn Ile
    50                  55                  60

Pro Thr Ser Gly Pro Ile Leu Ile Ala Pro Thr His Arg Ser Arg Trp
65                  70                  75                  80

Asp Ala Leu Ile Ile Pro Tyr Ala Val Gly Arg Ile Val Ser Gly Arg
                85                  90                  95

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

Trp Phe Ile Arg Arg Met Gly Gly Phe Pro Val Asp Leu Lys Arg Pro
        115                 120                 125

Gly Ala Ser Ser Leu Glu His Ser Val Glu Ile Leu Lys Gln Gly Glu
    130                 135                 140

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

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

Gln Gln Pro Gly Cys Gly Met Lys Ile Leu Pro Val Ser Ile Glu Tyr
            180                 185                 190

Thr Lys Pro Phe Pro Ser Trp Gly Thr His Ile Thr Val Asn Ile Gly
        195                 200                 205

Cys Ser Leu Asp Val Ala Ser Tyr Asp Thr Thr Thr Leu Lys Arg Ser
    210                 215                 220

Ser Gln Lys Leu Thr Gln Asp Leu Glu Ser His Leu Lys Trp Ile Phe
225                 230                 235                 240

Gln Gly Lys Leu Glu Gln Ile Val Ile Gln Lys
                245                 250
```

<210> 5
<211> 234
<212> PRT
<213> Cyanothece CCY0110

<400> 5

```
Met Asn Ile Ser Thr Ile Lys Ser Asp Gln Val Lys Ser Arg Ile Ser
1               5                  10                  15

Pro Trp Leu Ile Arg Ile Thr Tyr Pro Leu Gly Ser Trp Leu Val Leu
            20                  25                  30

Pro Leu Tyr Phe Gly Arg Ile Lys Ile Thr Gly Gln Glu Asn Ile Pro
        35                  40                  45

Asp Thr Gly Pro Val Ile Ile Ala Pro Thr His Arg Ser Arg Trp Asp
    50                  55                  60

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

Val Arg Phe Met Val Thr Ser Ser Glu Met Glu Gly Ile Gln Gly Trp
                85                  90                  95

Phe Ile Arg Arg Leu Gly Gly Phe Pro Val Asp Val Lys Arg Pro Gly
            100                 105                 110

Pro Ser Ser Leu Glu His Ser Ile Glu Val Leu Lys Gln Gly Glu Met
        115                 120                 125

Leu Val Ile Phe Pro Glu Gly Gly Ile Phe Arg Asp Thr Glu Val His
    130                 135                 140

Pro Leu Lys Arg Gly Val Ala Arg Ile Ala Leu Asp Val Glu Ser Gln
145                 150                 155                 160

Gln Pro Gly Cys Gly Met Lys Ile Leu Pro Ile Gly Ile Glu Tyr Ser
                165                 170                 175

Gln Pro Phe Pro Ser Trp Gly Thr Asp Val Thr Val Asn Ile Gly Ser
            180                 185                 190

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

Glu Lys Leu Thr Ile Asp Leu Glu Ser Asp Leu Lys Gln Val Phe His
    210                 215                 220

His Gln Ser Glu Ala Ile Val Phe Gln Ser
                225                 230
```

<210> 6

<211> 236
<212> PRT
<213> Microcystis aeruginosa

<400> 6

```
Met Ser Gln Ser Thr Ile Val Asp Glu Ser Leu Thr Ile Arg Lys Ser
1               5                   10                  15

Pro Ala Val Asn Ser Arg Ile Asn Pro Cys Leu Ile Arg Phe Cys Tyr
            20                  25                  30

Phe Leu Gly Asn Tyr Leu Val Leu Pro Ala Phe Phe Ser Lys Ile Thr
        35                  40                  45

Val Thr Gly Gln Glu Asn Ile Pro Leu Thr Gly Gly Val Ile Leu Ala
    50                  55                  60

Pro Thr His Arg Ser Arg Trp Asp Ala Leu Ile Leu Pro Tyr Ala Val
65                  70                  75                  80

Gly Arg Leu Val Ser Gly Arg Asp Leu Arg Phe Met Val Ser Ala Asn
                85                  90                  95

Glu Ile Lys Gly Val Gln Gly Trp Phe Ile Arg Arg Leu Gly Gly Phe
            100                 105                 110

Pro Val Asn Thr Asp His Pro Gly Met Gly Ser Leu Val His Ser Val
        115                 120                 125

Glu Leu Leu Ala Ala Gly Glu Met Val Ala Ile Phe Pro Glu Gly Gly
    130                 135                 140

Ile Cys Arg Asp Arg Val Val His Pro Leu Lys Pro Gly Val Ala Arg
145                 150                 155                 160

Ile Ala Leu Glu Val Lys Thr Met Lys Pro Asp Ala Asp Ile Lys Ile
                165                 170                 175

Leu Pro Val Ser Ile Ser Tyr Asn Gln Pro Tyr Pro Gly Trp Gly Ser
            180                 185                 190

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

Gln Asn Ser Leu Lys Arg Glu Thr Val Arg Leu Thr Lys Thr Leu Ser
```

```
                210                 215                 220

Asp Arg Leu Lys Leu Leu His Glu Asp Gln Ser Pro
225                 230                 235
```

<210> 7
<211> 287
<212> PRT
<213> Nostoc PCC 7120

<400> 7

```
Met Ile Glu Phe Phe Ser Val Ser Asp Thr His Gln Thr Gln Pro Thr
1               5                  10                  15

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

Pro Trp Leu Ser Pro Val Leu Tyr Phe Val Gly Asn Tyr Phe Leu Leu
        35                  40                  45

Pro Ser Phe Phe Gly Arg Ile Ser Ile Thr Gly Gln Glu Asn Ile Pro
    50                  55                  60

Lys Thr Gly Pro Val Ile Leu Ala Pro Thr His Arg Ala Arg Trp Asp
65                  70                  75                  80

Ser Leu Leu Leu Pro Tyr Ala Thr Gly Arg Tyr Val Thr Gly Arg Asp
                85                  90                  95

Leu Lys Phe Met Val Thr Lys Thr Glu Cys Arg Gly Leu Gln Gly Trp
            100                 105                 110

Phe Val Arg Arg Met Gly Gly Phe Pro Ile Asp Thr Gln His Pro Ala
        115                 120                 125

Val Ser Thr Leu Arg His Ala Val Glu Leu Leu Gln Gln Gly Gln Met
    130                 135                 140

Leu Val Ile Tyr Pro Glu Gly Asn Ile Phe Arg Asp Gly Lys Leu His
145                 150                 155                 160

Pro Leu Lys Ser Gly Ile Ser Arg Leu Ala Leu Ser Ala Glu Ser Ser
                165                 170                 175

His Pro Gly Leu Gly Val Lys Ile Leu Pro Ile Ser Ile Asn Tyr Ser
            180                 185                 190

Gln Pro Tyr Pro Cys Trp Gly Thr Asp Val Ser Ile His Ile Gly Ser
```

```
          195                 200                 205

Pro Ile Asn Val Gln Asp Tyr Thr Lys Gly Lys Val Lys Gln Asn Ala
    210                 215                 220

Lys Arg Leu Thr Glu Asp Leu Ala Arg Asp Leu Gln Ser Leu Ser Asn
225                 230                 235                 240

Pro Glu Leu Ala Phe Ser His His Asp Ala Ala Glu Val Ser Thr Ser
                245                 250                 255

Pro Ile Val Ser Val Ser Asp Ser Thr Glu Arg Ser Pro Asn Val Arg
            260                 265                 270

Gly Gly Ala Thr Ser Arg Lys Arg Glu Leu Leu Thr Ile Asp Tyr
        275                 280                 285
```

<210> 8
<211> 309
<212> PRT
<213> Anabaena variabilis

<400> 8

```
Met Ile Glu Phe Phe Ser Val Ser Asp Thr Arg Gln Thr Gln Pro Thr
1               5                  10                  15

Asn Thr Leu Phe Pro Pro Gln Val Ala Val Thr Thr Ser Lys Val Ser
            20                  25                  30

Pro Trp Leu Ser Pro Val Leu Tyr Phe Val Gly Asn Tyr Leu Leu Leu
        35                  40                  45

Pro Ser Phe Phe Gly Arg Ile Ser Ile Thr Gly Gln Glu Asn Ile Pro
    50                  55                  60

Asn Thr Gly Pro Val Ile Leu Ala Pro Thr His Arg Ala Arg Trp Asp
65                  70                  75                  80

Ser Leu Leu Leu Pro Tyr Ala Thr Gly Arg Tyr Val Thr Gly Arg Asp
                85                  90                  95

Leu Lys Phe Met Val Thr Lys Thr Glu Cys Arg Gly Leu Gln Gly Trp
            100                 105                 110

Phe Val Arg Arg Met Gly Gly Phe Pro Ile Asp Thr Gln His Pro Ala
        115                 120                 125

Ile Ser Thr Leu Arg His Ala Val Glu Leu Leu Gln Gln Gly Gln Met
```

```
    130                 135                 140

Leu Val Ile Tyr Pro Glu Gly Asn Ile Phe Arg Asp Gly Lys Leu His
145                 150                 155                 160

Pro Leu Lys Ser Gly Ile Ser Arg Leu Ala Leu Ser Ala Glu Ser Ser
                165                 170                 175

His Pro Gly Leu Gly Val Lys Ile Leu Pro Val Ser Ile Asn Tyr Ser
            180                 185                 190

Gln Pro Tyr Pro Cys Trp Gly Thr Asp Val Ser Ile His Ile Gly Thr
        195                 200                 205

Ala Ile Asn Val Gln Asp Tyr Thr Asn Gly Lys Val Lys Gln Asn Ala
    210                 215                 220

Lys Arg Leu Thr Glu Asp Leu Ala Arg Asp Leu Gln Ser Leu Ser Asn
225                 230                 235                 240

Pro Glu Leu Ala Phe Ser His His Asp Val Ala Glu Val Ser Asn Pro
                245                 250                 255

Pro Lys Gly Lys Gly Asp Glu Gly Ala Glu Ala Val Glu Glu Gln Cys
            260                 265                 270

Gly Leu Gly Leu Cys Pro Ser Gly Ala Ile Ala Glu Glu Gln Ala Ser
        275                 280                 285

Ile Val Ser Val Ser Arg Thr Gly Ala Thr Ser Arg Arg Arg Glu Leu
    290                 295                 300

Leu Thr Val Asp Tyr
305
```

<210> 9
<211> 256
<212> PRT
<213> Nostoc azollae

<400> 9

```
Met Met Glu Phe Tyr Ser Ala Ser Thr Lys Pro Gln Ser Gln Ile Thr
1               5                  10                  15

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

Pro Trp Leu Thr Pro Leu Ala Tyr Leu Leu Gly Arg Asn Ile Val Leu
```

```
        35                  40                  45

Pro Leu Phe Phe Gly Asp Ile His Ile Thr Gly Gln Glu Asn Ile Pro
    50                  55                  60

Thr Ser Gly Pro Ile Ile Leu Ala Pro Thr His Arg Ser Arg Trp Asp
65                  70                  75                  80

Ser Leu Leu Leu Pro Tyr Ala Thr Gly Arg Cys Val Thr Gly Arg Asp
                85                  90                  95

Met Arg Phe Met Val Thr Ser Ser Glu Cys Lys Gly Val Gln Gly Trp
            100                 105                 110

Phe Val Arg Arg Met Gly Gly Phe Pro Val Asp Thr Gln Arg Pro Ala
        115                 120                 125

Ile Ala Thr Leu Arg His Thr Val Glu Leu Met Val Glu Gly Glu Met
    130                 135                 140

Leu Val Ile Tyr Pro Glu Gly Gly Ile Tyr Arg Asp Gly Lys Val His
145                 150                 155                 160

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

His Pro Gly Leu Glu Ile Lys Ile Leu Thr Val Gly Ile Asn Tyr Ser
            180                 185                 190

Gln Pro Tyr Pro Asn Trp Gly Thr Asp Val Asn Ile His Ile Gly Lys
        195                 200                 205

Pro Ile Lys Val Thr Asp Tyr Ile Ser Gly Cys Leu Lys Lys Glu Ala
    210                 215                 220

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

Gln Gly Leu Glu Val Lys Ser His Ala Phe Ala Glu Ile Pro Asn Ser
                245                 250                 255
```

<210> 10
<211> 246
<212> PRT
<213> Raphidiopsis brookii

<400> 10

```
Met Lys Glu Pro His Ser Ala Ser Thr Ser Phe Ala Asp Gln Ser Pro
```

```
1               5                   10                  15

Asn Ile Cys Leu Asp Pro Pro Phe Ser His Thr Arg Ser Trp Ile Ser
            20                  25                  30

Pro Trp Leu Thr Pro Leu Ala Tyr Trp Leu Gly Tyr His Leu Val Ile
        35                  40                  45

Pro Leu Phe Phe Gly Ser Ile His Val Glu Gly Gln Glu His Ile Pro
    50                  55                  60

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

Pro Leu Leu Leu Ala Tyr Ala Ala Gly Arg Tyr Ile Thr Gly Arg Asp
                85                  90                  95

Leu Arg Phe Met Val Met Ser Ser Glu Cys Arg Gly Ile Gln Gly Trp
            100                 105                 110

Phe Ile Leu Ser Met Gly Gly Phe Ala Val Asn Leu Gln Arg Pro Gly
        115                 120                 125

Ile Lys Ser Leu Arg His Val Ile Asp Leu Met Leu Gly Gly Glu Met
    130                 135                 140

Leu Val Ile Tyr Pro Glu Gly Gly Ile Phe Arg Asp Gly Lys Ile His
145                 150                 155                 160

Pro Leu Lys Ser Gly Ile Ala Arg Leu Ser Ile Asn Ala Gln Ser Leu
                165                 170                 175

Asp Pro Asn Leu Asp Ile Lys Ile Ile Pro Val Ser Ile Asn Tyr Asp
            180                 185                 190

Arg Pro Tyr Pro Asn Trp Gly Thr Lys Val Lys Ile Cys Ile Gly Glu
        195                 200                 205

Pro Val Lys Val Ala Asp Tyr Met His Gly Cys Leu Lys Gln Asn Ser
    210                 215                 220

Gln Tyr Leu Thr Arg Asp Leu Arg Ser Lys Leu Gln Glu Val Ala Asp
225                 230                 235                 240

Ser Gln Pro Val Ser Leu
                245
```

<210> 11

<211> 253
<212> PRT
<213> Chthonoplastes microcoleus

<400> 11

```
Met Gln Leu Gln Ser Arg Glu Gln Asn Gln Thr Pro Ser Thr Leu Asp
1               5                  10                  15

Lys Pro Asp Ser Lys Thr Val Asn Ser Ala Lys Ile Val Pro Val Asn
            20                  25                  30

Ser Asn Val Ser Pro Trp Leu Thr Ser Ile Phe Tyr Pro Leu Gly Arg
        35                  40                  45

Arg Ile Leu Met Pro Leu Tyr Phe Gly Arg Leu Gln Val Thr Gly Gln
    50                  55                  60

Asp Asn Ile Pro Lys Thr Gly Pro Val Ile Leu Ala Pro Thr His Arg
65                  70                  75                  80

Ser Arg Trp Asp Gly Leu Thr Met Ala Tyr Ala Ala Gly Lys Pro Val
                85                  90                  95

Thr Gly Arg Asp Leu Arg Phe Met Val Ser Gln Asp Glu Met Lys Gly
            100                 105                 110

Leu Gln Gly Trp Phe Ile Arg Arg Leu Gly Gly Phe Pro Val Asn Thr
        115                 120                 125

Lys His Pro Gly Ile Ser Ser Ile Arg His Ser Val Glu Leu Leu Arg
    130                 135                 140

Arg Gly Glu Ala Leu Val Met Phe Pro Glu Gly Asn Ile Phe Arg His
145                 150                 155                 160

Gly Asp Val Asn Pro Leu Lys Pro Gly Met Ala Arg Ile Ala Leu Gln
                165                 170                 175

Ala Glu Ser Asn Gln Pro Gly Leu Gly Leu Lys Ile Val Pro Val Ser
            180                 185                 190

Ile Arg Tyr Ser Asp Pro Val Pro Gln Trp Gly Ser Asp Ile Asp Ile
        195                 200                 205

Gln Ile Gly Ser Ala Leu Asp Val Ala Lys Tyr Cys Asn Gln Ser Ala
    210                 215                 220
```

```
Lys Lys Gly Ala Gln Gln Leu Thr Thr Asp Leu Glu Ile Ala Leu Lys
225                 230                 235                 240

Gln Val Asp Gln Lys Leu Glu Val Ile Lys Glu Ala Val
                245                 250
```

<210> 12
<211> 255
<212> PRT
<213> Cylindrospermopsis raciborskii

<400> 12

```
Met Lys Glu Ser His Ser Ala Ser Thr Ser Phe Ala Gly Gln Ser Pro
1               5                   10                  15

Asn Ile Cys Leu Asp Pro Pro Phe Ser His Thr Arg Ser Trp Ile Ser
            20                  25                  30

Pro Trp Leu Thr Pro Leu Ala Tyr Trp Leu Gly Tyr Asn Leu Val Ile
        35                  40                  45

Pro Ser Phe Phe Gly Arg Ile Asp Val Ala Gly Gln Gln His Ile Pro
    50                  55                  60

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

Pro Leu Leu Leu Ala Tyr Ala Ala Gly Arg Tyr Ile Thr Gly Arg Asp
                85                  90                  95

Leu Arg Phe Met Val Met Ser Ser Glu Cys Arg Gly Ile Gln Gly Trp
            100                 105                 110

Phe Ile Leu Ser Met Gly Gly Phe Ala Val Asn Leu Gln Arg Pro Gly
        115                 120                 125

Ile Arg Ser Leu Arg His Ala Val Asp Leu Met Leu Gly Gly Glu Met
    130                 135                 140

Leu Val Ile Tyr Pro Glu Gly Gly Ile Phe Arg Asp Gly Lys Ile His
145                 150                 155                 160

Pro Leu Lys Ser Gly Ile Ala Arg Leu Ser Ile Asn Ala Gln Ser Leu
                165                 170                 175

Asp Pro Asn Leu Asp Ile Lys Ile Ile Pro Val Ala Ile Asn Tyr Asp
            180                 185                 190
```

```
Arg Pro Tyr Pro Ser Trp Gly Thr Lys Val Lys Ile Arg Ile Gly Glu
        195                 200                 205

Pro Val Lys Val Thr Asp Tyr Ile His Gly Cys Leu Lys Gln Asn Ala
    210                 215                 220

Lys Tyr Leu Thr Arg Asp Leu Arg Ser Lys Leu Gln Glu Val Ala Asp
225                 230                 235                 240

Ser Gln Pro Asp Phe Ser Leu Arg Ala Asn Cys Leu Thr Asn Ser
                245                 250                 255
```

<210> 13
<211> 236
<212> PRT
<213> Oscillatoria PCC 6506

<400> 13

```
Met Ile Thr Phe Asp Ser Thr Asn Ser Ser Leu Glu Thr Gln Met Pro
1               5                  10                  15

Ala Lys Ala Thr Ser Ile Asn Ser His Val Ser Pro Trp Leu Ala Pro
            20                  25                  30

Leu Val Tyr Pro Leu Gly Arg Tyr Leu Ile Leu Pro Phe Tyr Phe Arg
        35                  40                  45

Ser Ile Glu Val Leu Gly Gln Glu His Ile Pro Arg Glu Gly Pro Val
    50                  55                  60

Ile Leu Ala Pro Thr His Arg Ser Arg Trp Asp Ser Gln Met Leu Pro
65                  70                  75                  80

Tyr Ser Thr Gly Arg Tyr Val Thr Gly Arg Asp Leu Arg Phe Met Val
                85                  90                  95

Ser Val Asn Glu Val Lys Gly Leu Gln Gly Trp Phe Ile Arg Arg Leu
            100                 105                 110

Gly Gly Phe Pro Val Asp Pro Lys His Pro Ala Ile Ala Thr Leu Arg
        115                 120                 125

His Gly Val Glu Leu Leu Leu Asp Gly Gln Met Phe Val Ile Phe Pro
    130                 135                 140

Glu Gly Gly Ile Phe Gln Asp Asn Gln Val His Pro Leu Lys Pro Gly
145                 150                 155                 160

Leu Ala Arg Leu Ala Ile Gln Ala Glu Ser Ser His Pro Gly Leu Asp
                165                 170                 175

Val Lys Ile Val Pro Met Ser Ile Arg Tyr His Pro Cys Ile Pro Arg
            180                 185                 190

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

Asp Tyr Cys Thr Gly Ala Gly Lys Lys Asp Ala Gln Leu Leu Thr Ala
    210                 215                 220

Asp Leu Glu Thr Ala Leu Lys Arg Ile Asp Arg Glu
225                 230                 235
```

<210> 14

<211> 250
<212> PRT
<213> Nodularia spumigena

<400> 14

```
Met Glu Ile Tyr Thr Ser His Gln Thr Ser Gln Glu Thr Pro Ala Ser
1               5                   10                  15

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

Trp Leu Ser Pro Leu Leu Tyr Leu Leu Gly Gln His Leu Leu Leu Pro
        35                  40                  45

Phe Phe Phe Arg Gln Ile Glu Ile Thr Gly Gln Glu Asn Ile Pro Leu
    50                  55                  60

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

Leu Leu Leu Pro Tyr Ala Thr Gly Arg Cys Val Thr Gly Arg Asp Leu
                85                  90                  95

Arg Phe Met Val Thr Ile Asn Glu Cys Gln Gly Leu Gln Gly Trp Leu
            100                 105                 110

Val Arg Ser Met Gly Gly Phe Pro Val Asn Pro Gln Arg Pro Ala Ile
        115                 120                 125

Thr Thr Leu Arg His Ala Val Glu Leu Leu Gln Gln Gly Glu Ile Leu
    130                 135                 140
```

```
Val Ile Tyr Pro Glu Gly Asn Ile Tyr Arg Asp Gly Lys Leu His Pro
145                 150                 155                 160

Leu Lys Pro Gly Ile Ala Arg Leu Ser Leu Thr Ala Glu Ser Ser His
                165                 170                 175

Pro Gly Leu Asp Val Lys Ile Leu Pro Val Ser Ile Asn Tyr Ser Gln
            180                 185                 190

Pro Tyr Pro Gln Trp Gly Thr Ser Val Lys Ile Ser Ile Gly Thr Thr
        195                 200                 205

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

Asn Leu Thr Ser Asp Leu Ser Leu Val Leu Gln Gln Leu Ser Pro Gln
225                 230                 235                 240

Glu Ser Val Ile Pro Glu Asn Val Asn Ser
                245                 250
```

<210> 15
<211> 242
<212> PRT
<213> Microcoleus vaginatus

<400> 15

```
Met Pro His Thr Ala Thr Ile Ala Val Glu Ser Ser Lys Ser Ser Leu
1               5                   10                  15

Val Lys Pro Met Pro Ala Lys Val Ala Ser Asn Thr Ser Arg Val Ser
            20                  25                  30

Pro Trp Leu Thr Ser Leu Leu Tyr Pro Leu Gly His Tyr Val Val Leu
        35                  40                  45

Pro Phe His Phe Gly Lys Ile Glu Ile Thr Gly Gln Glu His Leu Pro
    50                  55                  60

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

Ala Leu Met Met Pro Tyr Ser Thr Gly Gln Lys Val Thr Gly Arg Asp
                85                  90                  95

Leu Arg Phe Met Val Thr Ala Asp Glu Val Lys Gly Leu Gln Gly Trp
            100                 105                 110
```

```
Phe Ile Arg Asn Leu Gly Gly Phe Pro Val Asp Leu Lys His Pro Ala
        115                 120                 125

Ile Gly Thr Leu Arg Tyr Gly Val Glu Leu Leu Leu Asp Arg Glu Met
    130                 135                 140

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

Pro Ile Lys Pro Gly Leu Ala Arg Leu Ala Met Gln Ala Glu Leu Ser
                165                 170                 175

Lys Pro Gly Leu Gly Val Lys Ile Val Pro Met Ser Ile Arg Tyr Arg
            180                 185                 190

Pro Arg Ile Pro Gln Trp Gly Cys Gln Val Lys Ile Ala Ile Gly Ser
        195                 200                 205

Pro Ile Ser Val Ala Asp Tyr Ala Asn Gly Gly Gly Gly Lys Lys Glu
    210                 215                 220

Ala Arg Leu Leu Ser Ala Gln Leu Glu Met Asp Leu Lys Lys Leu Asp
225                 230                 235                 240

Glu Ser
```

<210> 16
<211> 245
<212> PRT
<213> *Escherichia coli*

<400> 16

```
Met Leu Tyr Ile Phe Arg Leu Ile Ile Thr Val Ile Tyr Ser Ile Leu
1               5                   10                  15

Val Cys Val Phe Gly Ser Ile Tyr Cys Leu Phe Ser Pro Arg Asn Pro
            20                  25                  30

Lys His Val Ala Thr Phe Gly His Met Phe Gly Arg Leu Ala Pro Leu
        35                  40                  45

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

Gly Asn Ala Ile Tyr Ile Ala Asn His Gln Asn Asn Tyr Asp Met Val
65                  70                  75                  80
```

```
Thr Ala Ser Asn Ile Val Gln Pro Pro Thr Val Thr Val Gly Lys Lys
                85                  90                  95

Ser Leu Leu Trp Ile Pro Phe Phe Gly Gln Leu Tyr Trp Leu Thr Gly
            100                 105                 110

Asn Leu Leu Ile Asp Arg Asn Asn Arg Thr Lys Ala His Gly Thr Ile
        115                 120                 125

Ala Glu Val Val Asn His Phe Lys Lys Arg Arg Ile Ser Ile Trp Met
    130                 135                 140

Phe Pro Glu Gly Thr Arg Ser Arg Gly Arg Gly Leu Leu Pro Phe Lys
145                 150                 155                 160

Thr Gly Ala Phe His Ala Ala Ile Ala Ala Gly Val Pro Ile Ile Pro
                165                 170                 175

Val Cys Val Ser Thr Thr Ser Asn Lys Ile Asn Leu Asn Arg Leu His
            180                 185                 190

Asn Gly Leu Val Ile Val Glu Met Leu Pro Pro Ile Asp Val Ser Gln
        195                 200                 205

Tyr Gly Lys Asp Gln Val Arg Glu Leu Ala Ala His Cys Arg Ser Ile
    210                 215                 220

Met Glu Gln Lys Ile Ala Glu Leu Asp Lys Glu Val Ala Glu Arg Glu
225                 230                 235                 240

Ala Ala Gly Lys Val
                245
```

<210> 17
<211> 231
<212> PRT
<213> Marinobacter adhaerens

<400> 17

```
Met Leu Arg Leu Leu Lys Ala Ala Trp Ile Leu Phe Trp Ala Ile Leu
1               5                  10                  15

Leu Thr Leu Ile Leu Phe Phe Pro Ile Val Ile Ala Ala Leu Leu Gly
            20                  25                  30

Lys Arg Gly Asp Ala Ala Phe His Gly Thr Gln Ile Tyr Ala Trp Ile
        35                  40                  45

   Ile Leu Lys Val Cys Gly Ile Arg Leu Lys Val Arg Gly Arg Glu Asn
       50                  55                  60

   Ile Glu Pro Gly Gln Arg Tyr Val Ile Leu Ser Asn His Ala Ser Tyr
   65                  70                  75                  80

   Leu Asp Pro Pro Ala Leu Val Leu Ala Leu Gly Leu Gln Tyr Arg Trp
                   85                  90                  95

   Val Ile Lys Lys Glu Leu Arg Lys Val Pro Leu Phe Gly Leu Ala Leu
               100                 105                 110

   Glu Ala Ser Arg Asn Leu Phe Ile Asp Arg Ser Lys Gly Ser Asp Ala
           115                 120                 125

   Leu Glu Ser Ile Lys Arg Gly Val Gly Gln Leu Pro Asp Gly Thr Gly
       130                 135                 140

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

   Pro Phe Lys Lys Gly Gly Phe Val Ile Ala Gln Asp Gly Glu Leu Pro
                   165                 170                 175

   Ile Leu Pro Val Thr Ile Cys Gly Ser His Gln Arg Leu Pro Lys Gly
               180                 185                 190

   Ser Ala Ala Phe Ser Arg Gly Asp Ile Glu Ile Val Ile His Pro Pro
           195                 200                 205

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

   Asn Ser Ile Ala Ser Ser Leu
   225                 230
```

<210> 18
<211> 332
<212> PRT

<213> *Oryza sativa*

<400> 18

```
Met Gly Thr Leu Leu Arg Pro Arg Pro Leu Ala His Ala Ala Gly Ala
1               5                   10                  15

Gly Asp Ala Thr Pro Ser Thr Ala His Ala Val Val Val Ser Gly Gly
            20                  25                  30
```

```
Arg Gly Arg Gly Val Glu Cys Gln Pro His Arg Val Arg Arg Arg Pro
        35                  40                  45

Gly Pro Gln Val Ala Val Ala Thr Ala Ser Trp Arg Arg Arg Arg Glu
    50                  55                  60

Thr Val Val Arg Ser Asp Phe Ala Ala Gly Gly Ala Ala Thr Met Gly
65                  70                  75                  80

Asp Ser Pro Gln Ala Leu Ser Asp Ile Asp Val Val Ser Arg Val Arg
                85                  90                  95

Gly Val Cys Phe Tyr Ala Val Thr Ala Val Ala Ala Ile Phe Leu Phe
            100                 105                 110

Val Ala Met Val Val Val His Pro Leu Val Leu Leu Phe Asp Arg Tyr
        115                 120                 125

Arg Arg Arg Ala Gln His Tyr Ile Ala Lys Ile Trp Ala Thr Leu Thr
    130                 135                 140

Ile Ser Met Phe Tyr Lys Leu Asp Val Glu Gly Met Glu Asn Leu Pro
145                 150                 155                 160

Pro Asn Ser Ser Pro Ala Val Tyr Val Ala Asn His Gln Ser Phe Leu
                165                 170                 175

Asp Ile Tyr Thr Leu Leu Thr Leu Gly Arg Cys Phe Lys Phe Ile Ser
            180                 185                 190

Lys Thr Ser Ile Phe Met Phe Pro Ile Ile Gly Trp Ala Met Tyr Leu
        195                 200                 205

Leu Gly Val Ile Pro Leu Arg Arg Met Asp Ser Arg Ser Gln Leu Asp
    210                 215                 220

Cys Leu Lys Arg Cys Val Asp Leu Val Lys Lys Gly Ala Ser Val Phe
225                 230                 235                 240

Phe Phe Pro Glu Gly Thr Arg Ser Lys Asp Gly Lys Leu Gly Ala Phe
                245                 250                 255

Lys Arg Gly Ala Phe Ser Val Ala Thr Lys Thr Gly Ala Pro Val Ile
            260                 265                 270

Pro Ile Thr Leu Leu Gly Thr Gly Lys Leu Met Pro Ser Gly Met Glu
        275                 280                 285
```

```
Gly Ile Leu Asn Ser Gly Ser Val Lys Leu Ile Ile His His Pro Ile
    290                 295                 300

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

Ala Asp Thr Leu Ile Leu Asn Gly Tyr Gly Val His
                325                 330
```

<210> 19
<211> 345
<212> PRT
<213> *Glycine max*

<400> 19

```
Met Glu Val Thr Pro Leu Ser Ser Pro Ser Pro Ile His Arg Leu His
1               5                   10                  15

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

Cys Thr Arg Arg Gly Thr Thr Thr Tyr Val His Pro Ile Leu Arg Thr
        35                  40                  45

Ser His Asn Ser Pro Pro Cys Ser Leu Gln Ala Ile Ser Lys Lys His
    50                  55                  60

Glu Asn Val Ser Trp Leu Ser Val Ser Pro Lys Leu His Val Gln Asn
65                  70                  75                  80

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

Ser Ala Gly Asp Gly Tyr Leu Leu Pro Glu Leu Lys Leu Glu Ser Lys
            100                 105                 110

Val Arg Gly Val Cys Phe Tyr Val Val Thr Ala Phe Ser Ala Ile Phe
        115                 120                 125

Leu Phe Met Leu Met Leu Val Gly His Pro Ser Val Leu Leu Phe Asp
    130                 135                 140

Arg Tyr Arg Arg Met Phe His His Phe Val Ala Lys Val Trp Ala Ala
145                 150                 155                 160

Leu Thr Val Ala Pro Phe Tyr Lys Ile Glu Phe Glu Gly Leu Glu Asn
                165                 170                 175
```

```
Leu Pro Pro Pro Asp Thr Pro Ala Val Tyr Val Ser Asn His Gln Ser
            180                 185                 190

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

Ile Ser Lys Thr Gly Ile Phe Leu Phe Pro Ile Ile Gly Trp Ala Met
    210                 215                 220

Phe Leu Leu Gly Val Ile Pro Leu Lys Arg Met Asp Ser Arg Ser Gln
225                 230                 235                 240

Leu Asp Cys Leu Lys Arg Cys Met Asp Leu Ile Lys Lys Gly Ala Ser
                245                 250                 255

Val Phe Phe Phe Pro Glu Gly Thr Arg Ser Lys Asp Gly Lys Leu Gly
            260                 265                 270

Thr Phe Lys Lys Gly Ala Phe Ser Val Ala Ala Lys Thr Asn Ala Pro
        275                 280                 285

Val Val Pro Ile Ser Leu Ile Gly Thr Gly Gln Ile Met Pro Ala Gly
    290                 295                 300

Lys Glu Gly Ile Val Asn Leu Gly Ser Val Lys Val Val Ile His Lys
305                 310                 315                 320

Pro Ile Val Gly Lys Asp Pro Asp Met Leu Cys Lys Glu Ala Arg Lys
                325                 330                 335

Thr Ile Ala Ser Val Leu Thr Gln Ser
            340                 345
```

<210> 20
<211> 344
<212> PRT
<213> Brassica napus

<400> 20

```
Met Asp Val Ala Ser Ala Arg Gly Val Ser Ser His Pro Pro Tyr Tyr
1               5                   10                  15

Ser Lys Pro Ile Cys Ser Ser Gln Ser Ser Leu Ile Arg Ile Pro Ile
            20                  25                  30

Ser Lys Gly Cys Cys Phe Ala Arg Ser Ser Asn Leu Ile Thr Ser Leu
        35                  40                  45
```

```
His Ala Ala Ser Arg Gly Val Thr Arg Arg Thr Ser Gly Val Gln Trp
    50                  55                  60

Cys Tyr Arg Ser Ile Arg Phe Asp Pro Phe Lys Val Asn Asp Lys Asn
65                  70                  75                  80

Ser Arg Thr Val Thr Val Arg Ser Asp Leu Ser Gly Ala Ala Thr Pro
                85                  90                  95

Glu Ser Thr Tyr Pro Glu Pro Glu Ile Lys Leu Ser Ser Arg Leu Arg
            100                 105                 110

Gly Ile Cys Phe Cys Leu Val Ala Gly Ile Ser Ala Ile Val Leu Ile
        115                 120                 125

Val Leu Met Ile Ile Gly His Pro Phe Val Leu Leu Phe Asp Arg Tyr
    130                 135                 140

Arg Arg Lys Phe His His Phe Ile Ala Lys Leu Trp Ala Ser Ile Ser
145                 150                 155                 160

Ile Tyr Pro Phe Tyr Lys Thr Asp Ile Gln Gly Leu Glu Asn Leu Pro
                165                 170                 175

Ser Ser Asp Thr Pro Cys Val Tyr Val Ser Asn His Gln Ser Phe Leu
            180                 185                 190

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

Lys Thr Gly Ile Phe Val Ile Pro Val Ile Gly Trp Ala Met Ser Met
    210                 215                 220

Met Gly Val Val Pro Leu Lys Arg Met Asp Pro Arg Ser Gln Val Asp
225                 230                 235                 240

Cys Leu Lys Arg Cys Met Glu Leu Val Lys Lys Gly Ala Ser Val Phe
                245                 250                 255

Phe Phe Pro Glu Gly Thr Arg Ser Lys Asp Gly Arg Leu Gly Pro Phe
            260                 265                 270

Lys Lys Gly Ala Phe Thr Ile Ala Ala Lys Thr Gly Val Pro Val Val
        275                 280                 285

Pro Ile Thr Leu Met Gly Thr Gly Lys Ile Met Pro Thr Gly Ser Glu
    290                 295                 300
```

```
Gly Ile Leu Asn His Gly Asp Val Arg Val Ile Ile His Lys Pro Ile
305                 310                 315                 320

Tyr Gly Ser Lys Ala Asp Val Leu Cys Glu Glu Ala Arg Asn Lys Ile
                325                 330                 335

Ala Glu Ser Met Asn Leu Leu Ser
            340
```

<210> 21
<211> 356
<212> PRT
<213> Arabidopsis thaliana

<400> 21

```
Met Asp Val Ala Ser Ala Arg Ser Ile Ser Ser His Pro Ser Tyr Tyr
1               5                   10                  15

Gly Lys Pro Ile Cys Ser Ser Gln Ser Ser Leu Ile Arg Ile Ser Arg
            20                  25                  30

Asp Lys Val Cys Cys Phe Gly Arg Ile Ser Asn Gly Met Thr Ser Phe
        35                  40                  45

Thr Thr Ser Leu His Ala Val Pro Ser Glu Lys Phe Met Gly Glu Thr
    50                  55                  60

Arg Arg Thr Gly Ile Gln Trp Ser Asn Arg Ser Leu Arg His Asp Pro
65                  70                  75                  80

Tyr Arg Phe Leu Asp Lys Lys Ser Pro Arg Ser Ser Gln Leu Ala Arg
                85                  90                  95

Asp Ile Thr Val Arg Ala Asp Leu Ser Gly Ala Ala Thr Pro Asp Ser
            100                 105                 110

Ser Phe Pro Glu Pro Glu Ile Lys Leu Ser Ser Arg Leu Arg Gly Ile
        115                 120                 125

Phe Phe Cys Val Val Ala Gly Ile Ser Ala Thr Phe Leu Ile Val Leu
    130                 135                 140

Met Ile Ile Gly His Pro Phe Val Leu Leu Phe Asp Pro Tyr Arg Arg
145                 150                 155                 160

Lys Phe His His Phe Ile Ala Lys Leu Trp Ala Ser Ile Ser Ile Tyr
                165                 170                 175
```

```
Pro Phe Tyr Lys Ile Asn Ile Glu Gly Leu Glu Asn Leu Pro Ser Ser
            180                 185                 190

Asp Thr Pro Ala Val Tyr Val Ser Asn His Gln Ser Phe Leu Asp Ile
        195                 200                 205

Tyr Thr Leu Leu Ser Leu Gly Lys Ser Phe Lys Phe Ile Ser Lys Thr
    210                 215                 220

Gly Ile Phe Val Ile Pro Ile Ile Gly Trp Ala Met Ser Met Met Gly
225                 230                 235                 240

Val Val Pro Leu Lys Arg Met Asp Pro Arg Ser Gln Val Asp Cys Leu
                245                 250                 255

Lys Arg Cys Met Glu Leu Leu Lys Lys Gly Ala Ser Val Phe Phe Phe
            260                 265                 270

Pro Glu Gly Thr Arg Ser Lys Asp Gly Arg Leu Gly Ser Phe Lys Lys
        275                 280                 285

Gly Ala Phe Thr Val Ala Ala Lys Thr Gly Val Ala Val Val Pro Ile
    290                 295                 300

Thr Leu Met Gly Thr Gly Lys Ile Met Pro Thr Gly Ser Glu Gly Ile
305                 310                 315                 320

Leu Asn His Gly Asn Val Arg Val Ile Ile His Lys Pro Ile His Gly
                325                 330                 335

Ser Lys Ala Asp Val Leu Cys Asn Glu Ala Arg Ser Lys Ile Ala Glu
            340                 345                 350

Ser Met Asp Leu
        355
```

<210> 22
<211> 374
<212> PRT
<213> *Zea mays*

<400> 22

```
Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly Leu Leu Phe
1               5                   10                  15

Leu Leu Ser Gly Leu Ile Val Asn Ala Ile Gln Ala Val Leu Phe Val
            20                  25                  30
```

```
Thr Ile Arg Pro Phe Ser Lys Ser Phe Tyr Arg Arg Ile Asn Arg Phe
        35                  40                  45

Leu Ala Glu Leu Leu Trp Leu Gln Leu Val Trp Val Val Asp Trp Trp
    50                  55                  60

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

Met Gly Lys Glu His Ala Leu Ile Ile Ser Asn His Arg Ser Asp Ile
                85                  90                  95

Asp Trp Leu Ile Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
            100                 105                 110

Ser Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
        115                 120                 125

Gly Trp Ser Met Trp Phe Ala Glu Tyr Leu Phe Leu Glu Arg Ser Trp
    130                 135                 140

Ala Lys Asp Glu Lys Thr Leu Lys Trp Gly Leu Gln Arg Leu Lys Asp
145                 150                 155                 160

Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
                165                 170                 175

Thr Pro Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ser Gln Gly
            180                 185                 190

Leu Pro Ala Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
        195                 200                 205

Ser Ala Val Ser Ile Met Arg Asp Phe Val Pro Ala Ile Tyr Asp Thr
    210                 215                 220

Thr Val Ile Val Pro Lys Asp Ser Pro Gln Pro Thr Met Leu Arg Ile
225                 230                 235                 240

Leu Lys Gly Gln Ser Ser Val Ile His Val Arg Met Lys Arg His Ala
                245                 250                 255

Met Ser Glu Met Pro Lys Ser Asp Glu Asp Val Ser Lys Trp Cys Lys
            260                 265                 270

Asp Ile Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Leu Ala Thr
```

```
          275                 280                 285

Gly Thr Phe Asp Glu Glu Ile Arg Pro Ile Gly Arg Pro Val Lys Ser
    290                 295                 300

Leu Leu Val Thr Leu Phe Trp Ser Cys Leu Leu Leu Phe Gly Ala Ile
305                 310                 315                 320

Glu Phe Phe Lys Trp Thr Gln Leu Leu Ser Thr Trp Arg Gly Val Ala
                325                 330                 335

Phe Thr Ala Ala Gly Met Ala Leu Val Thr Gly Val Met His Val Phe
            340                 345                 350

Ile Met Phe Ser Gln Ala Glu Arg Ser Ser Ser Ala Arg Ala Ala Arg
        355                 360                 365

Asn Arg Val Lys Lys Glu
    370
```

<210> 23
<211> 389
<212> PRT
<213> Arabidopsis thaliana

<400> 23

```
Met Val Ile Ala Ala Ala Val Ile Val Pro Leu Gly Leu Leu Phe Phe
1               5                  10                  15

Ile Ser Gly Leu Ala Val Asn Leu Phe Gln Ala Val Cys Tyr Val Leu
            20                  25                  30

Ile Arg Pro Leu Ser Lys Asn Thr Tyr Arg Lys Ile Asn Arg Val Val
        35                  40                  45

Ala Glu Thr Leu Trp Leu Glu Leu Val Trp Ile Val Asp Trp Trp Ala
    50                  55                  60

Gly Val Lys Ile Gln Val Phe Ala Asp Asn Glu Thr Phe Asn Arg Met
65                  70                  75                  80

Gly Lys Glu His Ala Leu Val Val Cys Asn His Arg Ser Asp Ile Asp
                85                  90                  95

Trp Leu Val Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly Ser
            100                 105                 110

Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile Gly
```

```
        115                 120                 125

Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp Ala
    130                 135                 140

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

Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe Thr
                165                 170                 175

Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu Leu
            180                 185                 190

Pro Ile Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val Ser
        195                 200                 205

Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met Thr
    210                 215                 220

Val Thr Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu Phe
225                 230                 235                 240

Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser Met
                245                 250                 255

Lys Asp Leu Pro Glu Ser Asp Asp Ala Ile Ala Gln Trp Cys Arg Asp
            260                 265                 270

Gln Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala Asp
        275                 280                 285

Thr Phe Pro Gly Gln Gln Glu Gln Asn Ile Gly Arg Pro Ile Lys Ser
    290                 295                 300

Leu Ala Val Val Leu Ser Trp Ala Cys Val Leu Thr Leu Gly Ala Ile
305                 310                 315                 320

Lys Phe Leu His Trp Ala Gln Leu Phe Ser Ser Trp Lys Gly Ile Thr
                325                 330                 335

Ile Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile Leu
            340                 345                 350

Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Val Pro
        355                 360                 365
```

```
Ala Lys Pro Lys Asp Asn His His Pro Glu Ser Ser Ser Gln Thr Glu
    370                 375                 380

Thr Glu Lys Glu Lys
385
```

<210> 24
<211> 376
<212> PRT
<213> Arabidopsis thaliana

<400> 24

```
Met Lys Ile Pro Ala Ala Leu Val Phe Ile Pro Val Gly Val Leu Phe
1               5                   10                  15

Leu Ile Ser Gly Leu Ile Val Asn Ile Ile Gln Leu Val Phe Phe Ile
            20                  25                  30

Ile Val Arg Pro Phe Ser Arg Ser Leu Tyr Arg Arg Ile Asn Lys Asn
        35                  40                  45

Val Ala Glu Leu Leu Trp Leu Gln Leu Ile Trp Leu Phe Asp Trp Trp
    50                  55                  60

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

Ile Gly Lys Glu His Ala Leu Val Leu Ser Asn His Arg Ser Asp Ile
                85                  90                  95

Asp Trp Leu Ile Gly Trp Val Met Ala Gln Arg Val Gly Cys Leu Gly
            100                 105                 110

Ser Ser Leu Ala Ile Met Lys Lys Glu Ala Lys Tyr Leu Pro Ile Ile
        115                 120                 125

Gly Trp Ser Met Trp Phe Ser Asp Tyr Ile Phe Leu Glu Arg Ser Trp
    130                 135                 140

Ala Lys Asp Glu Asn Thr Leu Lys Ala Gly Phe Lys Arg Leu Glu Asp
145                 150                 155                 160

Phe Pro Met Thr Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
                165                 170                 175

Thr Gln Glu Lys Leu Glu Ala Ala Gln Glu Tyr Ala Ser Ile Arg Ser
            180                 185                 190
```

```
Leu Pro Ser Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
        195                 200                 205

Ser Ala Val Ser Glu Ile Arg Ser Phe Val Pro Ala Ile Tyr Asp Cys
    210                 215                 220

Thr Leu Thr Val His Asn Asn Gln Pro Thr Pro Thr Leu Leu Arg Met
225                 230                 235                 240

Phe Ser Gly Gln Ser Ser Glu Ile Asn Leu Gln Met Arg Arg His Lys
                245                 250                 255

Met Ser Glu Leu Pro Glu Thr Asp Asp Gly Ile Ala Gln Trp Cys Gln
            260                 265                 270

Asp Leu Phe Ile Thr Lys Asp Ala Gln Leu Glu Lys Tyr Phe Thr Lys
        275                 280                 285

Asp Val Phe Ser Asp Leu Glu Val His Gln Ile Asn Arg Pro Ile Lys
    290                 295                 300

Pro Leu Ile Val Val Ile Ile Trp Leu Gly Phe Leu Val Phe Gly Gly
305                 310                 315                 320

Phe Lys Leu Leu Gln Trp Leu Ser Ile Val Ala Ser Trp Lys Ile Ile
                325                 330                 335

Leu Leu Phe Val Phe Phe Leu Val Ile Ala Thr Ile Thr Met Gln Ile
            340                 345                 350

Leu Ile Gln Ser Ser Glu Ser Gln Arg Ser Thr Pro Ala Lys Arg Pro
        355                 360                 365

Leu Gln Glu Gln Leu Ile Ser Ala
    370                 375
```

<210> 25
<211> 390
<212> PRT
<213> Brassica napus

<400> 25

```
Met Ala Met Ala Ala Ala Val Ile Val Pro Leu Gly Ile Leu Phe Phe
1               5                   10                  15

Ile Ser Gly Leu Val Val Asn Leu Leu Gln Ala Val Cys Tyr Val Leu
            20                  25                  30
```

```
Val Arg Pro Met Ser Lys Asn Thr Tyr Arg Lys Ile Asn Arg Val Val
        35                  40                  45

Ala Glu Thr Leu Trp Leu Glu Leu Val Trp Ile Val Asp Trp Trp Ala
    50                  55                  60

Gly Val Lys Ile Gln Val Phe Ala Asp Asp Glu Thr Phe Asn Arg Met
65                  70                  75                  80

Gly Lys Glu His Ala Leu Val Val Cys Asn His Arg Ser Asp Ile Asp
                85                  90                  95

Trp Leu Val Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly Ser
            100                 105                 110

Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile Gly
        115                 120                 125

Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp Ala
    130                 135                 140

Lys Asp Glu Ser Thr Leu Gln Ser Gly Leu Gln Arg Leu Asn Asp Phe
145                 150                 155                 160

Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe Thr
                165                 170                 175

Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu Leu
            180                 185                 190

Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val Ser
        195                 200                 205

Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met Thr
    210                 215                 220

Val Ala Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu Phe
225                 230                 235                 240

Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser Met
                245                 250                 255

Lys Asp Leu Pro Glu Pro Glu Asp Glu Ile Ala Gln Trp Cys Arg Asp
            260                 265                 270

Gln Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala Asp
        275                 280                 285
```

```
Thr Phe Pro Gly Gln Lys Glu Gln Asn Ile Gly Arg Pro Ile Lys Ser
    290                 295                 300

Leu Ala Val Val Val Ser Trp Ala Cys Leu Leu Thr Leu Gly Ala Met
305                 310                 315                 320

Lys Phe Leu His Trp Ser Asn Leu Phe Ser Ser Trp Lys Gly Ile Ala
                325                 330                 335

Leu Ser Ala Phe Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile Leu
            340                 345                 350

Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Ala Pro
        355                 360                 365

Ala Lys Pro Lys Asp Asn His Gln Ser Gly Pro Ser Ser Gln Thr Glu
    370                 375                 380

Val Glu Glu Lys Gln Lys
385                 390
```

<210> 26
<211> 306
<212> PRT
<213> Prunus dulcis

<400> 26

```
Met Gly Lys Glu His Ala Leu Val Ile Ser Asn His Arg Ser Asp Ile
1               5                   10                  15

Asp Trp Leu Val Gly Trp Val Leu Ala Gln Arg Ser Gly Cys Leu Gly
            20                  25                  30

Ser Ser Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
        35                  40                  45

Gly Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Ser Trp
    50                  55                  60

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

Phe Pro Gln Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
                85                  90                  95

Thr Gln Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ala Thr Gly
            100                 105                 110
```

```
Leu Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
        115                 120                 125

Thr Ala Val Ser Gln Met Arg Ser Phe Ala Pro Ala Ile Tyr Asp Val
    130                 135                 140

Thr Val Ala Ile Pro Lys Ser Ser Pro Ala Pro Thr Met Leu Arg Leu
145                 150                 155                 160

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

Met Arg Asp Leu Pro Glu Thr Asp Glu Ala Val Ala Gln Trp Cys Lys
            180                 185                 190

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

Gln Thr Phe Gly Asp Gln Gln Leu Lys Val Thr Gly Arg Pro Leu Lys
    210                 215                 220

Ser Leu Leu Val Val Thr Ala Trp Ala Cys Leu Leu Ile Leu Gly Ala
225                 230                 235                 240

Leu Lys Phe Leu Tyr Trp Ser Ser Leu Leu Ser Ser Trp Lys Gly Ile
                245                 250                 255

Ala Phe Ser Ala Leu Gly Leu Gly Val Val Thr Val Leu Met Gln Ile
            260                 265                 270

Leu Ile Arg Phe Ser Gln Ser Glu Arg Ser Thr Pro Ala Pro Val Ala
        275                 280                 285

Pro Thr Asn Asn Lys Asn Lys Gly Glu Ser Ser Gly Lys Pro Glu Lys
    290                 295                 300

Gln Gln
305
```

<210> 27
<211> 7642
<212> ADN
<213> Secuencia artificial

<220>
<223> vector pSGI-NB55 que incluye el gen sll 1752 LPAAT

<400> 27

```
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc       60
```

catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 120

ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat 180

aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat 240

ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 300

caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 360

gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 420

taaaacgacg gccagtgaat tgggcccgac tgcctttggt ggtattaccg atgagtggca 480

cgttattttc accgctctgg ccgtgttgag catggtgctg ggcaacgtgg tggctttagc 540

ccaaaccagc atgaaacgga tgttggccta ctcttccatc ggtcaagcag gctttgtgat 600

gattggccta gtggccggca gtgaagatgg ttacgccagc atggttttct acatgctcat 660

ctatctgttt atgaacctgg gggcgtttag ttgcattatt ctcttcaccc tccgcactgg 720

cagtgaccaa attagtgatt acgctggtct gtaccacaaa gaccccttgt taaccttggg 780

cttgagcatt tgtttattat ccttgggggg cattcctcct ctggcgggct ttttcggcaa 840

aatttacatc ttctgggccg gttggcaatc gggattgtat ggcctagtcc tacttggtct 900

ggttaccagt gtagtttcca tctactacta catccgggtg gtgaaaatga tggtggtgaa 960

ggagccccag gaaatgtccg aagtaatcaa aaattacccg gccatcaaat ggaatttacc 1020

cggcatgcgt cccctacagg tgggcattgt cgctactttg gttgctacct cgctggcagg 1080

tattctggct aatcccctct ttaacctcgc caccgattcc gtggtcagca ccaagatgtt 1140

gcagacagcc ctccagcaaa caggagaaac tccggcgatc gccatttccc atgatttacc 1200

ctaggggtat caggaaatat tgctttgcag gcaaaagcca atgagtgtaa ctatagaaac 1260

cgatttaaag gagatccact agtcctgagg ctgaaatgag ctgttgacaa ttaatcatcc 1320

ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt cacactaagg aggaaaaaaa 1380

ccatgaccaa ttctcccctg gcgatcgacc ctgctgtcat tggcccttcc agggtttccc 1440

cctggctaat caaattaatt tatcccctag gcaccaggtt tttacattgg tattttggcc 1500

cgatcgccat ccatggtcag gagcatttac cccgcagtgg cccgattatt ctggccccta 1560

cccaccgttc ccgttgggat gcaattttgc tttctttggc cgccggtcgg ggggtaactg 1620

gtcgagacct ccgttttatg gtggcggtga cggaagtgca gggtttacag ggttggttta 1680

ttcgccattt ggggggattt cccgttgacg tgaaaaggcc ggaaatcagt agtgtgagtt 1740

atagtgtgca gttactccaa gcaggggaaa tgttggtcat ttttcctgag gggggcattt 1800

ttcgggatca acacactgtc catcccctca aacggggtat tgggcgcatt gccatggaag 1860

tttgtaagca aaaccccaac actgacatta aagtgatccc agtgacgatc gcctacagtg 1920

acccctaccc aggcaaaggc acttcagtgg aaattaattt tggccagggc atcgccgcta 1980

gggactacga ccccagcacc atcaaggcta gctcccaaaa actgacccgt tgtttggccc 2040

atagaatgca gaacctttac agttccgaga atgtctctct gtgcgaggcg atcgccgctt 2100

cgtgattcta gatatctgca ggcctaagct ttatgcttgt aaaccgtttt gtgaaaaaat 2160

ttttaaaata aaaaagggga cctctagggt ccccaattaa ttagtaatat aatctattaa 2220

aggtcattca aaaggtcatc caccggatca attcccctgc tcgcgcaggc tgggtgccag 2280

gcccgatcct tggagccctt gccctcccgc acgatgatcg tgccgtgatc gaaatccaga 2340

tccttgaccc gcagttgcaa accctcactg atccgcatgc ccgttccata cagaagctgg 2400

gcgaacaaac gatgctcgcc ttccagaaaa ccgaggatgc gaaccacttc atccggggtc 2460

agcaccaccg gcaagcgccg cgacggccga ggtcttccga tctcctgaag ccagggcaga 2520

tccgtgcaca gcaccttgcc gtagaagaac agcaaggccg ccaatgcctg acgatgcgtg 2580

gagaccgaaa ccttgcgctc gttcgccagc caggacagaa atgcctcgac ttcgctgctg 2640

cccaaggttg ccgggtgacg cacaccgtgg aaacggatga aggcacgaac ccagtggaca 2700

taagcctgtt cggttcgtaa gctgtaatgc aagtagcgta tgcgctcacg caactggtcc 2760

agaaccttga ccgaacgcag cggtggtaac ggcgcagtgg cggttttcat ggcttgttat 2820

gactgttttt ttggggtaca gtctatgcct cggtcgggca tccaagcagc aagcgcgtta 2880

cgccgtgggt cgatgtttga tgttatggag cagcaacgat gttacgcagc agggcagtcg 2940

ccctaaaaca aagttaaaca tcatgaggga agcggtgatc gccgaagtat cgactcaact 3000

atcagaggta gttggcgtca tcgagcgcca tctcgaaccg acgttgctgg ccgtacattt 3060

gtacggctcc gcagtggatg gcggcctgaa gccacacagt gatattgatt tgctggttac 3120

ggtgaccgta aggcttgatg aaacaacgcg gcgagctttg atcaacgacc ttttggaaac 3180

ttcggcttcc cctggagaga gcgagattct ccgcgctgta gaagtcacca ttgttgtgca 3240

cgacgacatc attccgtggc gttatccagc taagcgcgaa ctgcaatttg gagaatggca 3300

gcgcaatgac attcttgcag gtatcttcga gccagccacg atcgacattg atctggctat 3360

cttgctgaca aaagcaagag aacatagcgt tgccttggta ggtccagcgg cggaggaact 3420

ctttgatccg gttcctgaac aggatctatt tgaggcgcta aatgaaacct taacgctatg 3480

gaactcgccg cccgactggg ctggcgatga gcgaaatgta gtgcttacgt tgtcccgcat 3540

ttggtacagc gcagtaaccg gcaaaatcgc gccgaaggat gtcgctgccg actgggcaat 3600

ggagcgcctg ccggcccagt atcagcccgt catacttgaa gctagacagg cttatcttgg 3660

acaagaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg tccactacgt 3720

gaaaggcgag atcaccaagg tagtcggcaa ataatgtcta acaattcgtt caagccgacg 3780

ccgcttcgcg gcgcggctta actcaagcgt tagatgcact aagcacataa ttgctcacag 3840

ccaaactatc aggtcaagtc tgcttttatt atttttaagc gtgcataata agccctacac 3900

aaattgggag atatatcatg aaaggctggc tttttcttgt tatcgcaata gttggcgaag 3960

taatcgcaac atccgcatta aaatctagcg agggctttac taagctgatc cggtggatga 4020

ccttttgaat gacctttaat agattatatt actaattaat tggggaccct agaggtcccc 4080

ttttttattt taaaaatttt ttcacaaaac ggtttacaag cataaagctt ccgcggtacc 4140

cgggaattcg ccctttcaag cttcagatca attcgcgcta acttacatta attgcgttgc 4200

gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 4260

aacgcgcggg gagaggcggt ttgcgtattg ggcgccaggg tggtttttct tttcaccagt 4320

gagacgggca acagctgatt gcccttcacc gcctggccct gagagagttg cagcaagcgg 4380

tccacgctgg tttgccccag caggcgaaaa tcctgtttga tggtggttaa cggcgggata 4440

taacatgagc tgtcttcggt atcgtcgtat cccactaccg agatatccgc accaacgcgc 4500

agcccggact cggtaatggc gcgcattgcg cccagcgcca tctgatcgtt ggcaaccagc 4560

atcgcagtgg gaacgatgcc ctcattcagc atttgcatgg tttgttgaaa accggacatg 4620

gcactccagt cgccttcccg ttccgctatc ggctgaattt gattgcgagt gagatattta 4680

tgccagccag ccagacgcag acgcgccgag acagaactta atgggcccgc taacagcgcg 4740

atttgctggt gacccaatgc gaccagatgc tccacgccca gtcgcgtacc gtcttcatgg 4800

gagaaaataa tactgttgat gggtgtctgg tcagagacat caagaaataa cgccggaaca 4860

ttagtgcagg cagcttccac agcaatggca tcctggtcat ccagcggata gttaatgatc 4920

agcccactga cgcgttgcgc gagaagattg tgcaccgccg ctttacaggc ttcgacgccg 4980

cttcgttcta ccatcgacac caccacgctg gcacccagtt gatcggcgcg agatttaatc 5040

gccgcgacaa tttgcgacgg cgcgtgcagg gccagactgg aggtggcaac gccaatcagc 5100

aacgactgtt tgcccgccag ttgttgtgcc acgcggttgg gaatgtaatt cagctccgcc 5160

atcgccgctt ccacttttcc ccgcgtttc gcagaaacgt ggctggcctg gttcaccacg 5220

cgggaaacgg tctgataaga gacaccggca tactctgcga catcgtataa cgttactggt 5280

ttcacattca ccaccctgaa ttgactctct tccgggcgct atcatgccat accgcgaaag 5340

gttttgcacc attcgatggt gtcaacgtaa atgcatgccg cttcgccttc gcgcaagctt 5400

agaagggcga attccggaca tatggatctt gggggaaatt aagaccaaac tcgatgacct 5460

ccaaaaagat gtaacttctc ttaagatcga tatggcaacg gtgaaaaccg agttatctgc 5520

ggtcaggatg gagataggta cagtcaagga tgatgttaaa gatgtcaaag ggcgggctaa 5580

tgctcaaatt tgggcgttga ttcttgccgt catcggagcc ataattacca ccttggtgcg 5640

ttttggcatt ttccctaatc cctaacaaaa aagcgaccag gcttttcttt caattgcccg 5700

atcgcctttg atattttccc aaaggataaa agctagtcca ttcagaatcg agccttaaag 5760

tactcccata ttggctagcc ccagaattac tccagcgccg aggatgtggc caaagctagc 5820

ggtgcccagc acagccccta aaccaaagcc gccaaagaag ttagaggaag gcatgggggt 5880

gcccacattt tgttgtttga tggtcaattt accaaaggcg atcgccaaaa tgttgcaagc 5940

aatcatcacc ccagcaactt tagggctcca ggacagggtg gcgggaacgg cggtggccaa 6000

caaaaagcta tgcattgaga ttctccagaa taaagacggt ttttaaaggg atagccccac 6060

gctaatgggg gtctttaaaa atctcatctt acggggacgc tctgcccctg ggaaaccacc 6120

gttgcaatac ttaacaaatt ttcgttttta gcttggcaaa tgtctttggc aaaattggtt 6180

gatctggctt aaatcgtcag ttatttgccc tggaatagtc tggggacggg caattctgat 6240

cagatttacc cccaacgctt ccgccacttt ttgcttaacc aattctcccc cctgggcacc 6300

ggaggcttta gttaccaccc cttgaatttg ccattgttgc cacagggctt tttccaatgg 6360

ttcggctacg gggggggcgca aagcaatgat acggtcggaa gtaaacccag cggcgatcgc 6420

ctgggctagg gcttggggat agggcagaat acgggcaaat agggcccagc ttggcgtaat 6480

catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 6540

gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa 6600

ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 6660

gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc 6720

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 6780

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 6840

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 6900

gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 6960

gactataaag ataccaggcg tttcccccctg gaagctccct cgtgcgctct cctgttccga 7020

ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 7080

atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 7140

tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 7200

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 7260

gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 7320

ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 7380

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 7440

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 7500

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 7560

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 7620

tatatgagta aacttggtct ga 7642

<210> 28
<211> 90
<212> ADN
<213> Secuencia artificial

<220>
<223> promotor TrcY

<400> 28

ctgaaatgag ctgttgacaa ttaatcatcc ggctcgtata atgtgtggaa ttgtgagcgg 60

ataacaattt cacactaagg aggaaaaaaa 90

<210> 29
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB395

<400> 29
cactaaggag gaaaaaaaat gaccaattct cccctggcg 39

<210> 30
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB396

<400> 30
cctgcagata tctagaatca cgaagcggcg atcg 34

<210> 31
<211> 8572
<212> ADN
<213> Secuencia artificial

<220>
<223> vector YC63 de expresión/integración

<400> 31

cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 60

catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 120

ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat 180

aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat 240

ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 300

caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 360

gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 420

taaaacgacg gccagtgaat tgggcccgac tgcctttggt ggtattaccg atgagtggca 480

cgttattttc accgctctgg ccgtgttgag catggtgctg ggcaacgtgg tggctttagc 540

ccaaaccagc atgaaacgga tgttggccta ctcttccatc ggtcaagcag gctttgtgat 600

gattggccta gtggccggca gtgaagatgg ttacgccagc atggttttct acatgctcat 660

ctatctgttt atgaacctgg gggcgtttag ttgcattatt ctcttcaccc tccgcactgg 720

cagtgaccaa attagtgatt acgctggtct gtaccacaaa gacccccttgt taaccttggg 780

cttgagcatt tgtttattat ccttgggggg cattcctcct ctggcgggct ttttcggcaa 840

aatttacatc ttctgggccg gttggcaatc gggattgtat ggcctagtcc tacttggtct 900

ggttaccagt gtagtttcca tctactacta catccgggtg gtgaaaatga tggtggtgaa 960

ggagccccag gaaatgtccg aagtaatcaa aaattacccg gccatcaaat ggaatttacc 1020

cggcatgcgt cccctacagg tgggcattgt cgctactttg gttgctacct cgctggcagg 1080

tattctggct aatcccctct ttaacctcgc caccgattcc gtggtcagca ccaagatgtt 1140

gcagacagcc ctccagcaaa caggagaaac tccggcgatc gccatttccc atgatttacc 1200

ctaggggtat caggaaatat tgctttgcag gcaaaagcca atgagtgtaa ctatagaaac 1260

cgatttaaag gagatccact agtcctgagg ctgaaatgag ctgttgacaa ttaatcatcc 1320

ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt cacactaagg aggaaaaaaa 1380

ccatggaaga cgccaaaaac ataaagaaag gcccggcgcc attctatccg ctggaagatg 1440

gaaccgctgg agagcaactg cataaggcta tgaagagata cgccctggtt cctggaacaa 1500

ttgcttttac agatgcacat atcgaggtgg acatcactta cgctgagtac ttcgaaatgt 1560

ccgttcggtt ggcagaagct atgaaacgat atgggctgaa tacaaatcac agaatcgtcg 1620

tatgcagtga aaactctctt caattcttta tgccggtgtt gggcgcgtta tttatcggag 1680

ttgcagttgc gcccgcgaac gacatttata atgaacgtga attgctcaac agtatgggca 1740

tttcgcagcc taccgtggtg ttcgtttcca aaaaggggtt gcaaaaaatt ttgaacgtgc 1800

aaaaaaagct cccaatcatc caaaaaatta ttatcatgga ttctaaaacg gattaccagg 1860

gatttcagtc gatgtacacg ttcgtcacat ctcatctacc tcccggtttt aatgaatacg 1920

attttgtgcc agagtccttc gatagggaca agacaattgc actgatcatg aactcctctg 1980

gatctactgg tctgcctaaa ggtgtcgctc tgcctcatag aactgcctgc gtgagattct 2040

cgcatgccag agatcctatt tttggcaatc aaatcattcc ggatactgcg attttaagtg 2100

ttgttccatt ccatcacggt tttggaatgt ttactacact cggatatttg atatgtggat 2160

ttcgagtcgt cttaatgtat agatttgaag aagagctgtt tctgaggagc cttcaggatt 2220

acaagattca aagtgcgctg ctggtgccaa ccctattctc cttcttcgcc aaaagcactc 2280

tgattgacaa atacgattta tctaatttac acgaaattgc ttctggtggc gctcccctct 2340

ctaaggaagt cggggaagcg gttgccaaga ggttccatct gccaggtatc aggcaaggat 2400

atgggctcac tgagactaca tcagctattc tgattacacc cgaggggggat gataaaccgg 2460

gcgcggtcgg taaagttgtt ccatttttg aagcgaaggt tgtggatctg gataccggga 2520

aaacgctggg cgttaatcaa agaggcgaac tgtgtgtgag aggtcctatg attatgtccg 2580

gttatgtaaa caatccggaa gcgaccaacg ccttgattga caaggatgga tggctacatt 2640

ctggagacat agcttactgg gacgaagacg aacacttctt catcgttgac cgcctgaagt 2700

ctctgattaa gtacaaaggc tatcaggtgg ctcccgctga attggaatcc atcttgctcc 2760

aacaccccaa catcttcgac gcaggtgtcg caggtcttcc cgacgatgac gccggtgaac 2820

ttcccgccgc cgttgttgtt ttggagcacg gaaagacgat gacggaaaaa gagatcgtgg 2880

attacgtcgc cagtcaagta acaaccgcga aaaagttgcg cggaggagtt gtgtttgtgg 2940

acgaagtacc gaaaggtctt accggaaaac tcgacgcaag aaaaatcaga gagatcctca 3000

taaaggccaa gaagggcgga aagatcgccg tgtaattcta gatatctgca ggcctaagct 3060

ttatgcttgt aaaccgtttt gtgaaaaaat ttttaaaata aaaaagggga cctctagggt 3120

ccccaattaa ttagtaatat aatctattaa aggtcattca aaaggtcatc caccggatca 3180

attcccctgc tcgcgcaggc tgggtgccag gcccgatcct tggagccctt gccctcccgc 3240

acgatgatcg tgccgtgatc gaaatccaga tccttgaccc gcagttgcaa accctcactg 3300

atccgcatgc ccgttccata cagaagctgg gcgaacaaac gatgctcgcc ttccagaaaa 3360

ccgaggatgc gaaccacttc atccggggtc agcaccaccg gcaagcgccg cgacggccga 3420

ggtcttccga tctcctgaag ccagggcaga tccgtgcaca gcaccttgcc gtagaagaac 3480

agcaaggccg ccaatgcctg acgatgcgtg gagaccgaaa ccttgcgctc gttcgccagc 3540

caggacagaa atgcctcgac ttcgctgctg cccaaggttg ccgggtgacg cacaccgtgg 3600

aaacggatga aggcacgaac ccagtggaca taagcctgtt cggttcgtaa gctgtaatgc 3660

aagtagcgta tgcgctcacg caactggtcc agaaccttga ccgaacgcag cggtggtaac 3720

ggcgcagtgg cggttttcat ggcttgttat gactgttttt ttggggtaca gtctatgcct 3780

cggtcgggca tccaagcagc aagcgcgtta cgccgtgggt cgatgtttga tgttatggag 3840

cagcaacgat gttacgcagc agggcagtcg ccctaaaaca aagttaaaca tcatgaggga 3900

agcggtgatc gccgaagtat cgactcaact atcagaggta gttggcgtca tcgagcgcca 3960

tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg gcggcctgaa 4020

gccacacagt gatattgatt tgctggttac ggtgaccgta aggcttgatg aaacaacgcg 4080

gcgagctttg atcaacgacc ttttggaaac ttcggcttcc cctggagaga gcgagattct 4140

ccgcgctgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc gttatccagc 4200

taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgcag gtatcttcga 4260

gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag aacatagcgt 4320

tgccttggta ggtccagcgg cggaggaact ctttgatccg gttcctgaac aggatctatt 4380

tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg ctggcgatga 4440

gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg gcaaaatcgc 4500

gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt atcagcccgt 4560

catacttgaa gctagacagg cttatcttgg acaagaagaa gatcgcttgg cctcgcgcgc 4620

agatcagttg gaagaatttg tccactacgt gaaaggcgag atcaccaagg tagtcggcaa 4680

ataatgtcta acaattcgtt caagccgacg ccgcttcgcg gcgcggctta actcaagcgt 4740

tagatgcact aagcacataa ttgctcacag ccaaactatc aggtcaagtc tgcttttatt 4800

atttttaagc gtgcataata agccctacac aaattgggag atatatcatg aaaggctggc 4860

tttttcttgt tatcgcaata gttggcgaag taatcgcaac atccgcatta aaatctagcg 4920

agggctttac taagctgatc cggtggatga ccttttgaat gacctttaat agattatatt 4980

actaattaat tggggaccct agaggtcccc ttttttattt taaaaatttt ttcacaaaac 5040

ggtttacaag cataaagctt ccgcggtacc cgggaattcg ccctttcaag cttcagatca 5100

attcgcgcta acttacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 5160

tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 5220

ggcgccaggg tggtttttct tttcaccagt gagacgggca acagctgatt gcccttcacc 5280

gcctggccct gagagagttg cagcaagcgg tccacgctgg tttgccccag caggcgaaaa 5340

tcctgtttga tggtggttaa cggcgggata taacatgagc tgtcttcggt atcgtcgtat 5400

cccactaccg agatatccgc accaacgcgc agcccggact cggtaatggc gcgcattgcg 5460

cccagcgcca tctgatcgtt ggcaaccagc atcgcagtgg gaacgatgcc ctcattcagc 5520

atttgcatgg tttgttgaaa accggacatg gcactccagt cgccttcccg ttccgctatc 5580

ggctgaattt gattgcgagt gagatattta tgccagccag ccagacgcag acgcgccgag 5640

acagaactta atgggcccgc taacagcgcg atttgctggt gacccaatgc gaccagatgc 5700

tccacgccca gtcgcgtacc gtcttcatgg gagaaaataa tactgttgat gggtgtctgg 5760

tcagagacat caagaaataa cgccggaaca ttagtgcagg cagcttccac agcaatggca 5820

tcctggtcat ccagcggata gttaatgatc agcccactga cgcgttgcgc gagaagattg 5880

tgcaccgccg ctttacaggc ttcgacgccg cttcgttcta ccatcgacac caccacgctg 5940

gcacccagtt gatcggcgcg agatttaatc gccgcgacaa tttgcgacgg cgcgtgcagg 6000

gccagactgg aggtggcaac gccaatcagc aacgactgtt tgcccgccag ttgttgtgcc 6060

acgcggttgg gaatgtaatt cagctccgcc atcgccgctt ccacttttc ccgcgttttc 6120

gcagaaacgt ggctggcctg gttcaccacg cgggaaacgg tctgataaga gacaccggca 6180

tactctgcga catcgtataa cgttactggt ttcacattca ccaccctgaa ttgactctct 6240

tccgggcgct atcatgccat accgcgaaag gttttgcacc attcgatggt gtcaacgtaa 6300

atgcatgccg cttcgccttc gcgcaagctt agaagggcga attccggaca tatggatctt 6360

gggggaaatt aagaccaaac tcgatgacct ccaaaaagat gtaacttctc ttaagatcga 6420

tatggcaacg gtgaaaaccg agttatctgc ggtcaggatg gagataggta cagtcaagga 6480

tgatgttaaa gatgtcaaag ggcgggctaa tgctcaaatt tgggcgttga ttcttgccgt 6540

catcggagcc ataattacca ccttggtgcg ttttggcatt ttccctaatc cctaacaaaa 6600

aagcgaccag gcttttcttt caattgcccg atcgcctttg atattttccc aaaggataaa 6660

agctagtcca ttcagaatcg agccttaaag tactcccata ttggctagcc ccagaattac 6720

tccagcgccg aggatgtggc caaagctagc ggtgcccagc acagccccta aaccaaagcc 6780

gccaaagaag ttagaggaag gcatgggggt gcccacattt tgttgtttga tggtcaattt 6840

accaaaggcg atcgccaaaa tgttgcaagc aatcatcacc ccagcaactt tagggctcca 6900

ggacagggtg gcgggaacgg cggtggccaa caaaaagcta tgcattgaga ttctccagaa 6960

taaagacggt ttttaaaggg atagccccac gctaatgggg gtctttaaaa atctcatctt 7020

acggggacgc tctgcccctg ggaaaccacc gttgcaatac ttaacaaatt ttcgttttta 7080

gcttggcaaa tgtctttggc aaaattggtt gatctggctt aaatcgtcag ttatttgccc 7140

tggaatagtc tggggacggg caattctgat cagatttacc cccaacgctt ccgccacttt 7200

ttgcttaacc aattctcccc cctgggcacc ggaggcttta gttaccaccc cttgaatttg 7260

ccattgttgc cacagggctt tttccaatgg ttcggctacg ggggggcgca aagcaatgat 7320

acggtcggaa gtaaacccag cggcgatcgc ctgggctagg gcttggggat agggcagaat 7380

acgggcaaat agggcccagc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt 7440

gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg 7500

gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt 7560

cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt 7620

tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc 7680

tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 7740

ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 7800

ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac aaaaatcgac 7860

gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttcccccctg 7920

gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 7980

ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg 8040

tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 8100

gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac 8160

tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt 8220

tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc 8280

tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 8340

ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 8400

ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 8460

gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 8520

aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct ga 8572

<210> 32
<211> 1032
<212> ADN
<213> Synechocystis PCC6803

<400> 32

tatttgcccg tattctgccc tatccccaag ccctagccca ggcgatcgcc gctgggttta 60

cttccgaccg tatcattgct ttgcgccccc ccgtagccga accattggaa aaagccctgt 120

ggcaacaatg gcaaattcaa ggggtggtaa ctaaagcctc cggtgcccag gggggagaat 180

tggttaagca aaaagtggcg gaagcgttgg gggtaaatct gatcagaatt gcccgtcccc 240

agactattcc agggcaaata actgacgatt taagccagat caaccaattt tgccaaagac 300

atttgccaag ctaaaaacga aaatttgtta agtattgcaa cggtggtttc ccaggggcag 360

agcgtccccg taagatgaga tttttaaaga cccccattag cgtggggcta tccctttaaa 420

aaccgtcttt attctggaga atctcaatgc atagcttttt gttggccacc gccgttcccg 480

ccaccctgtc ctggagccct aaagttgctg gggtgatgat tgcttgcaac attttggcga 540

tcgcctttgg taaattgacc atcaaacaac aaaatgtggg caccccccatg ccttcctcta 600

acttctttgg cggctttggt ttaggggctg tgctgggcac cgctagcttt ggccacatcc 660

tcggcgctgg agtaattctg gggctagcca atatgggagt actttaaggc tcgattctga 720

atggactagc ttttatcctt tgggaaaata tcaaaggcga tcgggcaatt gaaagaaaag 780

cctggtcgct tttttgttag ggattaggga aaatgccaaa acgcaccaag gtggtaatta 840

tggctccgat gacggcaaga atcaacgccc aaatttgagc attagcccgc cctttgacat 900

ctttaacatc atccttgact gtacctatct ccatcctgac cgcagataac tcggttttca 960

ccgttgccat atcgatctta agagaagtta catcttttg gaggtcatcg agtttggtct 1020

taatttcccc ca 1032

<210> 33
<211> 824
<212> ADN
<213> Synechocystis PCC6803

<400> 33

cctttaaatc ggtttctata gttacactca ttggcttttg cctgcaaagc aatatttcct 60

gataccccta gggtaaatca tgggaaatgg cgatcgccgg agtttctcct gtttgctgga 120

gggctgtctg caacatcttg gtgctgacca cggaatcggt ggcgaggtta aagaggggat 180

tagccagaat acctgccagc gaggtagcaa ccaaagtagc gacaatgccc acctgtaggg 240

gacgcatgcc gggtaaattc catttgatgg ccgggtaatt tttgattact tcggacattt 300

cctggggctc cttcaccacc atcattttca ccacccggat gtagtagtag atggaaacta 360

cactggtaac cagaccaagt aggactaggc catacaatcc cgattgccaa ccggcccaga 420

agatgtaaat tttgccgaaa aagcccgcca gaggaggaat gccccccaag gataataaac 480

aaatgctcaa gcccaaggtt aacaaggggt ctttgtggta cagaccagcg taatcactaa 540

tttggtcact gccagtgcgg agggtgaaga gaataatgca actaaacgcc cccaggttca 600

taaacagata gatgagcatg tagaaaacca tgctggcgta accatcttca ctgccggcca 660

ctaggccaat catcacaaag cctgcttgac cgatggaaga gtaggccaac atccgtttca 720

tgctggtttg ggctaaagcc accacgttgc ccagcaccat gctcaacacg gccagagcgg 780

tgaaaataac gtgccactca tcggtaatac caccaaaggc agtc 824

<210> 34
<211> 31
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB393

<400> 34
ttctagatat ctgcaggcct aagctttatg c 31

<210> 35
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB394

<400> 35
ttttttcct ccttagtgtg aaattgttat ccgc 34

<210> 36
<211> 6414
<212> ADN
<213> Secuencia artificial

<220>
<223> vector pSGI-NB158

<400> 36

cctgaggctg aaatgagctg ttgacaatta atcatccggc tcgtataatg tgtggaattg 60

tgagcggata acaatttcac actaaggagg aaaaaaacca tggcgaacgg tagcgctgtc 120

tctctgaaga gcggctcctt gaatacgcaa gaggacactt cttcttcccc accgccacgc 180

gcgttcatca accaattacc cgactggtcc atgttattga cggcgattac cactgtcttt 240

gttgccgcag agaaacagtg gactatgtta gaccgcaaga gcaagcgctc cgatatgtta 300

gtggattctt ttggcatgga acgcattgtg caggatggct tagtgtttcg tcaatctttt 360

agcattcgtt cttatgaaat cggtgcagat cgtcgtgcat ccattgaaac cttaatgaac 420

catctgcagg aaactagctt gaatcattgc aaatccattc gcttgttgaa tgagggtttt 480

ggtcgcaccc ccgagatgtg caaacgtgac ttgatctggg tggttacccg catgcacatc 540

atggtcaacc gctaccctac ctggggtgat accgttgaga ttaacacttg ggtttcccaa 600

agcggcaaga atggtatggg tcgtgattgg ctgatttccg actgtaatac cggcgaaatc 660

ctgatccgcg cgacgtctgc atgggcgatg atgaaccaaa agacccgtcg tctgtctaaa 720

ctgccttacg aagtcagcca agagattgct ccgcacttcg tcgacagccc tcccgtgatc 780

gaggacggcg accgtaagtt acacaagttc gatgtgaaaa ccggcgacag catccgtaaa 840

ggtttgactc cgcgttggaa tgacttagat gttaatcagc acgttaacaa cgttaagtat 900

atcggctgga tcttagagag catgccgacc gaggtcttgg aaactcatga actgtgtttc 960

ttaactctgg agtatcgtcg cgagtgcggt cgcgatagcg tgctggaatc tgtgaccgcg 1020

atggatcctt ctaatgaagg tggtcgctcc cactaccagc atttactgcg cttggaggac 1080

ggtactgaca tcgttaaggg ccgcactgag tggcgtccaa agaatgcccg gaatattggt 1140

gccattagta ccggtaaaac cagtaatggt aatcccgcca gttaataacc tttcaagctt 1200

cagatcaatt cgcgctaact tacattaatt gcgttgcgct cactgcccgc tttccagtcg 1260

ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg 1320

cgtattgggc gccagggtgg tttttctttt caccagtgag acgggcaaca gctgattgcc 1380

cttcaccgcc tggccctgag agagttgcag caagcggtcc acgctggttt gccccagcag 1440

gcgaaaatcc tgtttgatgg tggttaacgg cgggatataa catgagctgt cttcggtatc 1500

gtcgtatccc actaccgaga tatccgcacc aacgcgcagc ccggactcgg taatggcgcg 1560

cattgcgccc agcgccatct gatcgttggc aaccagcatc gcagtgggaa cgatgccctc 1620

attcagcatt tgcatggttt gttgaaaacc ggacatggca ctccagtcgc cttcccgttc 1680

cgctatcggc tgaatttgat tgcgagtgag atatttatgc cagccagcca gacgcagacg 1740

cgccgagaca gaacttaatg ggcccgctaa cagcgcgatt tgctggtgac ccaatgcgac 1800

cagatgctcc acgcccagtc gcgtaccgtc ttcatgggag aaaataatac tgttgatggg 1860

tgtctggtca gagacatcaa gaaataacgc cggaacatta gtgcaggcag cttccacagc 1920

aatggcatcc tggtcatcca gcggatagtt aatgatcagc ccactgacgc gttgcgcgag 1980

aagattgtgc accgccgctt tacaggcttc gacgccgctt cgttctacca tcgacaccac 2040

cacgctggca cccagttgat cggcgcgaga tttaatcgcc gcgacaattt gcgacggcgc 2100

gtgcagggcc agactggagg tggcaacgcc aatcagcaac gactgtttgc ccgccagttg 2160

ttgtgccacg cggttgggaa tgtaattcag ctccgccatc gccgcttcca ctttttcccg 2220

cgttttcgca gaaacgtggc tggcctggtt caccacgcgg gaaacggtct gataagagac 2280

accggcatac tctgcgacat cgtataacgt tactggtttc acattcacca ccctgaattg 2340

actctcttcc gggcgctatc atgccatacc gcgaaaggtt ttgcaccatt cgatggtgtc 2400

aacgtaaatg catgccgctt cgccttcgcg caagcttaga agggcgaatt ccggaccaca 2460

agataaaaat atatcatcat gaacaataaa actgtctgct tacataaaca gtaatacaag 2520

gggtgttatg agccatattc aacgggaaac gtcttgctcg aggccgcgat taaattccaa 2580

catggatgct gatttatatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc 2640

gacaatctat cgattgtatg ggaagcccga tgcgccagag ttgtttctga aacatggcaa 2700

aggtagcgtt gccaatgatg ttacagatga gatggtcaga ctaaactggc tgacggaatt 2760

tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgcat ggttactcac 2820

cactgcgatc cccgggaaaa cagcattcca ggtattagaa gaatatcctg attcaggtga 2880

aaatattgtt gatgcgctgg cagtgttcct gcgccggttg cattcgattc ctgtttgtaa 2940

ttgtcctttt aacagcgatc gcgtatttcg tctcgctcag gcgcaatcac gaatgaataa 3000

cggtttggtt gatgcgagtg attttgatga cgagcgtaat ggctggcctg ttgaacaagt 3060

ctggaaagaa atgcataagc ttttgccatt ctcaccggat tcagtcgtca ctcatggtga 3120

tttctcactt gataacctta tttttgacga ggggaaatta ataggttgta ttgatgttgg 3180

acgagtcgga atcgcagacc gataccagga tcttgccatc ctatggaact gcctcggtga 3240

gttttctcct tcattacaga aacggctttt tcaaaaatat ggtattgata atcctgatat 3300

gaataaattg cagtttcatt tgatgctcga tgagtttttc taatcagaat tggttaattg 3360

gttgtaacac tggcagagca ttacgctgac ttgacgggac ggcggcccgg tatggatggc 3420

accgatgcgg aatcccaaca gattgccttt gacaacaatg tggcctggaa taacctgggg 3480

gatttgtcca ccaccaccca acgggcctac acttcggcta ttagcacaga cacagtgcag 3540

agtgtttatg gcgttaatct ggaaaaaaac gataacattc ccattgtttt tgcgtggccc 3600

atttttccca ccacccttaa tcccacagat tttcaggtaa tgcttaacac gggggaaatt 3660

gtcaccccgg tgatcgcctc tttgattccc aacagtgaat acaacgaacg gcaaacggta 3720

gtaattacgg gcaattttgg taatcgttta accccaggca cggagggagc gatttatccc 3780

gtttccgtag gcacagtgtt ggacagtact cctttggaaa tggtgggacc caacggcccg 3840

gtcagtgcgg tgggtattac cattgatagt ctcaacccct acgtggccgg caatggtccc 3900

aaaattgtcg ccgctaagtt agaccgcttc agtgacctgg gggaaggggc tcccctctgg 3960

ttagccacca atcaaaataa cagtggcggg gatttatatg gagaccaagc ccaatttcgt 4020

ttgcgaattt acaccagcgc cggtttttcc cccgatggca ttgccagttt actacccaca 4080

gaatttgaac ggtattttca actccaagcg gaagatatta cgggacggac agttatccta 4140

acccaaactg gtgttgatta tgaaattccc ggctttggtc tggtgcaggt gttggggctg 4200

gcggatttgg ccggggttca ggacagctat gacctgactt acatcgaaga tcatgacaac 4260

tattacgaca ttatcctcaa aggggacgaa gccgcagttc gccaaattaa gagggttgct 4320

ttgccctccg aaggggatta ttcggcggtt tataatcccg gtggccccgg caatgatcca 4380

gagaatggtc ccccatgttc agctactgac ggggtggtgc gtaacggcaa aagcaccgcc 4440

ggacatcagc gctagcggag tgtatactgg cttactatgt tggcactgat gagggtgtca 4500

gtgaagtgct tcatgtggca ggagaaaaaa ggctgcaccg gtgcgtcagc agaatatgtg 4560

atacaggata tattccgctt cctcgctcac tgactcgcta cgctcggtcg ttcgactgcg 4620

gcgagcggaa atggcttacg aacggggcgg agatttcctg gaagatgcca ggaagatact 4680

taacagggaa gtgagagggc cgcggcaaag ccgtttttcc ataggctccg cccccctgac 4740

aagcatcacg aaatctgacg ctcaaatcag tggtggcgaa acccgacagg actataaaga 4800

taccaggcgt ttcccccctgg cggctccctc gtgcgctctc ctgttcctgc ctttcggttt 4860

accggtgtca ttccgctgtt atggccgcgt ttgtctcatt ccacgcctga cactcagttc 4920

cgggtaggca gttcgctcca agctggactg tatgcacgaa ccccccgttc agtccgaccg 4980

ctgcgcctta tccggtaact atcgtcttga gtccaacccg gaaagacatg caaaagcacc 5040

actggcagca gccactggta attgatttag aggagttagt cttgaagtca tgcgccggtt 5100

aaggctaaac tgaaaggaca agttttggtg actgcgctcc tccaagccag ttacctcggt 5160

tcaaagagtt ggtagctcag agaaccttcg aaaaaccgcc ctgcaaggcg gttttttcgt 5220

tttcagagca agagattacg cgcagaccaa aacgatctca agaagatcat cttattaatc 5280

agataaaata tttctagatt tcagtgcaat ttatctcttc aaatgtagca cctgaagtca 5340

gccccatacg atataagttg taattctcat gtttgacagc ttatcatcga taagctttaa 5400

tgcggtagtt tatcacagtt aaattattgc tgaagcggaa tccctggtta atgccgccgc 5460

cgatgccaat tgcattctcc aagtggggca cattgaacgc ttcaacccgg catttttaga 5520

gctaaccaaa attctcaaaa cggaagagtt attggcgatc gaagcccatc gcatgagtcc 5580

ctattcccag cgggccaatg atgtctccgt ggtattggat ttgatgatcc atgacattga 5640

cctgttgctg gaattggtgg gttcggaagt ggttaaactg tccgccagtg gcagtcgggc 5700

ttctgggtca ggatatttgg attatgtcac cgctacgtta ggcttctcct ccggcattgt 5760

ggccaccctc accgccagta aggtcaccca tcgtaaaatt cgttccatcg ccgcccactg 5820

caaaaattcc ctcaccgaag cggattttct caataacgaa attttgatcc atcgccaaac 5880

caccgctgat tggagcgcgg actatggcca ggtattgtat cgccaggatg gtctaatcga 5940

aaaggtttac accagtaata ttgaacctct ccacgctgaa ttagaacatt ttattcattg 6000

tgttagggga ggtgatcaac cctcagtggg gggagaacag gccctcaagg ccctgaagtt 6060

agccagttta attgaagaaa tggccctgga cagtcaggaa tggcatgggg gggaagttgt 6120

gacagaatat caagatgcca ccctggccct cagtgcgagt gtttaaatca acttaattaa 6180

tgcaattatt gcgagttcaa actcgataac tttgtgaaat attactgttg aattaatcta 6240

tgactattca atacaccccc ctagccgatc gcctgttggc ctacctcgcc gccgatcgcc 6300

taaatctcag cgccaagagt agttccctca acaccagtat tctgctcagc agtgacctat 6360

tcaatcagga agggggaatt gtaacagcca actatggctt tgatggttat atgg 6414

<210> 37
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB470

<400> 37
agagaatggt cccccatgtt cagctactga cggggtg 37

<210> 38
<211> 56
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB471

<400> 38
ggattccgct tcagcaataa tttaactgtg ataaactacc gcattaaagc ttatcg 56

<210> 39
<211> 57
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB466

<400> 39
gggcgaattc cggaccacaa gataaaaata tatcatcatg aacaataaaa ctgtctg 57

<210> 40
<211> 34
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB467

<400> 40
ggtgccatcc ataccgggcc gccgtcccgt caag 34

<210> 41
<211> 989
<212> ADN
<213> Synechocystis PCC6803

<400> 41

attgctgaag cggaatccct ggttaatgcc gccgccgatg ccaattgcat tctccaagtg 60

gggcacattg aacgcttcaa cccggcattt ttagagctaa ccaaaattct caaaacggaa 120

gagttattgg cgatcgaagc ccatcgcatg agtccctatt cccagcgggc caatgatgtc 180

tccgtggtat tggatttgat gatccatgac attgacctgt tgctggaatt ggtgggttcg 240

gaagtggtta aactgtccgc cagtggcagt cgggcttctg ggtcaggata tttggattat 300

gtcaccgcta cgttaggctt ctcctccggc attgtggcca ccctcaccgc cagtaaggtc 360

acccatcgta aaattcgttc catcgccgcc cactgcaaaa attccctcac cgaagcggat 420

tttctcaata acgaaatttt gatccatcgc caaaccaccg ctgattggag cgcggactat 480

ggccaggtat tgtatcgcca ggatggtcta atcgaaaagg tttacaccag taatattgaa 540

cctctccacg ctgaattaga acattttatt cattgtgtta ggggaggtga tcaaccctca 600

gtgggggggag aacaggccct caaggccctg aagttagcca gtttaattga agaaatggcc 660

ctggacagtc aggaatggca tgggggggaa gttgtgacag aatatcaaga tgccaccctg 720

gccctcagtg cgagtgttta aatcaactta attaatgcaa ttattgcgag ttcaaactcg 780

ataactttgt gaaatattac tgttgaatta atctatgact attcaataca ccccccctagc 840

cgatcgcctg ttggcctacc tcgccgccga tcgcctaaat ctcagcgcca agagtagttc 900

cctcaacacc agtattctgc tcagcagtga cctattcaat caggaagggg gaattgtaac 960

agccaactat ggctttgatg gttatatgg 989

<210> 42
<211> 989
<212> ADN
<213> Synechocystis PCC6803

<400> 42

```
ccggtatgga tggcaccgat gcggaatccc aacagattgc ctttgacaac aatgtggcct       60
ggaataacct gggggatttg tccaccacca cccaacgggc ctacacttcg gctattagca      120
cagacacagt gcagagtgtt tatggcgtta atctggaaaa aaacgataac attcccattg      180
tttttgcgtg gcccattttt cccaccaccc ttaatcccac agattttcag gtaatgctta      240
acacgggggа aattgtcacc ccggtgatcg cctctttgat tcccaacagt gaatacaacg      300
aacggcaaac ggtagtaatt acgggcaatt ttggtaatcg tttaacccca ggcacggagg      360
gagcgattta tcccgtttcc gtaggcacag tgttggacag tactcctttg gaaatggtgg      420
gacccaacgg cccggtcagt gcggtgggta ttaccattga tagtctcaac ccctacgtgg      480
ccggcaatgg tcccaaaatt gtcgccgcta agttagaccg cttcagtgac ctgggggaag      540
gggctcccct ctggttagcc accaatcaaa ataacagtgg cggggattta tatggagacc      600
aagcccaatt tcgtttgcga atttacacca gcgccggttt ttcccccgat ggcattgcca      660
gtttactacc cacagaattt gaacggtatt ttcaactcca agcggaagat attacgggac      720
ggacagttat cctaacccaa actggtgttg attatgaaat tcccggcttt ggtctggtgc      780
aggtgttggg gctggcggat ttggccgggg ttcaggacag ctatgacctg acttacatcg      840
aagatcatga caactattac gacattatcc tcaaagggga cgaagccgca gttcgccaaa      900
ttaagagggt tgctttgccc tccgaagggg attattcggc ggtttataat cccggtggcc      960
ccggcaatga tccagagaat ggtccccca                                        989
```

<210> 43
<211> 1089
<212> ADN
<213> Cuphea carthagenensis

<400> 43

atggcgaacg gtagcgctgt ctctctgaag agcggctcct tgaatacgca agaggacact 60

tcttcttccc caccgccacg cgcgttcatc aaccaattac ccgactggtc catgttattg 120

acggcgatta ccactgtctt tgttgccgca gagaaacagt ggactatgtt agaccgcaag 180

agcaagcgct ccgatatgtt agtggattct tttggcatgg aacgcattgt gcaggatggc 240

ttagtgtttc gtcaatcttt tagcattcgt tcttatgaaa tcggtgcaga tcgtcgtgca 300

tccattgaaa ccttaatgaa ccatctgcag gaaactagct tgaatcattg caaatccatt 360

cgcttgttga atgagggttt tggtcgcacc cccgagatgt gcaaacgtga cttgatctgg 420

gtggttaccc gcatgcacat catggtcaac cgctacccta cctggggtga taccgttgag 480

attaacactt gggtttccca aagcggcaag aatggtatgg gtcgtgattg gctgatttcc 540

gactgtaata ccggcgaaat cctgatccgc gcgacgtctg catgggcgat gatgaaccaa 600

aagacccgtc gtctgtctaa actgccttac gaagtcagcc aagagattgc tccgcacttc 660

gtcgacagcc ctcccgtgat cgaggacggc gaccgtaagt tacacaagtt cgatgtgaaa 720

accggcgaca gcatccgtaa aggtttgact ccgcgttgga atgacttaga tgttaatcag 780

cacgttaaca acgttaagta tatcggctgg atcttagaga gcatgccgac cgaggtcttg 840

gaaactcatg aactgtgttt cttaactctg gagtatcgtc gcgagtgcgg tcgcgatagc 900

gtgctggaat ctgtgaccgc gatggatcct tctaatgaag gtggtcgctc ccactaccag 960

catttactgc gcttggagga cggtactgac atcgttaagg gccgcactga gtggcgtcca 1020

aagaatgccc ggaatattgg tgccattagt accggtaaaa ccagtaatgg taatcccgcc 1080

agttaataa 1089

<210> 44
<211> 361
<212> PRT
<213> Cuphea carthagenensis

<400> 44

```
Met Ala Asn Gly Ser Ala Val Ser Leu Lys Ser Gly Ser Leu Asn Thr
1               5                  10                  15

Gln Glu Asp Thr Ser Ser Ser Pro Pro Pro Arg Ala Phe Ile Asn Gln
            20                  25                  30

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

Ala Ala Glu Lys Gln Trp Thr Met Leu Asp Arg Lys Ser Lys Arg Ser
    50                  55                  60

Asp Met Leu Val Asp Ser Phe Gly Met Glu Arg Ile Val Gln Asp Gly
65                  70                  75                  80

Leu Val Phe Arg Gln Ser Phe Ser Ile Arg Ser Tyr Glu Ile Gly Ala
                85                  90                  95

Asp Arg Arg Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu Thr
            100                 105                 110

Ser Leu Asn His Cys Lys Ser Ile Arg Leu Leu Asn Glu Gly Phe Gly
        115                 120                 125

Arg Thr Pro Glu Met Cys Lys Arg Asp Leu Ile Trp Val Val Thr Arg
    130                 135                 140
```

```
Met His Ile Met Val Asn Arg Tyr Pro Thr Trp Gly Asp Thr Val Glu
145                 150                 155                 160

Ile Asn Thr Trp Val Ser Gln Ser Gly Lys Asn Gly Met Gly Arg Asp
                165                 170                 175

Trp Leu Ile Ser Asp Cys Asn Thr Gly Glu Ile Leu Ile Arg Ala Thr
            180                 185                 190

Ser Ala Trp Ala Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys Leu
        195                 200                 205

Pro Tyr Glu Val Ser Gln Glu Ile Ala Pro His Phe Val Asp Ser Pro
    210                 215                 220

Pro Val Ile Glu Asp Gly Asp Arg Lys Leu His Lys Phe Asp Val Lys
225                 230                 235                 240

Thr Gly Asp Ser Ile Arg Lys Gly Leu Thr Pro Arg Trp Asn Asp Leu
                245                 250                 255

Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu
            260                 265                 270

Glu Ser Met Pro Thr Glu Val Leu Glu Thr His Glu Leu Cys Phe Leu
        275                 280                 285

Thr Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Glu Ser
    290                 295                 300

Val Thr Ala Met Asp Pro Ser Asn Glu Gly Gly Arg Ser His Tyr Gln
305                 310                 315                 320

His Leu Leu Arg Leu Glu Asp Gly Thr Asp Ile Val Lys Gly Arg Thr
                325                 330                 335

Glu Trp Arg Pro Lys Asn Ala Arg Asn Ile Gly Ala Ile Ser Thr Gly
            340                 345                 350

Lys Thr Ser Asn Gly Asn Pro Ala Ser
        355                 360
```

<210> 45
<211> 42
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB462

<400> 45
ctttgatggt tatatggcct gaggctgaaa tgagctgttg ac 42

<210> 46
<211> 47
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB463

<400> 46
gatctgaagc ttgaaaggtt attaactggc gggattacca ttactgg 47

<210> 47
<211> 47
<212> PRT
<213> Secuencia artificial

<220>
<223> Cebador NB464

<400> 47

Gly Thr Ala Ala Thr Cys Cys Cys Gly Cys Cys Ala Gly Thr Thr Ala
1 5 10 15

Ala Thr Ala Ala Cys Cys Thr Thr Thr Cys Ala Ala Gly Cys Thr Thr
20 25 30

Cys Ala Gly Ala Thr Cys Ala Ala Thr Thr Cys Gly Cys Gly Cys
35 40 45

<210> 48
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB465

<400> 48
atatttttat cttgtggtcc ggaattcgcc cttctaagc 39

<210> 49
<211> 57
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB466

<400> 49
gggcgaattc cggaccacaa gataaaaata tatcatcatg aacaataaaa ctgtctg 57

<210> 50
<211> 34
<212> ADN
<213> Secuencia artificial

<220>

<223> Cebador NB467

<400> 50
ggtgccatcc ataccgggcc gccgtcccgt caag 34

<210> 51
<211> 6663
<212> ADN
<213> Secuencia artificial

<220>
<223> vector pSGI-NB19

<400> 51

ccccgtgct atgactagcg gcgatcgcca taccggccac gaccatttgc attggatccc 60

caacggcggc cacaacttcc atggcattga gatgcgggga atgatgttct agactctgac 120

gcaccaaagc caatttttgt tgatggttgc aatggggatg actactgttc actttgcccc 180

cagcgtcaat gcctagacct agcagtaccc ccagggctgt ggtagtgccc cccaccacgc 240

attcgcttag cactaagtaa ctttcggcat gttcctgggc taactgtgcg ccccactgca 300

aaccctgctg aaaaagatgc tccaccaggg ccaacggtaa cgcttgccct gtggaaagac 360

agcgggcggg ttgtccgtct agattgatga ctggcaccgc tgggggaatg ggtaaaccag 420

agttaaataa ataaaccgga gtatggaggg catccaccaa cgctttggtg atgaacactg 480

gggaaacccc agaaatgagg ggaggtaagg gataggttgc ccctgccgta gttcccttga 540

ttaaaaattc cgcatcggcg atcgccgtca attttcgatc agcgggggtt ttacccgccg 600

cagaaatgcc cggaattaaa ccagtttccg taaagcccaa cacacagaca aacaccggtg 660

gacagtggcc atggcgctca atccaggata aagcttggtc agactgggta taaactgtca 720

acatatttct gcaagagtgg gcccatgcac ccctatttgt ttatttttct aaatacattc 780

aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat acgaggagga 840

aaaaaaaatg cgctcacgca actggtccag aaccttgacc gaacgcagcg gtggtaacgg 900

cgcagtggcg gttttcatgg cttgttatga ctgtttttt gtacagtcta tgcctcgggc 960

atccaagcag caagcgcgtt acgccgtggg tcgatgtttg atgttatgga gcagcaacga 1020

tgttacgcag cagggcagtc gccctaaaac aaagttaggt ggctcaagta tgggcatcat 1080

tcgcacatgt aggctcggcc ctgaccaagt caaatccatg cgggctgctc ttgatctttt 1140

cggtcgtgag ttcggtgacg tagccaccta ctcccaacat cagccggact ccgattacct 1200

cgggaacttg ctccgtagta agacattcat cgcgcttgct gccttcgacc aagaagcggt 1260

tgttggcgct ctcgcggctt acgttctgcc caagtttgag cagccgcgta gtgagatcta 1320

tatctatgat ctcgcagtct ccggcgagca ccggaggcag ggcattgcca ccgcgctcat 1380

caatctcctc aagcatgagg ccaacgcgct tggtgcttat gtgatctacg tgcaagcaga 1440

ttacggtgac gatcccgcag tggctctcta tacaaagttg ggcatacggg aagaagtgat 1500

gcactttgat atcgacccaa gtaccgccac ctaagcagag cttacgctga gcgcgcagat 1560

cagttggaag aatttgtcca ctacgtgaaa ggcgagatca ccaaggtagt cggcaaataa 1620

tgtctaacaa ttcgttcaag ccgacgccgc ttcgcggcgc ggcttaactc aagcgttaga 1680

tgcactaagc acataattgc tcacagccaa actatcaggt caagtctgct tttattattt 1740

ttaagcgtgc ataataagcc ctacacaaat tgggagatat atcatgaaag gctggctttt 1800

tcttgttatc gcaatagttg gcgaagtaat cgcaacatcc gcattaaaat ctagcgaggg 1860

ctttactaag ctgatccggt ggatgacctt ttgaatgacc tttaatagat tatattacta 1920

attaattggg gaccctagag gtccccttt ttattttaaa aatttttca caaaacggtt 1980

tacaagcata aagcttaggc cgtttaaacc aattgcggga agaaattaca gcttttgccg 2040

ctggcctaca gagtttagga gttacccccc atcaacacct ggccattttc gccgacaaca 2100

gcccccggtg gtttatcgcc gatcaaggca gtatgttggc tggagccgtc aacgccgtcc 2160

gttctgccca agcagagcgc caggaattac tctacatcct agaagacagc aacagccgta 2220

ctttaatcgc agaaaatcgg caaaccctaa gcaaattggc cctagatggc gaaaccattg 2280

acctgaaact aatcatcctc ctcaccgatg aagaagtggc agaggacagc gccattcccc 2340

aatataactt tgcccaggtc atggccctag gggccggcaa aatccccact cccgttcccc 2400

gccaggaaga agatttagcc accctgatct acacctccgg caccacagga caacccaaag 2460

gggtgatgct cagccacggt aatttattgc accaagtacg ggaattggat tcggtgatta 2520

ttccccgccc cggcgatcag gtgttgagca ttttgccctg ttggcactcc ctagaaagaa 2580

gcgccgaata ttttcttctt tcccggggct gcacgatgaa ctacaccagc attcgccatt 2640

tcaaggggga tgtgaaggac attaaacccc atcacattgt cggtgtgccc cggctgtggg 2700

aatccctcta cgaaggggta caaaaaacgt tccgagggcg aattctgcag atatccatca 2760

cactggcggc cgctcgagca tgcatctaga gggcccaatt cgccctatag tgagtcgtat 2820

tacaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 2880

cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 2940

accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggacgcg ccctgtagcg 3000

gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg 3060

ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc 3120

cccgtcaagc tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc 3180

tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga 3240

cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa 3300

ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg attttgccga 3360

tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca 3420

aaattcaggg cgcaagggct gctaaaggaa gcggaacacg tagaaagcca gtccgcagaa 3480

acggtgctga ccccggatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag 3540

cgcaaagaga aagcaggtag cttgcagtgg gcttacatgg cgatagctag actgggcggt 3600

tttatggaca gcaagcgaac cggaattgcc agctggggcg ccctctggta aggttgggaa 3660

gccctgcaaa gtaaactgga tggctttctt gccgccaagg atctgatggc gcaggggatc 3720

aagatctgat caagagacag gatgaggatc gtttcgcatg attgaacaag atggattgca 3780

cgcaggttct ccggccgctt gggtggagag gctattcggc tatgactggg cacaacagac 3840

aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt 3900

tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc 3960

gtggctggcc acgacgggcg ttccttgcgc agctgtgctc gacgttgtca ctgaagcggg 4020

aagggactgg ctgctattgg gcgaagtgcc ggggcaggat ctcctgtcat cccaccttgc 4080

tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc 4140

ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcggat 4200

ggaagccggt cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc 4260

cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg tcgtgaccca 4320

tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga 4380

ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta cccgtgatat 4440

tgctgaagag cttggcggcg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc 4500

tcccgattcg cagcgcatcg ccttctatcg ccttcttgac gagttcttct gaattgaaaa 4560

aggaagagta tgagtattca acatttccgt gtcgccctta ttccctttt tgcggcattt 4620

tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag 4680

ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt 4740

tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg 4800

gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag 4860

aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta 4920

agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg 4980

acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta 5040

actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac 5100

accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt 5160

actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca 5220

cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag 5280

cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 5340

gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag 5400

ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt 5460

tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat 5520

aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 5580

gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa 5640

acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 5700

tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag 5760

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 5820

atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 5880

agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag 5940

cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa 6000

agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga 6060

acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc 6120

gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc 6180

ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt 6240

gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt 6300

gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag 6360

gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa 6420

tgcagctggc acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat 6480

gtgagttagc tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg 6540

ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac 6600

gccaagcttg gtaccgagct cggatccact agtaacggcc gccagtgtgc tggaattcgc 6660

cct 6663

<210> 52
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB1

<400> 52
ctcgagcccc cgtgctatga ctagc 25

<210> 53
<211> 28
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador NB9

<400> 53
ctcgagcccg gaacgttttt tgtacccc 28

<210> 54
<211> 362
<212> PRT
<213> Cuphea decandra

<400> 54

```
Met Gly Ile Asn Gly Ser Ser Val Gly Leu Lys Ser Gly Ser Leu Lys
1               5                  10                  15

Thr Gln Glu Asp Thr Pro Ser Ser Pro Pro Pro Arg Thr Phe Ile Asn
            20                  25                  30

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

Leu Ala Ala Glu Lys Gln Trp Met Met Leu Asp Trp Lys Pro Lys Arg
    50                  55                  60

Pro Asp Met Leu Val Asp Pro Phe Gly Leu Gly Arg Ile Val Gln Asp
65                  70                  75                  80

Gly Leu Val Phe Arg Gln Asn Phe Ser Ile Arg Ser Tyr Glu Ile Gly
                85                  90                  95

Ala Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu
            100                 105                 110

Thr Ala Leu Asn His Val Lys Ser Ala Gly Leu Leu Asn Asp Gly Phe
        115                 120                 125

Gly Arg Thr Pro Glu Met Tyr Lys Arg Asp Leu Ile Trp Val Val Ala
    130                 135                 140

Lys Met Gln Val Met Val Asn Arg Tyr Pro Thr Trp Gly Asp Thr Val
145                 150                 155                 160
```

```
Glu Val Asn Thr Trp Val Ala Lys Ser Gly Lys Asn Gly Met Arg Arg
                165                 170                 175

Asp Trp Leu Ile Ser Asp Cys Asn Thr Gly Glu Ile Leu Thr Arg Ala
            180                 185                 190

Ser Ser Val Trp Val Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys
        195                 200                 205

Ile Pro Asp Glu Val Arg His Glu Ile Glu Pro His Phe Val Asp Ser
    210                 215                 220

Pro Pro Val Ile Glu Asp Asp Asp Arg Lys Leu Pro Lys Leu Asp Glu
225                 230                 235                 240

Lys Thr Ala Asp Ser Ile Arg Lys Gly Leu Thr Pro Arg Trp Asn Asp
                245                 250                 255

Leu Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile
            260                 265                 270

Leu Glu Ser Thr Pro Gln Glu Val Leu Glu Thr Gln Glu Leu Cys Ser
        275                 280                 285

Leu Thr Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Glu
    290                 295                 300

Ser Leu Thr Ala Val Asp His Ser Gly Lys Gly Ser Gly Ser Asn Phe
305                 310                 315                 320

Gln His Leu Leu Arg Leu Glu Asp Gly Gly Glu Ile Val Lys Gly Arg
                325                 330                 335

Thr Glu Trp Arg Pro Lys Asn Ala Val Ile Asn Gly Ala Val Ala Pro
            340                 345                 350

Gly Glu Thr Ser Pro Gly Asn Ser Val Ser
        355                 360
```

<210> 55
<211> 356
<212> PRT
<213> Cuphea paucipetela

<400> 55

```
Met Ala Asn Gly Ser Ala Val Ser Leu Lys Asp Gly Ser Leu Glu Thr
1               5                   10                  15
```

```
Gln Glu Gly Thr Ser Ser Ser Ser His Pro Pro Arg Thr Phe Ile Asn
            20                  25                  30

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

Val Ala Ala Glu Lys Gln Trp Thr Met Leu Asp Arg Lys Ser Lys Arg
    50                  55                  60

Pro Asp Met Leu Val Glu Pro Phe Val Gln Asp Gly Val Ser Phe Arg
65                  70                  75                  80

Gln Ser Phe Ser Ile Arg Ser Tyr Glu Ile Gly Ala Asp Arg Thr Ala
                85                  90                  95

Ser Ile Glu Thr Leu Met Asn Ile Phe Gln Glu Thr Ser Leu Asn His
            100                 105                 110

Cys Lys Ser Leu Gly Leu Leu Asn Asp Gly Phe Gly Arg Thr Pro Glu
        115                 120                 125

Met Cys Lys Arg Asp Leu Ile Trp Val Val Thr Lys Met Gln Ile Glu
    130                 135                 140

Val Asn Arg Tyr Pro Thr Trp Gly Asp Thr Ile Glu Val Thr Thr Trp
145                 150                 155                 160

Val Ser Glu Ser Gly Lys Asn Gly Met Ser Arg Asp Trp Leu Ile Ser
                165                 170                 175

Asp Cys His Thr Gly Glu Ile Leu Ile Arg Ala Thr Ser Val Trp Ala
            180                 185                 190

Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys Ile Pro Asp Glu Val
        195                 200                 205

Arg Gln Glu Ile Val Pro Tyr Phe Val Asp Ser Ala Pro Val Ile Glu
    210                 215                 220

Asp Asp Arg Lys Leu His Lys Leu Asp Val Lys Thr Gly Asp Ser Ile
225                 230                 235                 240

Arg Asn Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln His
                245                 250                 255

Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu Lys Ser Val Pro Thr
```

```
                260                 265                 270

Glu Val Phe Val Thr Gln Glu Leu Cys Gly Leu Thr Leu Glu Tyr Arg
        275                 280                 285

Arg Glu Cys Arg Arg Asp Ser Val Leu Glu Ser Val Thr Ala Met Asp
    290                 295                 300

Pro Ser Lys Glu Gly Asp Arg Ser Leu Tyr Gln His Leu Leu Arg Leu
305                 310                 315                 320

Glu Asn Gly Ala Asp Ile Ala Leu Gly Arg Thr Glu Trp Arg Pro Lys
                325                 330                 335

Asn Ala Gly Thr Asn Gly Ala Ile Ser Thr Thr Lys Thr Ser Pro Gly
            340                 345                 350

Asn Ser Val Ser
        355
```

<210> 56
<211> 267
<212> PRT
<213> Elaeis oleifera

<400> 56

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

Phe Leu Ala Ala Glu Lys Gln Trp Thr Leu Leu Asp Trp Lys Pro Arg
            20                  25                  30

Arg Pro Asp Met Leu Thr Asp Ala Phe Ser Leu Gly Glu Ile Val Gln
        35                  40                  45

Asp Gly Leu Val Phe Arg Gln Asn Phe Ser Ile Arg Ser Tyr Glu Ile
    50                  55                  60

Gly Ala Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln
65                  70                  75                  80

Glu Thr Val Leu Asn His Val Arg Asn Ala Gly Leu Leu Gly Asp Gly
                85                  90                  95

Phe Gly Ala Thr Pro Glu Met Ser Lys Arg Asn Leu Ile Trp Val Val
            100                 105                 110

Thr Lys Met Gln Val Leu Ile Glu His Tyr Pro Ser Trp Gly Asp Val
```

```
        115                 120                 125

Val Glu Val Asp Thr Trp Val Gly Ala Ser Gly Lys Asn Gly Met Arg
    130                 135                 140

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

Ala Thr Ser Ile Trp Val Met Met Asn Lys His Thr Arg Lys Leu Ser
                165                 170                 175

Lys Met Pro Glu Glu Val Arg Ala Glu Ile Gly Pro Tyr Phe Met Glu
            180                 185                 190

His Ala Thr Ile Val Asp Glu Asp Ser Gly Lys Leu Pro Lys Leu Asp
        195                 200                 205

Asp Asp Thr Ala Asp Tyr Ile Lys Trp Gly Leu Thr Pro Arg Trp Ser
    210                 215                 220

Asp Leu Asp Val Asn Gln His Val Asn His Val Lys Tyr Ile Gly Trp
225                 230                 235                 240

Ile Leu Glu Ser Ala Pro Ile Ser Ile Leu Glu Asn His Asn Trp Ala
                245                 250                 255

Ser Leu Ile Leu Asp Thr Gly Arg Asn Trp Asn
            260                 265
```

<210> 57
<211> 417
<212> PRT
<213> Elaeis guineensis

<400> 57

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

Ser Pro Thr Ala Ser Ala Lys Ala Ser Lys Thr Ile Gly Glu Gly Ser
            20                  25                  30

Glu Asn Leu Asn Val Arg Gly Ile Ile Ala Lys Pro Thr Ser Ser Ser
        35                  40                  45

Ala Ala Lys Gln Gly Lys Val Met Ala Gln Ala Val Pro Lys Ile Asn
    50                  55                  60

Gly Ala Lys Val Gly Leu Lys Ala Glu Ser Gln Lys Ala Glu Glu Asp
```

65 70 75 80

Ala Ala Pro Ser Ser Ala Pro Arg Thr Phe Tyr Asn Gln Leu Pro Asp
85 90 95

Trp Ser Val Leu Leu Ala Ala Val Thr Thr Ile Phe Leu Ala Ala Glu
100 105 110

Lys Gln Trp Thr Leu Leu Asp Trp Lys Pro Arg Arg Pro Asp Met Leu
115 120 125

Thr Gly Ala Phe Ser Leu Gly Lys Ile Val Gln Asp Gly Leu Val Phe
130 135 140

Arg Gln Asn Phe Ser Ile Arg Ser Tyr Glu Ile Gly Ala Asp Arg Thr
145 150 155 160

Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu Thr Ala Leu Asn
165 170 175

His Val Arg Asn Ala Gly Leu Leu Gly Asp Gly Phe Gly Ala Thr Pro
180 185 190

Glu Met Ser Lys Arg Asn Leu Ile Trp Val Val Thr Lys Met Gln Val
195 200 205

Leu Ile Glu His Tyr Pro Ser Trp Gly Asp Val Val Glu Val Asp Thr
210 215 220

Trp Val Gly Ala Ser Gly Lys Asn Gly Met Arg Arg Asp Trp His Val
225 230 235 240

Arg Asp Tyr Arg Thr Gly Gln Thr Ile Leu Arg Ala Thr Ser Ile Trp
245 250 255

Val Met Met Asp Lys His Thr Arg Lys Leu Ser Lys Met Pro Glu Glu
260 265 270

Val Arg Ala Glu Ile Gly Pro Tyr Phe Met Glu His Ala Ala Ile Val
275 280 285

Asp Glu Asp Ser Arg Lys Leu Pro Lys Leu Asp Asp Asp Thr Ala Asp
290 295 300

Tyr Ile Lys Trp Gly Leu Thr Pro Arg Trp Ser Asp Leu Asp Val Asn
305 310 315 320

```
Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Ala
                325                 330                 335

Pro Ile Ser Ile Leu Glu Asn His Glu Leu Ala Ser Met Thr Leu Glu
            340                 345                 350

Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Gln Ser Leu Thr Ala
        355                 360                 365

Val Ala Asn Asp Cys Thr Gly Gly Leu Pro Glu Ala Ser Ile Glu Cys
    370                 375                 380

Gln His Leu Leu Gln Leu Glu Cys Gly Ala Glu Ile Val Arg Gly Arg
385                 390                 395                 400

Thr Gln Trp Arg Pro Arg Arg Ala Ser Gly Pro Thr Ser Ala Gly Ser
                405                 410                 415

Ala
```

REIVINDICACIONES

1. Una cianobacteria recombinante transformada con una secuencia de ácido nucleico no originario que codifica una acil-ACP tioesterasa procedente de una planta superior, unida operativamente a un promotor heterólogo y una secuencia de ácido nucleico no originario que codifica una ácido lisofosfatídico aciltransferasa (LPAAT) procedente de una cianobacteria unida operativamente a un promotor heterólogo, en donde la cianobacteria produce al menos un ácido graso libre o al menos un derivado de ácido graso, y en donde la LPAAT tiene una identidad de al menos 85 % con la SEQ ID NO: 2, y la tioesterasa tiene una identidad de al menos 85 % con la SEQ ID NO: 44.

2. La cianobacteria recombinante de la reivindicación 1, en donde la LPAAT transfiere grupos acilo sobre sustratos C16 o C18.

3. La cianobacteria recombinante de la reivindicación 1 o 2, en donde la cianobacteria recombinante comprende además una construcción para regular negativamente la expresión de un gen de acil-ACP sintetasa o de un gen de acil-CoA sintetasa endógeno, en donde la construcción es una construcción antisentido, una construcción de micro ARN, una construcción ribozimática o una construcción de alteración génica.

4. La cianobacteria recombinante de una cualquiera de las reivindicaciones anteriores, en donde la cianobacteria recombinante comprende además una construcción para regular negativamente la expresión de un gen de acil-ACP sintetasa o de un gen de acil-CoA sintetasa endógeno, en donde la construcción es una construcción de ARN antisentido, una construcción de microARN o una construcción ribozimática.

5. La cianobacteria recombinante de una cualquiera de las reivindicaciones 1 a 4, en donde la cianobacteria recombinante es del género *Agmenellum, Anabaena, Anabaenopsis, Anacystis, Aphanizomenon, Arthrospira, Asterocapsa, Borzia, Calothrix, Chamaesiphon, Chlorogloeopsis, Chroococcidiopsis, Chroococcus, Crinalium, Cyanobacterium, Cyanobium, Cyanocystis, Cyanospira, Cyanothece, Cylindrospermopsis, Cylindrospermum, Dactylococcopsis, Dermocarpella, Fischerella, Fremyella, Geitleria, Geitlerinema, Gloeobacter, Gloeocapsa, Gloeothece, Halospirulina, Iyengariella, Leptolyngbya, Limnothrix, Lyngbya, Microcoleus, Microcystis, Myxosarcina, Nodularia, Nostoc, Nostochopsis, Oscillatoria, Phormidium, Planktothrix, Pleurocapsa, Prochlorococcus, Prochloron, Prochlorothrix, Pseudanabaena, Rivularia, Schizothrix, Scytonema, Spirulina, Stanieria, Starria, Stigonema, Symploca, Synechococcus, Synechocystis, Tolypothrix, Trichodesmium, Tychonema* o *Xenococcus.*

6. La cianobacteria recombinante de una cualquiera de las reivindicaciones anteriores, en donde la cianobacteria produce: al menos un 10 % más de producto de ácido graso en comparación con una cianobacteria de otra manera idéntica que carece de la secuencia de ácido nucleico no originario que codifica una ácido lisofosfatídico aciltransferasa (LPAAT), y opcionalmente en donde el producto de ácido graso es uno o más de un ácido graso, un aldehído graso, un alcohol graso, un éster de ácido graso, un éster de cera o un hidrocarburo.

7. La cianobacteria recombinante de la reivindicación 6, en donde el producto de ácido graso es un producto de ácido graso C8-C24, preferentemente un producto de ácido graso C12-C18 o un producto de ácido graso C12-C16, en donde, opcionalmente, el producto de ácido graso está saturado.

8. Un método de producción de un producto de ácido graso que comprende cultivar la cianobacteria recombinante de una cualquiera de las reivindicaciones 1-7 en condiciones en las que se produce al menos un producto de ácido graso, en donde el producto de ácido graso es uno o más de un ácido graso libre, un aldehído graso, un alcohol graso, un éster de ácido graso, un éster de cera, o un hidrocarburo, y en donde el producto de ácido graso se secreta en el medio.

9. El método de la reivindicación 8, en donde la cianobacteria recombinante se cultiva de manera fotoautótrofa, y en donde el medio de cultivo no incluye una cantidad sustancial de carbono reducido.

10. El método de la reivindicación 9, en donde se induce la expresión de una, o ambas, de la secuencia de ácido nucleico no originario que codifica una LPAAT y de la secuencia de ácido nucleico no originario que codifica una tioesterasa.

11. El procedimiento de una cualquiera de las reivindicaciones 8-10, en donde al menos un producto de ácido graso producido es un ácido graso libre C8-C24, un alcohol graso o un aldehído graso.

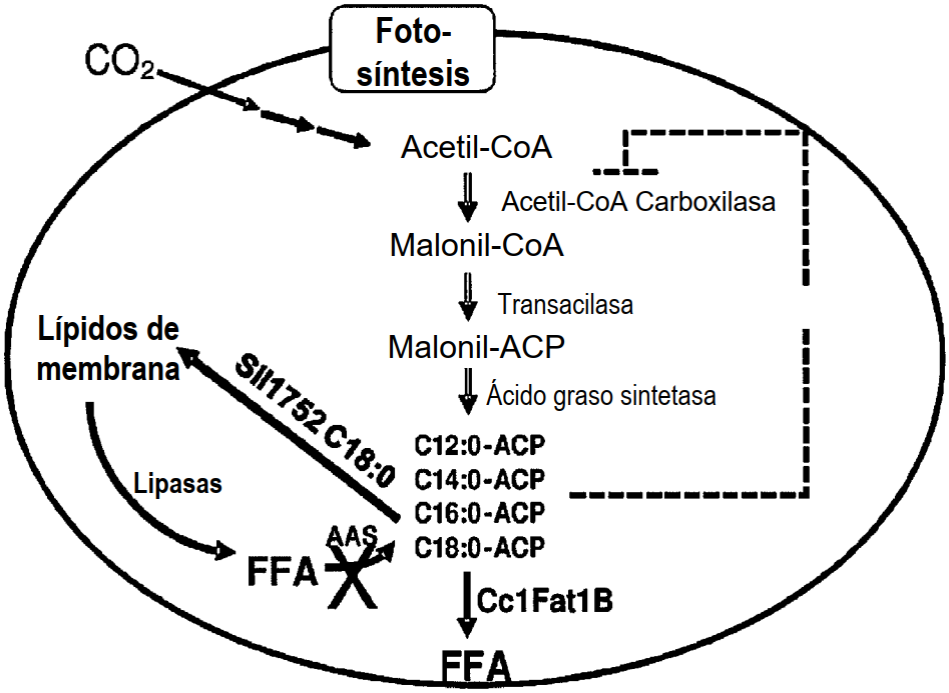
Foto-
síntesis
CO_2
Acetil-CoA
Acetil-CoA Carboxilasa
Malonil-CoA
Transacilasa
Malonil-ACP
Ácido graso sintetasa
Lípidos de
membrana
Sll1752 C18:0
C12:0-ACP
C14:0-ACP
C16:0-ACP
C18:0-ACP
Lipasas
AAS
FFA
Cc1Fat1B
FFA

Figura 1

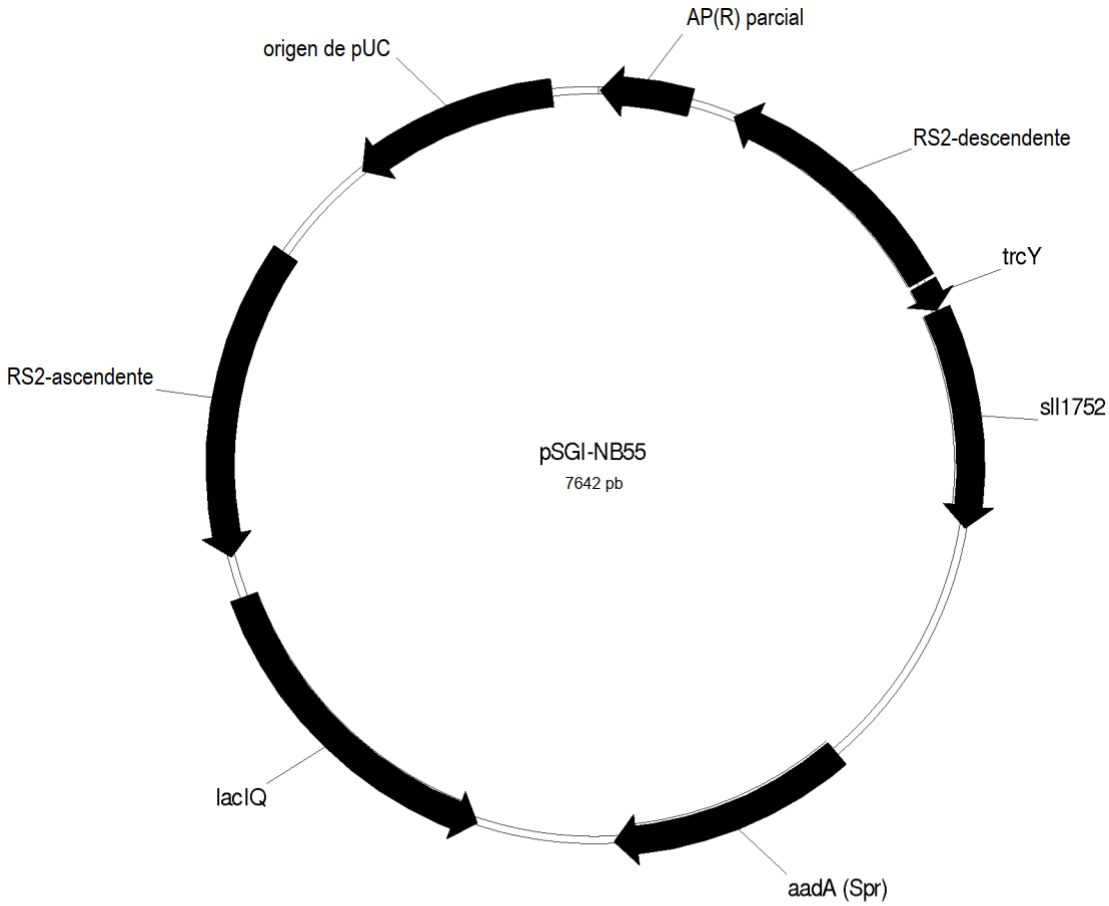
AP(R) parcial
origen de pUC
RS2-descendente
trcY
RS2-ascendente
sll1752
pSGI-NB55
7642 pb
lacIQ
aadA (Spr)

Figura 2

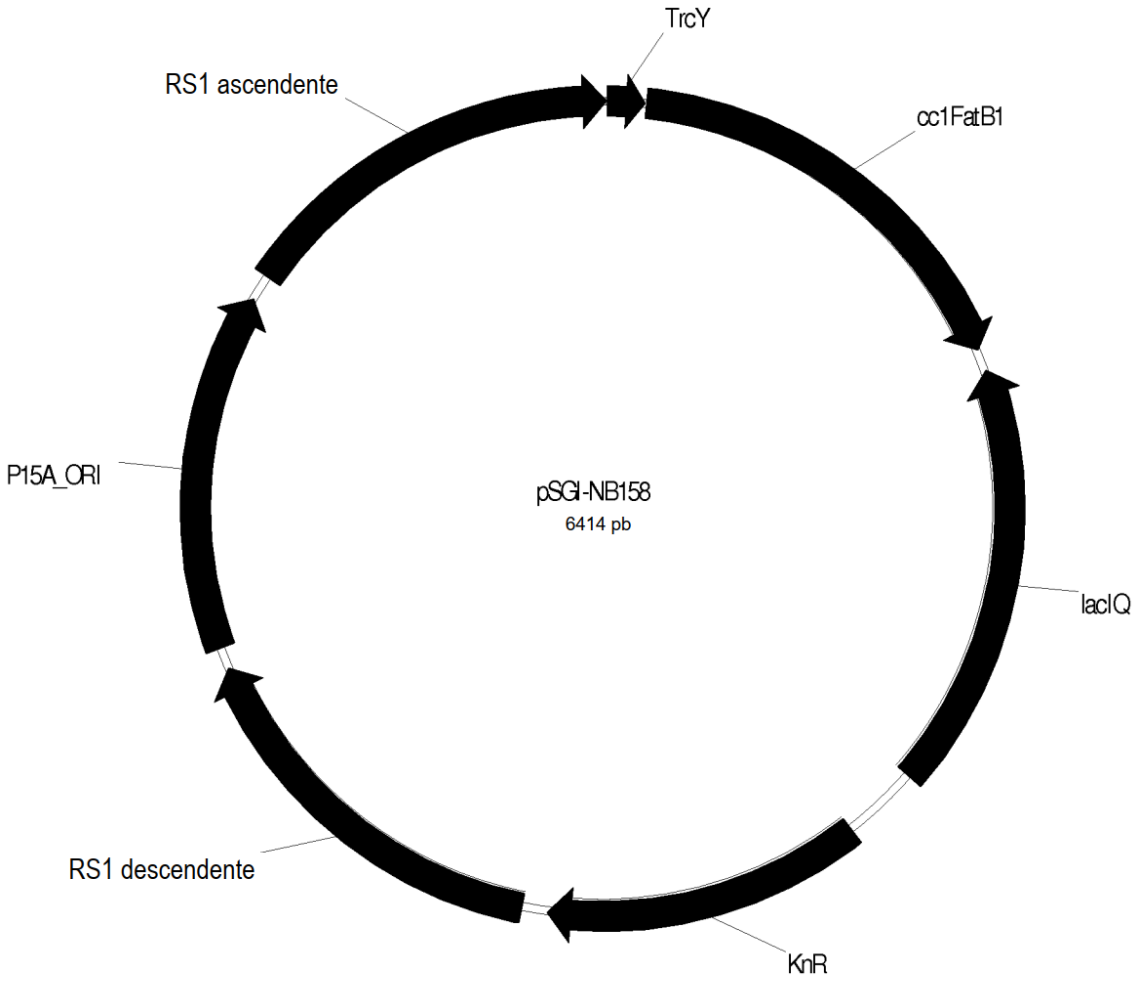
TrcY
RS1 ascendente
cc1FatB1
P15A_ORI
pSGI-NB158
6414 pb
lacIQ
RS1 descendente
KnR

Figura 3

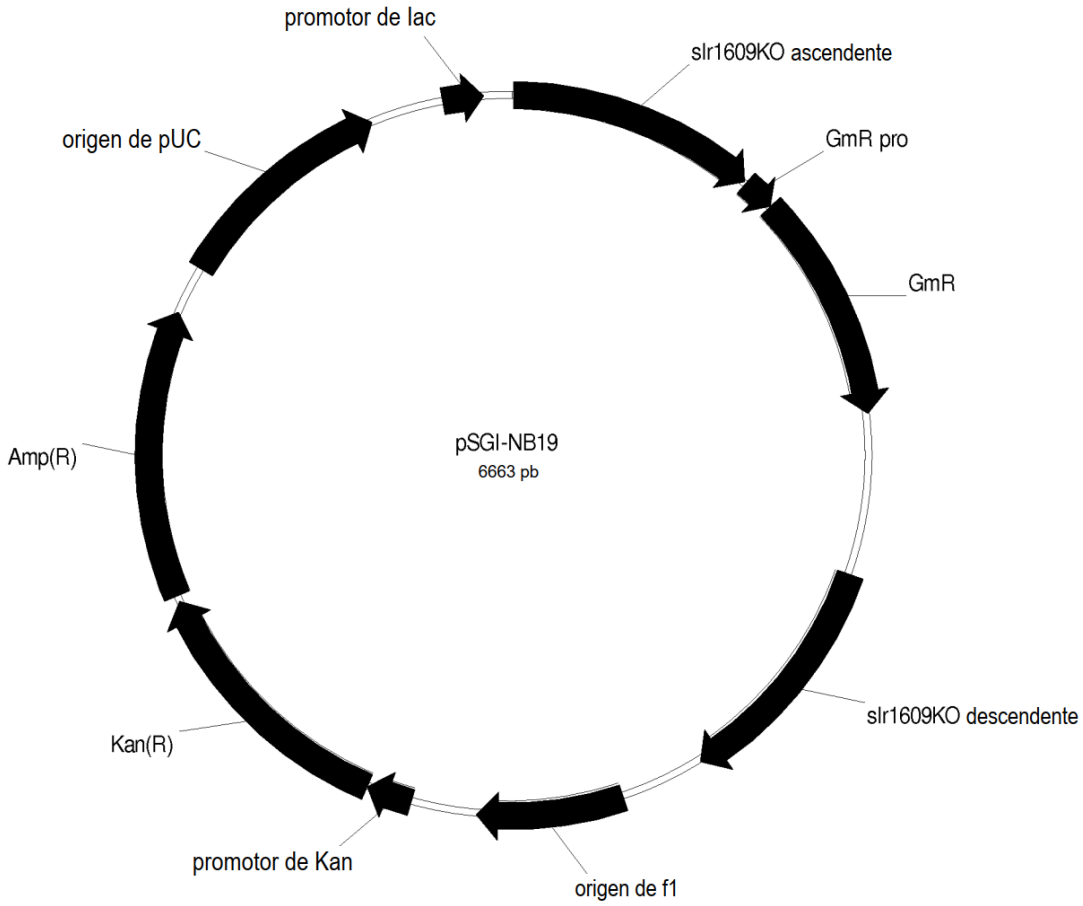
promotor de lac
slr1609KO ascendente
GmR pro
GmR
slr1609KO descendente
origen de f1
promotor de Kan
Kan(R)
Amp(R)
origen de pUC
pSGI-NB19
6663 pb

Figura 4

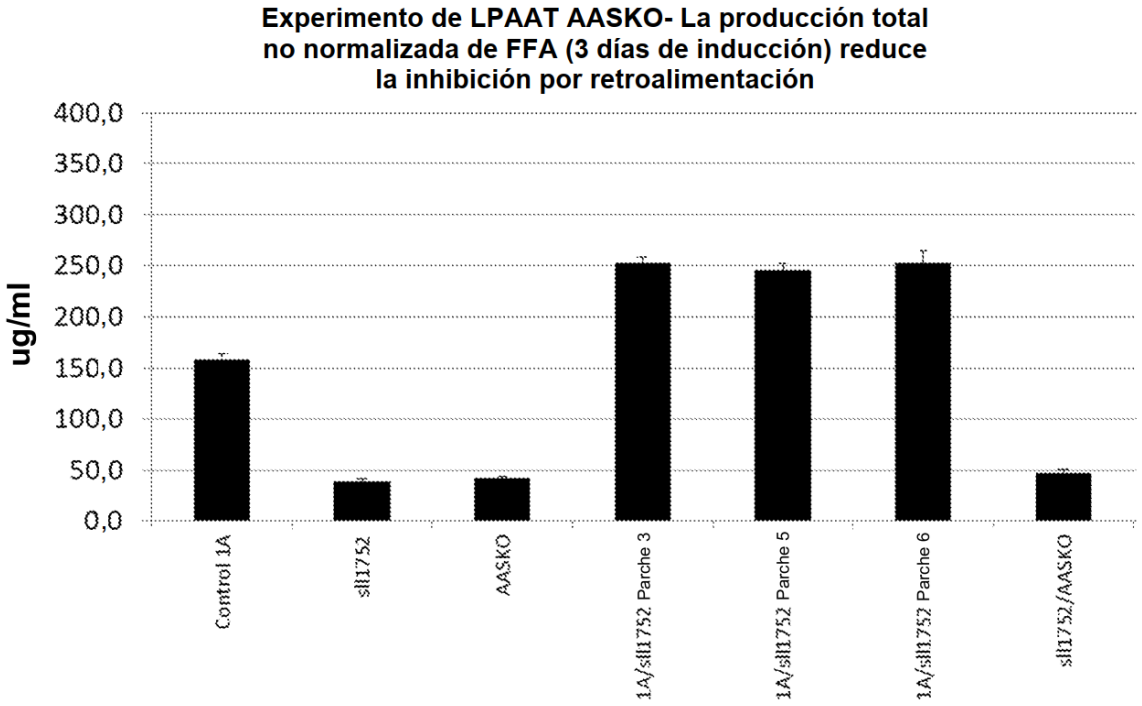
Experimento de LPAAT AASKO- La producción total no normalizada de FFA (3 días de inducción) reduce la inhibición por retroalimentación
ug/ml
400,0
350,0
300,0
250,0
200,0
150,0
100,0
50,0
0,0
Control 1A
slll1752
AASKO
1A/slll1752 Parche 3
1A/slll1752 Parche 5
1A/slll1752 Parche 6
slll1752/AASKO

Figura 5

Cepa	C8:0	C10:0	C12:0	C14:0	C16:0	C16:1 cis9	C18:0	C18:1 cis9	C18:2 cis9,12	C18:3 cis6,9,12	C18:3 cis9,12,15
Control 1A Rep 1	1,37	0,82	22,97	75,88	85,21	2,31	16,63	16,15	5,77	4,57	3,90
Control 1A Rep 2	0,90	0,75	23,61	78,32	86,03	3,71	16,67	20,28	6,82	4,32	3,80
Control 1A Rep 3	1,13	0,62	22,28	76,91	81,96	2,51	15,45	16,56	6,69	4,23	3,28
sll1752 Rep 1	0,53	0,00	0,64	1,77	13,46	1,40	14,12	16,45	4,87	1,33	3,22
sll1752 Rep 2	0,46	0,00	0,00	1,28	13,18	1,62	14,34	16,74	4,42	0,00	3,19
sll1752 Rep 3	0,68	0,00	0,00	1,93	13,71	1,45	16,29	18,22	4,71	1,58	3,00
AASKO Rep 1	0,41	0,00	0,00	1,72	14,43	2,02	14,27	18,23	5,75	3,08	3,16
AASKO Rep 2	0,79	0,00	0,00	1,38	14,90	2,07	13,42	17,75	5,77	2,81	3,24
AASKO Rep 3	0,62	0,00	0,00	1,41	15,21	2,64	13,37	18,47	4,92	2,78	3,43
1A/sll1752 Clon 3, Rep 1	1,10	0,85	28,51	112,56	155,83	3,19	30,77	22,35	13,17	5,57	3,35
1A/sll1752 Clon 3, Rep 2	0,93	0,76	30,64	111,08	159,53	3,98	31,21	21,88	12,38	5,80	3,22
1A/sll1752 Clon 3, Rep 3	0,97	0,76	31,48	114,99	156,71	3,23	29,84	22,48	12,74	4,98	3,66
1A/sll1752 Clon 5, Rep 1	1,05	0,77	29,86	111,62	155,21	2,90	30,53	21,05	12,36	5,69	4,04
1A/sll1752 Clon 5, Rep 2	1,04	0,60	29,26	104,85	150,17	2,40	30,17	20,72	12,05	5,27	3,39
1A/sll1752 Clon 5, Rep 3	0,98	0,88	29,08	109,42	154,44	2,58	31,04	22,16	12,59	6,10	3,87
1A/sll1752 Clon 6, Rep 1	0,91	0,78	31,26	108,22	145,40	2,82	28,43	20,77	12,76	6,40	3,21
1A/sll1752 Clon 6, Rep 2	1,21	1,07	30,79	113,20	160,87	2,84	32,69	21,68	12,60	5,81	4,06
1A/sll1752 Clon 6, Rep 3	1,08	0,71	32,38	117,78	159,10	2,94	32,07	21,88	12,81	5,69	4,40
sll1752/AASKO Rep 1	0,59	0,42	0,80	2,20	17,74	3,13	8,25	16,86	8,10	4,29	3,05
sll1752/AASKO Rep 2	0,56	0,00	0,78	2,07	19,08	3,23	10,58	20,58	9,04	5,82	2,83
sll1752/AASKO Rep 3	0,57	0,41	0,77	1,79	17,98	3,33	9,88	19,36	8,17	5,25	3,52

Figura 6