

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 808 287**

15 Folleto corregido: T3

Texto afectado: Lista de secuencias

48 Fecha de publicación de la corrección: 05.03.2021

51 Int. Cl.:

C07K 14/195 (2006.01)

C07K 14/245 (2006.01)

C12P 5/00 (2006.01)

C12P 7/64 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA CORREGIDA

T9

86 Fecha de presentación y número de la solicitud internacional: **02.04.2012** **PCT/US2012/031881**

87 Fecha y número de publicación internacional: **15.11.2012** **WO12154329**

96 Fecha de presentación y número de la solicitud europea: **02.04.2012** **E 12719914 (9)**

97 Fecha y número de publicación de la concesión europea: **27.05.2020** **EP 2906587**

54 Título: **Métodos y composiciones para la producción mejorada de ácidos grasos y derivados de los mismos**

30 Prioridad:

01.04.2011 US 201161470989 P

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

26.02.2021

73 Titular/es:

GENOMATICA, INC. (100.0%)
4757 Nexus Center Drive
San Diego, CA 92121, US

72 Inventor/es:

GREENFIELD, DEREK L.;
HOM, LOUIS G.;
SANCHEZ-RIERA, FERNANDO A.;
HU, ZHIHAO;
ARLAGADDA, VIKRANTH;
GROBAN, ELI S. y
FRYKMAN, SCOTT A.

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 808 287 T9

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

DESCRIPCIÓN

Métodos y composiciones para la producción mejorada de ácidos grasos y derivados de los mismos

5 Antecedentes de la invención

El petróleo crudo es un recurso natural limitado, que se encuentra en la Tierra en forma líquida, gaseosa y sólida. Aunque el petróleo crudo es un recurso valioso, se descubre y se extrae de la Tierra a un coste financiero y ambiental considerable. Además, en su forma natural, el petróleo crudo extraído de la Tierra tiene pocos usos comerciales. El petróleo crudo es una mezcla de hidrocarburos (por ejemplo, parafinas (o alcanos), olefinas (o alquenos), alquinos, naftenos (o cicloalcanos), compuestos alifáticos, compuestos aromáticos, etc.) de longitud y complejidad variables. Además, el petróleo crudo contiene otros compuestos orgánicos (por ejemplo, compuestos orgánicos que contienen nitrógeno, oxígeno, azufre, etc.) e impurezas (por ejemplo, azufre, sal, ácido, metales, etc.). Por tanto, el petróleo crudo debe refinarse y purificarse a un coste considerable antes de que pueda usarse comercialmente.

El petróleo crudo es también una fuente primaria de materias primas para la producción de productos petroquímicos. Las dos clases principales de materias primas derivadas del petróleo son olefinas de cadena corta (por ejemplo, etileno y propileno) y compuestos aromáticos (por ejemplo, isómeros de benceno y xileno). Estas materias primas derivan de hidrocarburos de cadena más larga del petróleo crudo mediante craqueo del mismo a un coste considerable usando una diversidad de métodos, tales como craqueo catalítico, craqueo por vapor o reforma catalítica. Estas materias primas pueden usarse para fabricar productos petroquímicos tales como monómeros, disolventes, detergentes y adhesivos, que de otra manera no pueden refinarse directamente a partir de petróleo crudo.

Los productos petroquímicos, a su vez, pueden usarse para fabricar productos químicos especiales, tales como plásticos, resinas, fibras, elastómeros, productos farmacéuticos, lubricantes y geles. Los productos químicos especiales que pueden producirse a partir de materias primas petroquímicas incluyen ácidos grasos, hidrocarburos (por ejemplo, de larga cadena, de cadena ramificada, saturados, insaturados, etc.), aldehídos grasos, alcoholes grasos, ésteres, cetonas, lubricantes, etc.

Debido a los desafíos inherentes que plantea el petróleo, existe la necesidad de una fuente de petróleo renovable que no sea necesario explorar, extraer, transportar a largas distancias o refinar sustancialmente como el petróleo crudo. También existe la necesidad de una fuente de petróleo renovable que pueda producirse económicamente sin crear el tipo de daño ambiental que produce la industria petrolera y la quema de combustibles a base de petróleo. Por razones similares, también existe la necesidad de una fuente renovable de productos químicos que normalmente derivan del petróleo.

Un método para producir petróleo renovable es mediante ingeniería de microorganismos para producir productos del petróleo renovables. Desde hace mucho tiempo se sabe que algunos microorganismos poseen una capacidad natural para producir productos del petróleo (por ejemplo, levadura para producir etanol). Más recientemente, el desarrollo de biotecnologías avanzadas ha hecho posible modificar mediante ingeniería metabólica un organismo para producir bioproductos y biocombustibles. Los bioproductos (por ejemplo, los productos químicos) y los biocombustibles (por ejemplo, el biodiésel) son alternativas renovables a los productos químicos y los combustibles a base de petróleo, respectivamente. Los bioproductos y los biocombustibles pueden derivar de fuentes renovables, tales como materia vegetal, materia animal y materia de desecho orgánica, conocidos colectivamente como biomasa.

Los biocombustibles pueden sustituir cualquier combustible a base de petróleo (por ejemplo, gasolina, diésel, combustible de aviación, aceite de calefacción, etc.) y ofrecen varias ventajas con respecto a los combustibles a base de petróleo. Los biocombustibles no requieren una exploración o extracción cara y arriesgada. Los biocombustibles pueden producirse localmente y por tanto no requieren transporte a largas distancias. Además, los biocombustibles pueden fabricarse directamente y requieren poco o ningún refinado adicional. Además, la combustión de biocombustibles provoca una menor carga sobre el medio ambiente puesto que la cantidad de emisiones nocivas (por ejemplo, gases de efecto invernadero, contaminación del aire, etc.) que se libera durante la combustión se reduce en comparación con la combustión de los combustibles a base de petróleo. Además, los biocombustibles mantienen un ciclo de carbono equilibrado porque los biocombustibles se producen a partir de biomasa, un recurso natural renovable. Aunque la combustión de biocombustibles libera carbono (por ejemplo, en forma de dióxido de carbono), este carbono se reciclará durante la producción de biomasa (por ejemplo, el cultivo de cultivos), equilibrando de este modo el ciclo del carbono, que no se consigue con el uso de combustibles a base de petróleo.

Los productos químicos derivados biológicamente ofrecen ventajas sobre los productos petroquímicos similares a las que los biocombustibles ofrecen sobre los combustibles a base de petróleo. En particular, los productos químicos derivados biológicamente pueden convertirse a partir de biomasa en el producto químico deseado directamente sin necesidad de un refinado exhaustivo, a diferencia de los productos petroquímicos, que deben producirse mediante el refinado de petróleo crudo para recuperar materias primas que después se procesan adicionalmente en el producto petroquímico deseado.

Los hidrocarburos tienen muchos usos comerciales. Por ejemplo, los alcanos de cadena más corta se usan como

combustibles. El metano y el etano son los componentes principales del gas natural. Se usan alcanos de cadena más larga (por ejemplo, de cinco a dieciséis carbonos) como combustibles para el transporte (por ejemplo, gasolina, diésel o combustible de aviación). Los alcanos que tienen más de dieciséis átomos de carbono son componentes importantes de aceites combustibles y aceites lubricantes. Los alcanos aún más largos, que son sólidos a temperatura ambiente, pueden usarse, por ejemplo, como cera de parafina. Se encuentran alcanos que contienen aproximadamente treinta y cinco carbonos en el betún, que se usa para el asfaltado de carreteras. Además, pueden craquearse alcanos de cadena más larga para producir hidrocarburos de cadena más corta comercialmente útiles.

Como los alcanos de cadena corta, los alquenos de cadena corta se usan en combustibles para el transporte. Se usan alquenos de cadena más larga en plásticos, lubricantes y lubricantes sintéticos. Además, se usan alquenos como materia prima para producir alcoholes, ésteres, plastificantes, tensioactivos, aminas terciarias, agentes de recuperación de aceite potenciada, ácidos grasos, tioles, epóxidos de anhídridos alqueniilsuccínicos, alcanos clorados, alquenos clorados, ceras, aditivos de combustible y reductores de flujo de arrastre.

Los ésteres tienen muchos usos comerciales. Por ejemplo, el biodiésel, un combustible alternativo, se compone de ésteres (por ejemplo, éster metílico de ácido graso, ésteres etílicos de ácidos grasos, etc.). Algunos ésteres de bajo peso molecular son volátiles con un olor agradable que los hace útiles como fragancias o agentes aromatizantes. Además, se usan ésteres como disolventes para lacas, pinturas y barnices. Además, algunas sustancias de origen natural, tales como ceras, grasas y aceites se componen de ésteres. Los ésteres también se usan como agentes suavizantes en resinas y plásticos, plastificantes, retardantes de llama y aditivos en la gasolina y el aceite. Además, pueden usarse ésteres en la fabricación de polímeros, películas, tejidos, colorantes y productos farmacéuticos.

Se usan aldehídos para producir muchos productos químicos especiales. Por ejemplo, se usan aldehídos para producir polímeros, resinas (por ejemplo, baquelita), colorantes, aromatizantes, plastificantes, perfumes, productos farmacéuticos y otros productos químicos, algunos de los cuales pueden usarse como disolventes, conservantes o desinfectantes. Además, determinados compuestos naturales y sintéticos, tales como vitaminas y hormonas, son aldehídos, y muchos azúcares contienen grupos aldehído. Los aldehídos grasos pueden convertirse en alcoholes grasos por reducción química o enzimática.

Los alcoholes grasos tienen muchos usos comerciales. Las ventas anuales mundiales de alcoholes grasos y sus derivados superan los 1.000 millones de dólares. Los alcoholes grasos de cadena más corta se usan en las industrias cosmética y alimentaria como emulsionantes, emolientes y espesantes. Debido a su naturaleza anfífila, los alcoholes grasos se comportan como tensioactivos no iónicos, que son útiles en productos de cuidado personal y para el hogar, tales como, por ejemplo, detergentes. Además, se usan alcoholes grasos en ceras, gomas, resinas, ungüentos y lociones farmacéuticas, aditivos de aceite lubricante, antiestáticos textiles y agentes de acabado, plastificantes, cosméticos, disolventes industriales y disolventes para grasas.

La Acil-CoA sintasa (ACS) esterifica los ácidos grasos libres a acil-CoA mediante un mecanismo de dos etapas. El ácido graso libre se convierte en primer lugar en un intermedio de acil-AMP (un adenilato) a través de la pirofosforólisis de ATP. El carbono de carbonilo activado del adenilato después se acopla con el grupo tiol de CoA, liberando AMP y el producto final de acil-CoA (Shockey *et al.*, *Plant. Physiol.*, 129: 1710-1722 (2002)).

FadR es un factor regulador clave que participa en la degradación de los ácidos grasos y las vías de biosíntesis de los ácidos grasos (Cronan *et al.*, *Mol. Microbiol.*, 29(4): 937-943 (1998)). La enzima ACS de *E. coli* FadD y la proteína de transporte de ácidos grasos FadI son componentes esenciales de un sistema de captación de ácidos grasos. FadI media el transporte de los ácidos grasos a la célula bacteriana, y FadD media la formación de ésteres de acil-CoA. Cuando no hay otra fuente de carbono disponible, los ácidos grasos exógenos son captados por bacterias y convertidos en ésteres de acil-CoA, que puede unirse al factor de transcripción FadR y desreprimir la expresión de los genes *fad* que codifican las proteínas responsables del transporte (FadI), la activación (FadD) y la β -oxidación de ácidos grasos (FadA, FadB, FadE y FadH). Cuando hay fuentes alternativas de carbono disponibles, las bacterias sintetizan ácidos grasos como acil-ACP, que se usan para la síntesis de fosfolípidos, pero no son sustratos para la β -oxidación. Por tanto, acil-CoA y acil-ACP son fuentes independientes de ácidos grasos que pueden dar lugar a diferentes productos finales (Caviglia *et al.*, *J. Biol. Chem.*, 279(12): 1163-1169 (2004)). El documento US 2010/242345 describe microorganismos modificados mediante ingeniería genética que producen productos de la vía biosintética de los ácidos grasos (derivados de ácidos grasos) y métodos para su uso.

Sigue existiendo la necesidad de métodos y composiciones para potenciar la producción de productos químicos derivados biológicamente, tales como ácidos grasos y derivados de ácidos grasos. La presente invención proporciona dichos métodos y composiciones. La invención proporciona adicionalmente productos derivados de los ácidos grasos y derivados de los mismos producidos mediante los métodos que se describen en el presente documento, tales como combustibles, tensioactivos y detergentes.

Breve resumen de la invención

La invención proporciona métodos mejorados para producir un ácido graso o un derivado de ácido graso del mismo en una célula hospedadora bacteriana. El método consiste en (a) proporcionar una célula hospedadora bacteriana

modificada por ingeniería genética que comprende un promotor heterólogo y/o un sitio de unión a ribosomas unido operativamente a una secuencia polinucleotídica que codifica un polipéptido FadR, provocando dicho promotor y/o sitio de unión a ribosomas la sobreexpresión del polipéptido FadR en dicha célula, (b) cultivar la célula hospedadora bacteriana modificada por ingeniería genética en un medio de cultivo en condiciones que permitan la producción de un ácido graso o un derivado de ácido graso del mismo y (c) aislar el ácido graso o un derivado de ácido graso del mismo de la célula hospedadora bacteriana modificada por ingeniería genética, en donde el ácido graso o derivado de ácido graso del mismo es un ácido graso, una acil-ACP, una acil-CoA, un aldehído graso, un alcohol de cadena corta, un alcohol de cadena larga, un alcohol graso, un hidrocarburo o un éster. Como resultado de este método, uno o más de entre el título, el rendimiento o la productividad del ácido graso o derivado de ácido graso producido por la célula hospedadora bacteriana modificada por ingeniería genética aumentan con respecto al de la célula hospedadora de tipo silvestre correspondiente.

También se desvelan ácidos grasos y derivados de ácidos grasos, tales como un acil-CoA, un aldehído graso, un alcohol de cadena corta, un alcohol de cadena larga, un alcohol graso, un hidrocarburo o un éster, producidos mediante los métodos de la invención. Se desvelan adicionalmente composiciones de biocombustible y composiciones de tensioactivo que comprenden un ácido graso o un derivado de ácido graso producidos mediante los métodos de la invención.

Breve descripción de las distintas vistas de los dibujos

La FIG. 1 es un cuadro de genes de ejemplo adecuados para su uso en la práctica de la invención. Los números de registro de polipéptidos y/o polinucleótidos son de la base de datos del Centro Nacional de Información Biotecnológica (NCBI, por sus siglas en inglés) y los números CE de enzimas son del Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (NC-IUBMB, por sus siglas en inglés).

La FIG. 2 es un gráfico de la producción de especies grasas en una cepa de *E. coli* de control (ALC310) o la cepa de inserción de transposón, D288.

La FIG. 3 es un diagrama que representa la ubicación de la inserción de transposón en la cepa D288.

La FIG. 4 es un gráfico de barras de títulos totales de especies grasas (FA, por sus siglas en inglés) de la biblioteca de expresión de las cepas de *E. coli* que tienen una expresión alterada de FadR de tipo silvestre o FadR[S219N] mutante en comparación con los títulos de FA de la cepa de *E. coli* de control (ALC487).

La FIG. 5 es un gráfico de barras de títulos totales de especies grasas (FA) en tres fermentaciones en matraz de agitación (SF, por sus siglas en inglés) separadas de la cepa D512 de *E. coli* que tiene una expresión alterada de FadR de tipo silvestre en comparación con los títulos de FA en la cepa ALC487 de control.

La FIG. 6 es un gráfico de barras de rendimiento total de especies grasas en carbono en fermentaciones en matraz de agitación de la cepa ALC487 de control o de la cepa D512 de *E. coli* que tiene una expresión alterada de FadR de tipo silvestre.

La FIG. 7 es un gráfico de producción de ácido graso y alcohol graso y de rendimiento total de especies grasas en fermentaciones de biorreactores de 5 l de la cepa ALC487 de control alimentada a una tasa de glucosa de 10 g/l/h o de la cepa D512 que tiene una expresión alterada de FadR de tipo silvestre alimentada a una tasa de glucosa de 10 g/l/h o 15 g/l/h. Las barras representan el título de alcohol graso o ácido graso, y los círculos representan el rendimiento total de especies grasas en carbono.

La FIG. 8 es un gráfico de producción de ácido graso y alcohol graso y rendimiento total de especies grasas en fermentaciones en matraz de agitación de la cepa D512 o una cepa D512 en la que se suprimió el gen entD. Las barras representan el título de ácido graso o alcohol graso y los círculos representan el rendimiento de ácido graso.

La FIG. 9 es un gráfico de títulos y rendimientos totales de especies grasas (ácidos grasos y FAME) en dos cepas de *E. coli* de la biblioteca de sitios de unión de ribosomas (RBS) que tienen una expresión alterada de FadR[S219N] mutante (es decir, P1A4 y P1G7) en comparación con los títulos y rendimientos totales de especies grasas en la cepa de *E. coli* parental (DAM1-pDS57) en fermentaciones en matraz de agitación (SF) a 32 °C. Las barras representan los títulos totales de especies grasas después de 56 horas de cultivo y los cuadrados representan el rendimiento total de especies grasas después de 56 horas de cultivo.

La FIG. 10 es un gráfico de líneas de títulos combinados de FAME y ácidos grasos libres (FFA) en la cepa parental DAM1 pDS57 o cepas de biblioteca de RBS P1A4 o P1G7 en fermentaciones en biorreactor en varios puntos temporales tras la inducción de la producción de FAME y FFA, en donde DAM1 P1A4 y DAM1 P1G7 expresan FadR y DAM1 pDS57 no expresa FadR.

La FIG. 11 es un gráfico de líneas de rendimientos combinados de FAME y FFA en la cepa parental DAM1 pDS57 o cepas de biblioteca de RBS P1A4 o P1G7 en fermentaciones en biorreactor en varios puntos temporales tras la inducción de la producción de FAME y FFA, en donde DAM1 P1A4 y DAM1 P1G7 expresan FadR y DAM1 pDS57 no expresa FadR.

Descripción detallada de la invención

La invención se basa, al menos en parte, en el descubrimiento de que la alteración del nivel de expresión de FadR en una célula hospedadora facilita una producción potenciada de ácidos grasos y derivados de ácidos grasos por la célula hospedadora.

La invención proporciona métodos mejorados para producir un ácido graso o un derivado de ácido graso en una célula

hospedadora bacteriana. El método consiste en (a) proporcionar una célula hospedadora bacteriana modificada por ingeniería genética que comprende un promotor heterólogo y/o un sitio de unión a ribosomas unido operativamente a una secuencia polinucleotídica que codifica un polipéptido FadR, provocando dicho promotor y/o sitio de unión a ribosomas la sobreexpresión del polipéptido FadR en dicha célula, (b) cultivar la célula hospedadora bacteriana modificada por ingeniería genética en un medio de cultivo en condiciones que permitan la producción de un ácido graso o un derivado del mismo y (c) aislar el ácido graso o un derivado del mismo de la célula hospedadora bacteriana modificada por ingeniería genética, en donde el ácido graso o el derivado del mismo es un ácido graso, una acil-ACP, una acil-CoA, un aldehído graso, un alcohol de cadena corta, un alcohol de cadena larga, un alcohol graso, un hidrocarburo o un éster. Como resultado de este método, uno o más de entre el título, el rendimiento o la productividad del ácido graso o derivado de ácido graso producido por la célula hospedadora bacteriana modificada por ingeniería genética aumentan con respecto al de la célula hospedadora de tipo silvestre correspondiente.

Definiciones

Como se utilizan en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas en singular "un", "una" y "el" o "la" incluyen los referentes en plural salvo que el contexto indique claramente otra cosa. Por tanto, por ejemplo, la referencia a "una célula hospedadora recombinante" incluye dos o más de dichas células hospedadoras recombinantes, la referencia a "un alcohol graso" incluye uno o más alcoholes grasos o mezclas de alcoholes grasos, la referencia a "una secuencia codificante de ácido nucleico" incluye una o más secuencias codificantes de ácido nucleico, la referencia a "una enzima" incluye una o más enzimas, y similares.

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el que entiende habitualmente un experto en la materia a la que pertenece la invención. Aunque en la práctica de la presente invención pueden usarse otros métodos y materiales similares, o equivalentes, a los que se describen en el presente documento, los materiales y métodos preferidos se describen en el presente documento.

En la descripción y la reivindicación de la presente invención, se usará la siguiente terminología de acuerdo con las definiciones que se exponen a continuación.

Números de registro: Los números de registro de secuencias a lo largo de la presente descripción se obtuvieron de las bases de datos proporcionadas por el NCBI (Centro Nacional de Información Biotecnológica) mantenido por los Institutos Nacionales de Salud, EE.UU. (que se identifican en el presente documento como "Números de registro del NCBI" o, como alternativa, como "Números de registro del GenBank"), y de la base de conocimiento UniProt (UniProtKB) y las bases de datos Swiss-Prot proporcionadas por el Instituto Suizo de Bioinformática (que se identifican en el presente documento como "Números de registro del UniProtKB").

Números de la clasificación de enzimas (CE): Los números CE son establecidos por el Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (IUBMB), cuya descripción se encuentra disponible en el sitio web de la Nomenclatura de enzimas IUBMB en la World Wide Web. Los números CE clasifican las enzimas de acuerdo con la reacción catalizada.

La expresión "polipéptido FadR" se refiere a un polipéptido que tiene una actividad biológica que corresponde a la de FadR derivado de *E. coli* MG1655 (SEQ ID NO: 1).

Como se usa en el presente documento, la expresión "ácido graso o derivado del mismo" significa un "ácido graso" o un "derivado de ácido graso". El término "ácido graso" significa un ácido carboxílico que tiene la fórmula RCOOH. R representa un grupo alifático, preferentemente, un grupo alquilo. R puede comprender entre aproximadamente 4 y aproximadamente 22 átomos de carbono. Los ácidos grasos pueden estar saturados, monoinsaturados o poliinsaturados. En una realización preferida, el ácido graso se prepara a partir de una vía biosintética de ácidos grasos. Un "derivado de ácido graso" es un producto preparado en parte a partir de la vía biosintética de ácidos grasos del organismo hospedador de producción. "Derivados de ácidos grasos" incluye productos preparados en parte a partir de acil-ACP o derivados de acil-ACP. Los derivados de ácidos grasos de ejemplo incluyen, por ejemplo, acil-CoA, ácidos grasos, aldehídos grasos, alcoholes de cadena corta y larga, hidrocarburos, alcoholes grasos, ésteres (por ejemplo, ceras, ésteres de ácidos grasos o ésteres grasos), olefinas terminales, olefinas internas y cetonas.

Una "composición de derivado de ácido graso", como se denomina en el presente documento, es producida por una célula hospedadora recombinante y normalmente comprende una mezcla de derivado de ácido graso. En algunos casos, la mezcla incluye más de un tipo de producto (por ejemplo, ácidos grasos y alcoholes grasos, ácidos grasos y ésteres de ácidos grasos o alcanos y olefinas). En otros casos, las composiciones de derivado de ácido graso pueden comprender, por ejemplo, una mezcla de alcoholes grasos (u otro derivado de ácido graso) con diversas longitudes de cadena y características de saturación o ramificación. En otros casos más, la composición de derivado de ácido graso comprende una mezcla tanto de más de un tipo de producto como de productos con diversas longitudes de cadena y características de saturación o ramificación.

Como se usa en el presente documento "acil-CoA" se refiere a un aciltioéster formado entre el átomo de carbono

carbonílico de la cadena de alquilo y el grupo sulfhidrido de la fracción 4'-fosfopantetonilo de la coenzima A (CoA), que tiene la fórmula $R-C(O)S-CoA$, donde R es cualquier grupo alquilo que tiene al menos 4 átomos de carbono.

5 Como se usa en el presente documento "acil-ACP" se refiere a un aciltoéster formado entre el carbono carbonílico de una cadena de alquilo y el grupo sulfhidrido de la fracción fosfopanteteinilo de una proteína portadora de acilo (ACP). La fracción fosfopanteteinilo se une después de la traducción a un resto de serina conservado en la ACP mediante la acción de la proteína sintasa portadora de holo-acilo (ACPS), una fosfopanteteinil transferasa. En algunas realizaciones, una acil-ACP es un intermedio en la síntesis de acil-ACP completamente saturadas. En otras realizaciones, una acil-ACP es un intermedio en la síntesis de acil-ACP insaturadas. En algunas realizaciones, la
10 cadena de carbono tendrá aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o 26 carbonos. Cada una de estas acil-ACP son sustratos para las enzimas que los convierten en derivados de ácidos grasos.

15 Como se usa en el presente documento, la expresión "vía biosintética de los ácidos grasos" significa una vía biosintética que produce ácidos grasos. La vía biosintética de los ácidos grasos incluye ácido graso sintetas que pueden modificarse por ingeniería genética para producir ácidos grasos y, en algunas realizaciones, pueden expresarse con enzimas adicionales para producir ácidos grasos que tengan las características deseadas de la cadena de carbono.

20 Como se usa en el presente documento, "aldehído graso" significa un aldehído que tiene la fórmula $RCHO$ caracterizada por un grupo carbonilo ($C=O$). En algunas realizaciones, el aldehído graso es cualquier aldehído preparado a partir de un ácido graso o derivado de ácido graso.

25 Como se usa en el presente documento, "alcohol graso" significa un alcohol que tiene la fórmula ROH . En algunas realizaciones, el alcohol graso es cualquier alcohol preparado a partir de un ácido graso o derivado de ácido graso.

En determinadas realizaciones, el grupo R de un ácido graso, aldehído graso o alcohol graso tiene al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18 o al menos 19, carbonos de longitud. Como alternativa, o además,
30 el grupo R tiene 20 o menos, 19 o menos, 18 o menos, 17 o menos, 16 o menos, 15 o menos, 14 o menos, 13 o menos, 12 o menos, 11 o menos, 10 o menos, 9 o menos, 8 o menos, 7 o menos, o 6 o menos átomos de carbono de longitud. Por tanto, el grupo R puede tener un grupo R limitado por dos cualquiera de los criterios de valoración anteriores. Por ejemplo, el grupo R puede tener 6-16 átomos de carbono de longitud, 10-14 átomos de carbono de longitud, o 12-18 átomos de carbono de longitud. En algunas realizaciones, el ácido graso, aldehído graso o alcohol
35 graso es un ácido graso, aldehído graso o alcohol graso C_6 , C_7 , C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} , C_{20} , C_{21} , C_{22} , C_{23} , C_{24} , C_{25} o C_{26} . En determinadas realizaciones, el ácido graso, aldehído graso o alcohol graso es un ácido graso, aldehído graso o alcohol graso C_6 , C_8 , C_{10} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} o C_{18} .

40 El grupo R de un ácido graso, aldehído graso o alcohol graso puede ser una cadena lineal o una cadena ramificada. Las cadenas ramificadas pueden tener más de un punto de ramificación y pueden incluir ramificaciones cíclicas. En algunas realizaciones, el ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado es un ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado C_6 , C_7 , C_8 , C_9 , C_{10} , C_{11} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} , C_{18} , C_{19} , C_{20} , C_{21} , C_{22} , C_{23} , C_{24} , C_{25} o C_{26} . En realizaciones particulares, el ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado es un ácido graso ramificado, aldehído graso ramificado o alcohol graso
45 ramificado C_6 , C_8 , C_{10} , C_{12} , C_{13} , C_{14} , C_{15} , C_{16} , C_{17} o C_{18} . En determinadas realizaciones, el grupo hidroxilo del ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado está en la posición primaria (C_1).

En determinadas realizaciones, el ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado es un iso-ácido graso, iso-aldehído graso o iso-alcohol graso, o un anteiso-ácido graso, un anteiso-aldehído graso o anteiso-alcohol graso. En realizaciones de ejemplo, el ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado se selecciona entre ácido graso ramificado, aldehído graso ramificado o alcohol graso ramificado iso- $C_{7:0}$, iso- $C_{8:0}$, iso- $C_{9:0}$, iso- $C_{10:0}$, iso- $C_{11:0}$, iso- $C_{12:0}$, iso- $C_{13:0}$, iso- $C_{14:0}$, iso- $C_{15:0}$, iso- $C_{16:0}$, iso- $C_{17:0}$, iso- $C_{18:0}$, iso- $C_{19:0}$, anteiso- $C_{7:0}$, anteiso- $C_{8:0}$, anteiso- $C_{9:0}$, anteiso- $C_{10:0}$, anteiso- $C_{11:0}$, anteiso- $C_{12:0}$, anteiso- $C_{13:0}$, anteiso- $C_{14:0}$, anteiso- $C_{15:0}$, anteiso- $C_{16:0}$, anteiso- $C_{17:0}$, anteiso- $C_{18:0}$ y anteiso- $C_{19:0}$.
55

El grupo R de un ácido graso ramificado o no ramificado, aldehído graso ramificado o no ramificado o alcohol graso ramificado o no ramificado puede estar saturado o no saturado. Si está insaturado, el grupo R puede tener uno o más de un punto de insaturación. En algunas realizaciones, el ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado es un ácido graso monoinsaturado, aldehído graso monoinsaturado o alcohol graso monoinsaturado.
60 En determinadas realizaciones, el ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado es un ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado $C_6:1$, $C_7:1$, $C_8:1$, $C_9:1$, $C_{10:1}$, $C_{11:1}$, $C_{12:1}$, $C_{13:1}$, $C_{14:1}$, $C_{15:1}$, $C_{16:1}$, $C_{17:1}$, $C_{18:1}$, $C_{19:1}$, $C_{20:1}$, $C_{21:1}$, $C_{22:1}$, $C_{23:1}$, $C_{24:1}$, $C_{25:1}$ o $C_{26:1}$. En determinadas realizaciones preferidas, el ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado es $C_{10:1}$, $C_{12:1}$, $C_{14:1}$, $C_{16:1}$ o $C_{18:1}$. En otras realizaciones más, el ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado está insaturado en la posición omega-7. En determinadas realizaciones, el ácido graso insaturado, aldehído graso insaturado o alcohol graso insaturado comprende un doble enlace cis.
65

Como se usa en el presente documento, el término "alcano" significa hidrocarburos saturados o compuestos que consisten solamente en carbono (C) e hidrógeno (H), en donde estos átomos están unidos entre sí por enlaces sencillos (es decir, son compuestos saturados).

5 Los términos "olefina" y "alqueno" se usan indistintamente en el presente documento y se refieren a hidrocarburos que contienen al menos un doble enlace carbono-carbono (es decir, son compuestos insaturados).

10 Las expresiones "olefina terminal", " α -olefina", "alqueno terminal" y "1-alqueno" se usan indistintamente en el presente documento con referencia a α -olefinas o alquenos con una fórmula química C_xH_{2x} , se distinguen de otras olefinas con una fórmula molecular similar por la linealidad de la cadena hidrocarbonada y la posición del doble enlace en la posición primaria o alfa.

15 Como se usa en el presente documento, la expresión "éster graso" puede usarse en referencia a un éster. En una realización preferida, un éster graso es cualquier éster preparado a partir de un ácido graso, por ejemplo, un éster de ácido graso. En algunas realizaciones, un éster graso contiene un lado A y un lado B. Como se usa en el presente documento, un "lado A" de un éster se refiere a la cadena de carbono unida al oxígeno de carboxilato del éster. Como se usa en el presente documento, un "lado B" de un éster se refiere a la cadena de carbono que comprende el carboxilato parental del éster. En realizaciones en las que el éster graso deriva de la vía biosintética de ácidos grasos, el lado A es aportado por un alcohol y el lado B es aportado por un ácido graso.

20 Puede usarse cualquier alcohol para formar el lado A de los ésteres grasos. Por ejemplo, el alcohol puede derivar de la vía biosintética de ácidos grasos. Como alternativa, el alcohol puede producirse a través de vías biosintéticas de ácidos no grasos. Además, el alcohol puede proporcionarse exógenamente. Por ejemplo, el alcohol puede suministrarse en el caldo de fermentación en los casos en los que el éster graso es producido por un organismo. Como alternativa, un ácido carboxílico, tal como un ácido graso o ácido acético, puede suministrarse exógenamente en los casos en los que el éster graso es producido por un organismo que también puede producir alcohol.

30 Las cadenas de carbono que comprenden el lado A o el lado B pueden tener cualquier longitud. En una realización, el lado A del éster tiene al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16 o 18 carbonos de longitud. Cuando el éster graso es un éster metílico de ácido graso, el lado A del éster tiene 1 carbono de longitud. Cuando el éster graso es un éster etílico de ácido graso, el lado A del éster tiene 2 carbonos de longitud. El lado B del éster puede tener al menos aproximadamente 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 o 26 carbonos de longitud. El lado A y/o el lado B pueden ser de cadena lineal o ramificada. Las cadenas ramificadas pueden tener uno o más puntos de ramificación. Además, las cadenas ramificadas pueden incluir ramas cíclicas. Además, el lado A y/o el lado B pueden estar saturados o insaturados. Si está insaturado, el lado A y/o el lado B pueden tener uno o más puntos de insaturación.

40 En algunas realizaciones, el éster de ácido graso es un éster metílico de ácido graso (FAME) o un éster etílico de ácido graso (FAEE). En determinadas realizaciones, el FAME es un beta-hidroxi (B-OH) FAME. En una realización, el éster graso se produce biosintéticamente. En esta realización, primero el ácido graso se "activa". Son ejemplos no limitantes de ácidos grasos "activados" acil-CoA, acil ACP y acil fosfato. Acil-CoA puede ser un producto directo de la biosíntesis o degradación de ácidos grasos. Además, puede sintetizarse acil-CoA a partir de un ácido graso libre, una CoA y un trifosfato de nucleótido de adenosina (ATP). Un ejemplo de una enzima que produce acil-CoA es la acil-CoA sintasa.

45 Después de que un ácido graso se activa, puede transferirse fácilmente a un nucleófilo receptor. Son nucleófilos de ejemplo alcoholes, tioles o fosfatos.

50 En una realización, el éster graso es una cera. La cera puede derivar de un alcohol de cadena larga y un ácido graso de cadena larga. En otra realización, el éster graso es un tioéster de ácido graso, por ejemplo, acil Coenzima A (CoA) graso. En otras realizaciones, el éster graso es un acil pantotenato graso, una proteína portadora de acilo (ACP) o un éster de fosfato graso.

55 Como se usa en el presente documento "acil CoA" se refiere a un aciltioéster formado entre el átomo de carbono carbonílico de la cadena de alquilo y el grupo sulfhidrido de la fracción 4'-fosfopantetionilo de la coenzima A (CoA), que tiene la fórmula $R-C(O)S-CoA$, donde R es cualquier grupo alquilo que tiene al menos 4 átomos de carbono. En algunos casos, una acil CoA será un intermedio en la síntesis de acil CoA totalmente saturadas, incluyendo, pero sin limitación, 3-ceto-acil CoA, una 3-hidroxi acil CoA, una delta-2-trans-enoil-CoA o una alquil acil CoA. En algunas realizaciones, la cadena de carbono tendrá aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o 26 átomos de carbono. En otras realizaciones la acil CoA estará ramificada. En una realización, la acil CoA ramificada es una isoacil CoA, en otra, es una anti-isoacil CoA. Cada una de estas "acil CoA" son sustratos para enzimas que las convierten en derivados de ácidos grasos tales como los que se describen en el presente documento.

65 La expresión "nivel de expresión alterado" y "nivel de expresión modificado" se usan indistintamente, y significan que un polinucleótido, polipéptido o hidrocarburo está presente a una concentración diferente en una célula hospedadora modificada en comparación con su concentración en una célula de tipo silvestre correspondiente en las mismas

condiciones.

"Polinucleótido" se refiere a un polímero de ADN o ARN, que puede ser monocatenario o bicatenario y que puede contener nucleótidos no naturales o modificados. Las expresiones "polinucleótido", "ácido nucleico", y "molécula de ácido nucleico" se usan en el presente documento indistintamente para referirse a una forma polimérica de nucleótidos de cualquier longitud, ya sea ribonucleótidos (ARN) o desoxirribonucleótidos (ADN). Estos términos se refieren a la estructura primaria de la molécula y, por lo tanto, incluyen ADN bicatenario y monocatenario, y ARN bicatenario y monocatenario. Los términos incluyen, como equivalentes, análogos de ARN o ADN hechos de análogos de nucleótidos y polinucleótidos modificados tales como, pero sin limitación, polinucleótidos metilados y/o protegidos terminalmente. El polinucleótido puede estar en cualquier forma, incluyendo, pero sin limitación, plásmido, vírico, cromosómico, EST, ADNc, ARNm y ARNr.

El término "nucleótido", como se usa en el presente documento, se refiere a una unidad monomérica de un polinucleótido que consiste en una base heterocíclica, un azúcar y uno o más grupos fosfato. Las bases naturales (guanina, (G), adenina, (A), citosina, (C), timina, (T) y uracilo (U)) normalmente derivan de purina o pirimidina, aunque debe entenderse que también se incluyen análogos de bases naturales y no naturales. El azúcar natural es la pentosa (azúcar de cinco átomos de carbono) desoxirribosa (que forma el ADN) o ribosa (que forma el ARN), aunque debe entenderse que también se incluyen análogos de azúcares naturales y no naturales. Los ácidos nucleicos normalmente están unidos mediante enlaces fosfato para formar ácidos nucleicos o polinucleótidos, aunque se conocen muchos otros enlaces en la técnica (por ejemplo, fosforotioatos, boranofosfatos y similares).

Los polinucleótidos que se describen en el presente documento pueden comprender nucleótidos degenerados que se definen de acuerdo con el código IUPAC para la degeneración de nucleótidos en donde B es C, G o T; D es A, G o T; H es A, C o T; K es G o T; M es A o C; N es A, C, G o T; R es A o G; S es C o G; V es A, C o G; W es A o T; y Y es C o T.

Los términos "polipéptido" y "proteína" se refieren a un polímero de restos de aminoácidos. La expresión "polipéptido recombinante" se refiere a un polipéptido que se produce mediante técnicas de ADN recombinante, en donde generalmente el ADN que codifica la proteína expresada o el ARN se inserta en un vector de expresión adecuado que a su vez se usa para transformar una célula hospedadora para producir el polipéptido o ARN.

En algunas realizaciones, el polipéptido, polinucleótido o hidrocarburo que tiene un nivel de expresión alterado o modificado está "sobreexpresado" o tiene un "nivel de expresión aumentado". Como se usa en el presente documento, "sobreexpresar" y "aumentar el nivel de expresión" significan expresar o hacer que se exprese un polinucleótido, polipéptido o hidrocarburo en una célula a una concentración superior a la que se expresa normalmente en una célula de tipo silvestre correspondiente en las mismas condiciones. Por ejemplo, un polipéptido puede "sobreexpresarse" en una célula hospedadora modificada por ingeniería genética cuando el polipéptido está presente en una concentración mayor en la célula hospedadora modificada por ingeniería genética en comparación con su concentración en una célula hospedadora no modificada por ingeniería genética de la misma especie en las mismas condiciones.

En otras realizaciones, el polipéptido, polinucleótido o hidrocarburo que tiene un nivel de expresión alterado está "atenuado" o tiene un "nivel de expresión disminuido". Como se usa en el presente documento, "atenuar" y "disminuir el nivel de expresión" significan expresar o hacer que se exprese un polinucleótido, polipéptido o hidrocarburo en una célula a una concentración inferior a la que se expresa normalmente en una célula de tipo silvestre correspondiente en las mismas condiciones.

El grado de sobreexpresión o atenuación puede ser de 1,5 veces o más, por ejemplo, 2 veces o más, 3 veces o más, 5 veces o más, 10 veces o más, o 15 veces o más. Como alternativa, o además, el grado de sobreexpresión o atenuación puede ser de 500 veces o menos, por ejemplo, 100 veces o menos, 50 veces o menos, 25 veces o menos, o 20 veces o menos. Por tanto, el grado de sobreexpresión o atenuación puede estar limitada por dos cualquiera de los criterios de valoración anteriores. Por ejemplo, el grado de sobreexpresión o atenuación puede ser de 1,5-500 veces, 2-50 veces, 10-25 veces o 15-20 veces.

En algunas realizaciones, un polipéptido que se describe en el presente documento tiene "un nivel de actividad aumentado". Por "nivel de actividad aumentado" se entiende que un polipéptido tiene un nivel mayor de función bioquímica o biológica (por ejemplo, unión a ADN o actividad enzimática) en una célula hospedadora modificada por ingeniería genética en comparación con su nivel de función bioquímica y/o biológica en una célula hospedadora de tipo silvestre correspondiente en las mismas condiciones. El grado de actividad potenciada puede ser de aproximadamente el 10 % o más, aproximadamente el 20 % o más, aproximadamente el 50 % o más, aproximadamente el 75 % o más, aproximadamente el 100 % o más, aproximadamente el 200 % o más, aproximadamente el 500 % o más, aproximadamente el 1000 % o más, o cualquier intervalo incluido en los mismos.

Un polinucleótido o polipéptido puede atenuarse usando métodos conocidos en la técnica. En algunas realizaciones, la expresión de un gen o polipéptido codificado por el gen se atenúa mediante la mutación de las secuencias polinucleótidas reguladoras que controlan la expresión del gen. En otras realizaciones, la expresión de un gen o polipéptido codificado por el gen se atenúa mediante la sobreexpresión de una proteína represora o proporcionando

un elemento regulador exógeno que active una proteína represora. En aún otras realizaciones, se usan métodos de silenciamiento de genes basados en ADN o ARN para atenuar la expresión de un gen o polinucleótido. En algunas realizaciones, la expresión de un gen o polipéptido se atenúa totalmente, por ejemplo, suprimiendo toda o una porción de la secuencia polinucleotídica de un gen.

5 Un polinucleótido o polipéptido puede sobreexpresarse usando métodos conocidos en la técnica. En algunos aspectos, la sobreexpresión de un polipéptido se consigue mediante el uso de un elemento regulador exógeno. La expresión "elemento regulador exógeno" se refiere generalmente a un elemento regulador que se origina fuera de la célula hospedadora. Sin embargo, en determinados aspectos, la expresión "elemento regulador exógeno" puede referirse a
10 un elemento regulador derivado de la célula hospedadora cuya función se replica o se usurpa con el fin de controlar la expresión de un polipéptido endógeno. Por ejemplo, si la célula hospedadora es una célula de *E. coli* y el polipéptido FadR es codificado por un gen endógeno *fadR*, entonces la expresión del *fadR* endógeno puede controlarse mediante un promotor derivado de otro gen de *E. coli*.

15 En algunos aspectos, el elemento regulador exógeno es un compuesto químico, tal como una molécula pequeña. Como se usa en el presente documento, la expresión "molécula pequeña" se refiere a una sustancia o compuesto que tiene un peso molecular inferior a aproximadamente 1.000 g/mol.

20 En algunos aspectos, el elemento regulador exógeno que controla la expresión de un gen *fadR* endógeno es una secuencia de control de la expresión que se une operativamente al gen *fadR* endógeno mediante integración recombinante en el genoma de la célula hospedadora. En determinados aspectos, la secuencia de control de la expresión se integra en un cromosoma de la célula hospedadora mediante recombinación homóloga, usando métodos conocidos en la técnica (por ejemplo, Datsenko *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 97(12): 6640-6645 (2000)).

25 Las secuencias de control de la expresión son conocidas en la técnica e incluyen, por ejemplo, promotores, potenciadores, señales de poliadenilación, terminadores de la transcripción, sitios de entrada de ribosomas internos (IRES), sitios de unión de ribosomas (RBS) y similares, que proporcionan la expresión de la secuencia polinucleotídica en una célula hospedadora. Las secuencias de control de la expresión interactúan específicamente con proteínas celulares que participan en la transcripción (Maniatis *et al.*, *Science*, 236: 1237-1245 (1987)). Se describen ejemplos
30 de secuencias de control de la expresión en, por ejemplo, Goeddel, *Gene Expression Technology: Methods in Enzymology*, Vol. 185, Academic Press, San Diego, Calif. (1990). En los métodos de la invención la secuencia de control de la expresión es un promotor heterólogo y/o un sitio de unión de ribosomas.

35 En los métodos de la invención, una secuencia de control de la expresión está unida operativamente a una secuencia polinucleotídica. Por "unidas operativamente" se entiende que una secuencia polinucleotídica y una secuencia o secuencias de control de la expresión están conectadas de manera que permiten la expresión génica cuando las moléculas apropiadas (por ejemplo, proteínas activadoras de la transcripción) se unen a la secuencia o secuencias de control de la expresión. Los promotores unidos operativamente se encuentran en dirección 5' de la secuencia polinucleotídica seleccionada en términos de la dirección de transcripción y traducción. Los potenciadores unidos
40 operativamente se pueden ubicar en dirección 5', dentro o en dirección 3' del polinucleótido seleccionado.

En algunas realizaciones, la secuencia polinucleotídica se proporciona a la célula hospedadora por medio de un vector recombinante, que comprende un promotor unido operativamente a la secuencia polinucleotídica. En determinadas realizaciones, el promotor es un promotor regulado por el desarrollo, específico de orgánulo, específico de tejido,
45 inducible, constitutivo o específico de la célula.

Como se usa en el presente documento, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico, es decir, una secuencia polinucleotídica, al que se ha unido. Un tipo de vector útil es un episoma (es decir, un ácido nucleico capaz de realizar la replicación extracromosómica). Los vectores útiles son
50 aquellos capaces de realizar la replicación autónoma y/o expresión de ácidos nucleicos a los que están unidos. Los vectores capaces de dirigir la expresión de los genes a los que están unidos operativamente se denominan, en el presente documento, "vectores de expresión". En general, los vectores de expresión útiles en las técnicas de ADN recombinante se encuentran normalmente en forma de "plásmidos" que, en general, se refieren a bucles de ADN bicatenarios circulares que, en su forma vectorial, no están unidos al cromosoma. Los términos "plásmido" y "vector" se usan indistintamente en el presente documento, ya que el plásmido es la forma más comúnmente utilizada de vector. Sin embargo, también se incluyen otras formas de vectores de expresión que tienen funciones equivalentes y que se dan a conocer en la técnica a continuación.

60 La expresión "secuencias reguladoras", como se usa en el presente documento, normalmente se refiere a una secuencia de bases en el ADN, unida operativamente a secuencias de ADN que codifican una proteína que finalmente controla la expresión de la proteína. Los ejemplos de secuencias reguladoras incluyen, pero sin limitación, secuencias promotoras de ARN, secuencias de unión de factor de transcripción, secuencias de terminación de la transcripción, moduladores de la transcripción (tales como elementos potenciadores), secuencias de nucleótidos que afectan a la estabilidad del ARN y secuencias reguladoras de la traducción (tales como sitios de unión a ribosomas (por ejemplo, secuencias de Shine-Dalgarno en procariotas o secuencias de Kozak en eucariotas), codones de inicio, codones de
65 terminación).

Como se usa en el presente documento, la expresión "la expresión de dicha secuencia de nucleótidos se modifica con respecto a la secuencia de nucleótidos de tipo silvestre", significa un aumento o una disminución en el nivel de expresión y/o la actividad de una secuencia de nucleótidos endógena o la expresión y/o actividad de una secuencia de nucleótidos que codifica un polipéptido heterólogo o no nativo.

Como se usa en el presente documento, el término "expresar" con respecto a un polinucleótido es hacer que funcione. Un polinucleótido que codifica un polipéptido (o una proteína), cuando se expresa, se transcribirá o se traducirá para producir ese polipéptido (o esa proteína). Como se usa en el presente documento, el término "sobrexexpresar" significa expresar o hacer que se exprese un polinucleótido o polipéptido en una célula a una concentración superior a la expresada normalmente en una célula de tipo silvestre correspondiente en las mismas condiciones.

En algunos aspectos, el vector recombinante comprende al menos una secuencia seleccionada entre el grupo que consiste en (a) una secuencia de control de la expresión acoplada operativamente a la secuencia polinucleotídica; (b) un marcador de selección acoplado operativamente a la secuencia polinucleotídica; (c) una secuencia marcadora acoplada operativamente a la secuencia polinucleotídica; (d) una fracción de purificación acoplada operativamente a la secuencia polinucleotídica; (e) una secuencia de secreción acoplada operativamente a la secuencia polinucleotídica; y (f) una secuencia de direccionamiento acoplada operativamente a la secuencia polinucleotídica.

Los vectores de expresión descritos en el presente documento incluyen una secuencia polinucleotídica que se describe en el presente documento en una forma adecuada para la expresión de la secuencia polinucleotídica en una célula hospedadora. Los expertos en la materia apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula hospedadora que se vaya a transformar, el nivel de expresión del polipéptido deseado, etc. Los vectores de expresión descritos en el presente documento pueden introducirse en células hospedadoras para producir polipéptidos, incluyendo polipéptidos de fusión, codificados por las secuencias polinucleotídicas como se describe en el presente documento.

La expresión de genes que codifican polipéptidos en procariotas, por ejemplo, *E. coli*, se realiza con mayor frecuencia con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de polipéptidos de fusión o no fusión. Los vectores de fusión agregan una cantidad de aminoácidos a un polipéptido codificado en los mismos, por lo general al extremo amino o carboxilo del polipéptido recombinante. Dichos vectores de fusión normalmente sirven para uno o más de los tres fines siguientes: (1) aumentar la expresión del polipéptido recombinante; (2) aumentar la solubilidad del polipéptido recombinante; y (3) ayudar en la purificación del polipéptido recombinante actuando como ligando en la purificación por afinidad. A menudo, en los vectores de expresión de fusión, se introduce un sitio de escisión proteolítica en la unión de la fracción de fusión y el polipéptido recombinante. Esto permite la separación del polipéptido recombinante de la fracción de fusión después de la purificación del polipéptido de fusión. Los ejemplos de dichas enzimas, y sus secuencias de reconocimiento afines, incluyen el factor Xa, la trombina y la enterocinasa. Los vectores de expresión de fusión de ejemplo incluyen pGEX (Pharmacia Biotech, Inc., Piscataway, NJ; Smith *et al.*, *Gene*, 67: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) y pRITS (Pharmacia Biotech, Inc., Piscataway, N.J.), que fusionan glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, a la proteína recombinante diana.

Se conocen bien en la técnica sistemas de expresión adecuados para células procariotas y eucariotas; véase, por ejemplo, Sambrook *et al.*, "*Molecular Cloning: A Laboratory Manual*", segunda edición, Cold Spring Harbor Laboratory (1989). Los ejemplos de vectores de expresión de *E. coli* no de fusión inducibles incluyen pTrc (Amann *et al.*, *Gene*, 69: 301-315 (1988)) y PET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA, págs. 60-89 (1990)). En determinadas realizaciones, una secuencia polinucleotídica de la invención está unida operativamente a un promotor derivado del bacteriófago T5. Los ejemplos de vectores para la expresión en levadura incluyen pYepSec1 (Baldari *et al.*, *EMBO J.*, 6: 229-234 (1987)), pMFa (Kurjan *et al.*, *Cell*, 30: 933-943 (1982)), pJRY88 (Schultz *et al.*, *Gene*, 54: 113-123 (1987)), pYES2 (Invitrogen Corp., San Diego, CA) y picZ (Invitrogen Corp., San Diego, CA). Los vectores de baculovirus disponibles para la expresión de proteínas en células de insecto cultivadas (por ejemplo, células Sf9) incluyen, por ejemplo, la serie pAc (Smith *et al.*, *Mol. Cell Biol.*, 3: 2156-2165 (1983)) y la serie de pVL (Lucklow *et al.*, *Virology*, 170: 31-39 (1989)). Los ejemplos de vectores de expresión de mamífero incluyen pCDM8 (Seed, *Nature*, 329: 840 (1987)) y pMT2PC (Kaufman *et al.*, *EMBO J.*, 6: 187-195 (1987)).

Los vectores pueden introducirse en células procariotas o eucariotas mediante técnicas convencionales de transformación o transfección. Como se usa en el presente documento, los términos "transformación" y "transfección" se refieren a diversas técnicas reconocidas por la técnica para introducir ácido nucleico extraño (por ejemplo, ADN) en una célula hospedadora, incluyendo la coprecipitación con fosfato de calcio o cloruro de calcio, la transfección mediada por DEAE-dextrano, la lipofección o la electroporación. Se pueden encontrar métodos adecuados para transformar o transfectar células hospedadoras en, por ejemplo, Sambrook *et al.* (citado anteriormente).

Para la transformación estable de células bacterianas, se sabe que, según el vector de expresión y la técnica de transformación utilizada, solo una fracción pequeña de células captará y replicará el vector de expresión. Para identificar y seleccionar estos transformantes, se puede introducir un gen que codifique un marcador seleccionable (por ejemplo, resistencia a un antibiótico) en las células hospedadoras junto con el gen de interés. Los marcadores

seleccionables incluyen aquellos que confieren resistencia a fármacos tales como, pero sin limitación, ampicilina, kanamicina, cloranfenicol o tetraciclina. Los ácidos nucleicos que codifican un marcador seleccionable pueden introducirse en una célula hospedadora en el mismo vector que el que codifica un polipéptido descrito en el presente documento o pueden introducirse en un vector separado. Las células transformadas de manera estable con el ácido nucleico introducido pueden identificarse mediante crecimiento en presencia de un fármaco de selección apropiado.

De la misma manera, para la transfección estable en células de mamífero, se sabe que, según el vector de expresión y la técnica de transfección utilizada, solo una pequeña fracción de las células puede integrar el ADN extraño en su genoma. Con el fin de identificar y seleccionar estos integrantes, se puede introducir un gen que codifique un marcador seleccionable (por ejemplo, resistencia a un antibiótico) en las células hospedadoras junto con el gen de interés. Los marcadores seleccionables preferidos incluyen los que confieren resistencia a fármacos, tal como G418, higromicina y metotrexato. Los ácidos nucleicos que codifican un marcador seleccionable pueden introducirse en una célula hospedadora en el mismo vector que el que codifica un polipéptido descrito en el presente documento o pueden introducirse en un vector separado. Las células transfectadas de manera estable con el ácido nucleico introducido pueden identificarse mediante crecimiento en presencia de un fármaco de selección apropiado.

En algunas realizaciones, el polipéptido FadR tiene la secuencia de aminoácidos de la SEQ ID NO: 1.

En otras realizaciones, el polipéptido FadR está codificado por un gen *fadR* obtenido de microorganismos del género *Escherichia*, *Salmonella*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Cronobacter*, *Yersinia*, *Serratia*, *Erwinia*, *Pectobacterium*, *Photobacterium*, *Edwardsiella*, *Shewanella* o *Vibrio*.

En otras realizaciones, el polipéptido FadR es un homólogo de FadR que tiene una secuencia de aminoácidos al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 % o al menos un 99 % idéntica a la secuencia de aminoácidos de la SEQ ID NO: 1.

La identidad de un polipéptido FadR que tiene al menos un 80 % de identidad con la secuencia de aminoácidos de la SEQ ID NO: 1 no está particularmente limitada y un experto habitual en la técnica puede identificar fácilmente los homólogos de FadR derivado de *E. coli* MG1655 usando los métodos que se describen en el presente documento así como métodos conocidos en la técnica.

Como se usan en el presente documento, los términos "homólogo", y "homólogos" se refieren a un polinucleótido o un polipéptido que comprende una secuencia que es al menos un 50 % idéntica a la secuencia polinucleotídica o polipeptídica correspondiente. Preferentemente, los polinucleótidos o polipéptidos homólogos tienen secuencias de polinucleótidos o secuencias de aminoácidos que tienen al menos aproximadamente un 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o al menos aproximadamente un 99 % de homología con la secuencia de aminoácidos o la secuencia polinucleotídica correspondientes. Como se usan en el presente documento, los términos "homología" de secuencia e "identidad" de secuencia se usan indistintamente.

En resumen, los cálculos de "homología" entre dos secuencias se pueden realizar de la siguiente manera. Se alinean las secuencias con fines comparativos óptimos (por ejemplo, se pueden introducir huecos en una o ambas de una primera y una segunda secuencia de aminoácidos o de ácido nucleico para un alineamiento óptimo, y las secuencias no homólogas se pueden descartar para fines comparativos). Luego se comparan los restos de aminoácidos o nucleótidos de las posiciones de aminoácidos o posiciones de nucleótidos correspondientes de la primera y segunda secuencias. Cuando una posición en la primera secuencia está ocupada por el mismo resto de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición (como se usa en el presente documento, "identidad" de aminoácido o ácido nucleico es equivalente a "homología" de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos y la longitud de cada hueco, que es necesario introducir para una alineación óptima de las dos secuencias.

Puede realizarse la comparación de secuencias y la determinación del porcentaje de homología entre dos secuencias usando un algoritmo matemático, tal como BLAST (Altschul *et al.* *J. Mol. Biol.*, 215(3): 403-410 (1990)). El porcentaje de homología entre dos secuencias de aminoácidos también se puede determinar usando el algoritmo de Needleman y Wunsch, que se ha incorporado al programa GAP en el paquete de software GCG, usando una matriz Blossum 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 o 4 y un peso de longitud de 1, 2, 3, 4, 5 o 6 (Needleman y Wunsch, *J. Mol. Biol.*, 48: 444-453 (1970)). El porcentaje de homología entre dos secuencias de nucleótidos también se puede determinar usando el programa GAP en el paquete de software GCG, usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 o 6. Un experto en la materia puede realizar cálculos iniciales de homología y ajustar los parámetros del algoritmo en consecuencia. Un conjunto preferido de parámetros (y el que debería usarse si un profesional no estuviera seguro acerca de qué parámetros deben aplicarse para determinar si una molécula está dentro de una limitación de homología de las reivindicaciones) es una matriz de puntuación Blossum 62 con una penalización por hueco de 12, una penalización por hueco extendido de 4 y una penalización por hueco de desplazamiento de fase de 5. Se conocen

métodos adicionales de alineamiento de secuencias en las técnicas de biotecnología (véase, por ejemplo, Rosenberg, *BMC Bioinformatics*, 6: 278 (2005); Altschul *et al.*, *FEBS J.*, 272(20): 5101-5109 (2005)).

Como se usa en el presente documento, la expresión "hibrida en condiciones de rigurosidad bajas, rigurosidad media, rigurosidad alta o rigurosidad muy alta" describe las condiciones para la hibridación y el lavado. Puede encontrarse orientación para llevar a cabo las reacciones de hibridación en *"Current Protocols in Molecular Biology"*, John Wiley & Sons, N. Y. (1989), 6.3.1 - 6.3.6. En esa referencia, se describen métodos acuosos y no acuosos, y puede usarse cualquier método. Las condiciones de hibridación específicas citadas en el presente documento son las siguientes: 1) condiciones de hibridación de baja rigurosidad en cloruro de sodio/citrato de sodio 6X (SSC) a aproximadamente 45 °C, seguido de dos lavados en SSC 0,2X, SDS al 0,1 % al menos a 50 °C (la temperatura de los lavados puede aumentarse a 55 °C para las condiciones de baja rigurosidad); 2) condiciones de hibridación de rigurosidad media en SSC 6X a aproximadamente 45 °C, seguida de uno o más lavados en SSC 0,2X, SDS al 0,1 % a 60 °C; 3) condiciones de hibridación de rigurosidad elevada en SSC 6X a aproximadamente 45 °C, seguido de uno o más lavados en 0,2X SSC, SDS al 0,1 % a 65 °C; y preferentemente 4) condiciones de hibridación de muy alta rigurosidad son fosfato de sodio 0,5M, SDS al 7 % a 65 °C, seguida de uno o más lavados en SSC 0,2X, SDS al 1 % a 65 °C. Las condiciones de rigurosidad muy alta (4) son las condiciones preferidas a menos que se especifique lo contrario.

En algunas realizaciones, el polipéptido es un fragmento de cualquiera de los polipéptidos que se describen en el presente documento. El término "fragmento" se refiere a una parte más corta de un polipéptido o de una proteína de longitud completa que varía en tamaño desde cuatro restos de aminoácidos hasta la secuencia de aminoácidos completa menos un resto de aminoácido. En determinadas realizaciones de la invención, un fragmento se refiere a la secuencia de aminoácidos completa de un dominio de un polipéptido o de una proteína (por ejemplo, un dominio de unión al sustrato o un dominio catalítico).

Un polipéptido "endógeno" se refiere a un polipéptido codificado por el genoma de la célula microbiana parental (también denominada "célula hospedadora") a partir de la que la célula recombinante se modifica por ingeniería genética (o se "deriva").

Un polipéptido "exógeno" se refiere a un polipéptido que no está codificado por el genoma de la célula microbiana parental. Una variante de polipéptido (es decir, mutante) es un ejemplo de un polipéptido exógeno.

El término "heterólogo", como se usa en el presente documento, se refiere normalmente a una secuencia de nucleótidos o a una proteína que no está presente de forma natural en un organismo. Por ejemplo, una secuencia polinucleotídica endógena a una planta puede introducirse en una célula hospedadora mediante métodos recombinantes y el polinucleótido vegetal es entonces un polinucleótido heterólogo en una célula hospedadora recombinante.

En algunas realizaciones, el polipéptido es un mutante o una variante de cualquiera de los polipéptidos descritos en el presente documento. Los términos "mutante" y "variante", como se usan en el presente documento, se refieren a un polipéptido que tiene una secuencia de aminoácidos que difiere de un polipéptido de tipo silvestre en al menos un aminoácido. Por ejemplo, el mutante puede comprender una o más de las siguientes sustituciones de aminoácidos conservadoras: el reemplazo de un aminoácido alifático, tal como alanina, valina, leucina e isoleucina, con otro aminoácido alifático; el reemplazo de una serina con una treonina; el reemplazo de una treonina con una serina; el reemplazo de un resto ácido, tal como el ácido aspártico y el ácido glutámico, con otro resto ácido; el reemplazo de un resto portador de un grupo amida, tal como la asparagina y la glutamina, con otro resto portador de un grupo amida; el intercambio de un resto básico, tal como lisina y arginina, con otro resto básico; y el reemplazo de un resto aromático, tal como fenilalanina y tirosina, con otro resto aromático. En algunas realizaciones, el polipéptido mutante tiene aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 o más sustituciones, adiciones, inserciones o supresiones de aminoácidos.

Como se usa en el presente documento, el término "mutagénesis" se refiere a un proceso mediante el que la información genética de un organismo se cambia de manera estable. La mutagénesis de una secuencia de ácido nucleico que codifica una proteína produce una proteína mutante. La mutagénesis también se refiere a cambios en las secuencias de ácido nucleico no codificantes que producen una actividad proteica modificada.

Como se usa en el presente documento, el término "gen" se refiere a secuencias de ácido nucleico que codifican un producto de ARN o un producto proteico, así como secuencias de ácido nucleico unidas operativamente que afectan a la expresión del ARN o proteína (por ejemplo, dichas secuencias incluyen, pero sin limitación, secuencias promotoras o potenciadoras) o secuencias de ácido nucleico unidas operativamente que codifican secuencias que afectan a la expresión del ARN o proteína (por ejemplo, dichas secuencias incluyen, pero sin limitación, sitios de unión a ribosomas o secuencias de control de la traducción).

En determinadas realizaciones, el polipéptido FadR comprende una mutación en un resto de aminoácido que corresponde al aminoácido 219 de la SEQ ID NO: 1. En determinadas realizaciones, la mutación da como resultado una sustitución del resto de aminoácido correspondiente al aminoácido 219 de la SEQ ID NO: 1 con un resto de asparagina. La mutación FadR(S219N) se ha descrito anteriormente (Raman *et al.*, *J. Biol. Chem.*, 270: 1092-1097

(1995)).

Los fragmentos preferidos o mutantes de un polipéptido conservan una parte o la totalidad de la función biológica (por ejemplo, la actividad enzimática) del polipéptido de tipo silvestre correspondiente. En algunas realizaciones, el fragmento o mutante conserva al menos un 75 %, al menos un 80 %, al menos un 90 %, al menos un 95 % o al menos un 98 % o más de la función biológica del polipéptido de tipo silvestre correspondiente. En otras realizaciones, el fragmento o mutante conserva aproximadamente el 100 % de la función biológica del polipéptido de tipo silvestre correspondiente. Puede encontrarse orientación para determinar qué restos de aminoácidos pueden sustituirse, insertarse o suprimirse sin afectar a la actividad biológica usando programas informáticos bien conocidos en la técnica, por ejemplo, software LASERGENE™ (DNASTAR, Inc., Madison, WI).

En otras realizaciones más, un fragmento o mutante presenta una función biológica aumentada en comparación con el polipéptido de tipo silvestre correspondiente. Por ejemplo, un fragmento o mutante puede mostrar al menos un 10 %, al menos un 25 %, al menos un 50 %, al menos un 75 % o al menos un 90 % de mejora en la actividad enzimática en comparación con el polipéptido de tipo silvestre correspondiente. En otras realizaciones, el fragmento o mutante presenta una mejora de al menos el 100 % (por ejemplo, al menos el 200 % o al menos el 500 %) de la actividad enzimática en comparación con el polipéptido de tipo silvestre correspondiente.

Se entiende que los polipéptidos descritos en el presente documento pueden tener sustituciones de aminoácidos conservativas o no esenciales adicionales, que no tengan un efecto sustancial sobre la función del polipéptido. Si se tolerará o no una determinada sustitución (es decir, si no afectará negativamente a la función biológica deseada), como la unión del ADN o la actividad enzimática) puede determinarse como se describe en Bowie *et al.* (*Science*, 247: 1306-1310 (1990)).

Una "sustitución de aminoácidos conservativa" es una en la que se reemplaza el resto de aminoácido por un resto de aminoácido que tenga una cadena lateral similar. Se han definido en la técnica familias de restos de aminoácidos que tienen cadenas laterales similares. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tiroides, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales beta-ramificadas (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina).

Las variantes pueden ser naturales o creadas *in vitro*. En particular, dichas variantes se pueden crear usando técnicas de ingeniería genética, tales como mutagénesis dirigida, mutagénesis química aleatoria, procedimientos de eliminación de exonucleasa III o técnicas de clonación convencionales. Como alternativa, dichas variantes, fragmentos, análogos o derivados se pueden crear mediante síntesis química o procedimientos de modificación.

Los métodos para crear variantes son bien conocidos en la técnica. Estos incluyen procedimientos en los que se modifican las secuencias de ácido nucleico obtenidas a partir de aislados naturales para generar ácidos nucleicos que codifican polipéptidos que tienen características que potencian su valor en aplicaciones industriales o de laboratorio. En dichos procedimientos, se genera y se caracteriza un gran número de secuencias variantes que tienen una o más diferencias de nucleótidos con respecto a la secuencia obtenida del aislado natural. Normalmente, estas diferencias de nucleótidos dan como resultado cambios en los aminoácidos con respecto a los polipéptidos codificados por los ácidos nucleicos de los aislados naturales.

Por ejemplo, las variantes pueden prepararse usando mutagénesis aleatoria y dirigida (véase, por ejemplo, Arnold, *Curr. Opin. Biotech.*, 4: 450-455 (1993)). La mutagénesis aleatoria se puede lograr usando PCR propensa a errores (véase, por ejemplo, Leung *et al.*, *Technique*, 1: 11-15 (1989); y Caldwell *et al.*, *PCR Methods Applic.*, 2: 28-33 (1992)). La mutagénesis dirigida se puede realizar usando mutagénesis dirigida a oligonucleótidos para generar mutaciones específicas del sitio en cualquier ADN clonado de interés (véase, por ejemplo, Reidhaar-Olson *et al.*, *Science*, 241: 53-57 (1988)). Otros métodos para generar variantes incluyen, por ejemplo, ensamblaje PCR (véase, por ejemplo, la Patente de los EE.UU. 5.965.408), mutagénesis sexual por PCR (véase, por ejemplo, Stemmer, *Proc. Natl. Acad. Sci., U.S.A.*, 91: 10747-10751 (1994) y la Patentes de los EE.UU. 5.965.408 y 5.939.250), mutagénesis de conjunto recursiva (véase, por ejemplo, Arkin *et al.*, *Proc. Natl. Acad. Sci., U.S.A.*, 89: 7811-7815 (1992)) y mutagénesis de conjunto exponencial (véase, por ejemplo, Delegrave *et al.*, *Biotech. Res.*, 11: 1548-1552 (1993)).

También pueden crearse variantes por mutagénesis *in vivo*. En algunas realizaciones, se generan mutaciones aleatorias en una secuencia de ácido nucleico propagando la secuencia en una cepa bacteriana, tal como una cepa de *E. coli*, que porta mutaciones en una o más de las vías de reparación del ADN. Dichas cepas "mutadoras" tienen una tasa de mutación aleatoria más alta que la de una cepa de tipo silvestre. La propagación de una secuencia de ADN (por ejemplo, una secuencia polinucleotídica que codifica una PPTasa) en una de estas cepas finalmente generará mutaciones aleatorias en el ADN. Se describen cepas mutadoras adecuadas para su uso en mutagénesis *in vivo*, por ejemplo, en la Publicación de Solicitud de Patente Internacional N.º WO 1991/016427.

Las variantes también se pueden generar usando mutagénesis en casete. En la mutagénesis en casete, una pequeña

región de una molécula de ADN bicatenaria se reemplaza con un "casete" de oligonucleótido sintético que difiere de la secuencia nativa. El oligonucleótido suele contener una secuencia nativa completa y/o parcialmente aleatoria.

Como se usa en el presente documento, una "célula hospedadora" es una célula utilizada para producir un producto que se describe en el presente documento (por ejemplo, un aldehído graso o un alcohol graso). En cualquiera de los aspectos de la divulgación descritos en el presente documento, la célula hospedadora puede seleccionarse del grupo que consiste en una célula de mamífero, célula vegetal, célula de insecto, célula fúngica (por ejemplo, una célula fúngica filamentosa o una célula de levadura) y célula bacteriana. En la presente invención, la célula hospedadora es una célula bacteriana. Una célula hospedadora se denomina "célula hospedadora modificada por ingeniería genética" o "célula hospedadora recombinada" si la expresión de uno o más polinucleótidos o polipéptidos en la célula hospedadora se altera o modifica en comparación con su expresión en una célula hospedadora de tipo silvestre correspondiente en las mismas condiciones.

En algunas realizaciones, la célula hospedadora es una célula bacteriana grampositiva. En otras realizaciones, la célula hospedadora es una célula bacteriana gramnegativa.

En algunos aspectos, la célula hospedadora se selecciona entre el género *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*, *Pichia*, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Stenotrophomonas*, *Schizosaccharomyces*, *Yarrowia* o *Streptomyces*.

En otras realizaciones, la célula hospedadora es una célula de *Bacillus lentus*, una célula de *Bacillus brevis*, una célula de *Bacillus stearothermophilus*, una célula de *Bacillus licheniformis*, una célula de *Bacillus alkalophilus*, una célula de *Bacillus coagulans*, una célula de *Bacillus circulans*, una célula de *Bacillus pumilis*, una célula de *Bacillus thuringiensis*, una célula de *Bacillus clausii*, una célula de *Bacillus megaterium*, una célula de *Bacillus subtilis* o una célula de *Bacillus amyloliquefaciens*.

En otros aspectos, la célula hospedadora es una célula de *Trichoderma koningii*, una célula de *Trichoderma viride*, una célula de *Trichoderma reesei*, una célula de *Trichoderma longibrachiatum*, una célula de *Aspergillus awamori*, una célula de *Aspergillus fumigates*, una célula de *Aspergillus foetidus*, una célula de *Aspergillus nidulans*, una célula de *Aspergillus niger*, una célula de *Aspergillus oryzae*, una célula de *Humicola insolens*, una célula de *Humicola lanuginosa*, una célula de *Rhodococcus opacus*, una célula de *Rhizomucor miehei* o una célula de *Mucor michei*.

En otras realizaciones más, la célula hospedadora es una célula de *Streptomyces lividans* o una célula de *Streptomyces murinus*.

En otras realizaciones más, la célula hospedadora es una célula de *Actinomycetes*.

En algunos aspectos, la célula hospedadora es una célula de *Saccharomyces cerevisiae*. En algunos aspectos, la célula hospedadora es una célula de *Saccharomyces cerevisiae*.

En otros aspectos más, la célula hospedadora es una célula CHO, una célula COS, una célula VERO, una célula BHK, una célula HeLa, una célula Cvl, una célula MDCK, una célula 293, una célula 3T3 o una célula PC12.

En otros aspectos, la célula hospedadora es una célula de una planta eucariota, alga, cianobacteria, bacteria verde de azufre, bacteria verde no de azufre, bacteria púrpura de azufre, bacteria púrpura no de azufre, extremófilo, levadura, hongo, un organismo modificado por ingeniería genética de la misma o un organismo sintético. En algunos aspectos, la célula hospedadora depende de la luz o fija el carbono. En algunos aspectos, la célula hospedadora tiene actividad autótrofa. En algunos aspectos, la célula hospedadora tiene actividad fotoautótrofa, tal como en presencia de luz. En algunos aspectos, la célula hospedadora es heterótrofa o mixotrófica en ausencia de luz. En determinados aspectos, la célula hospedadora es una célula de *Avicennia marina*, *Panicum virgatum*, *Miscanthus giganteus*, *Zea mays*, *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Dunaliella salina*, *Synechococcus* Sp. PCC 7002, *Synechococcus* Sp. PCC 7942, *Synechocystis* Sp. PCC 6803, *Thermosynechococcus elongates* BP-1, *Chlorobium tepidum*, *Chloroflexus auranticus*, *Chromatium vinosum*, *Rhodospirillum rubrum*, *Rhodobacter capsulatus*, *Rhodopseudomonas palustris*, *Clostridium ljungdahlii*, *Clostridium thermocellum*, *Penicillium chrysogenum*, *Pichia pastoris*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pseudomonas fluorescens* o *Zymomonas mobilis*.

En determinadas realizaciones preferidas, la célula hospedadora es una célula de *E. coli*. En algunas realizaciones, la célula de *E. coli* es una célula de *E. coli* cepa B, una cepa C, una cepa K o cepa W.

En otras realizaciones, la célula hospedadora es una célula de *Pantoea citrea*.

Como se usa en el presente documento, la expresión "condiciones permisivas para la producción" significa cualesquier condiciones que permitan a una célula hospedadora producir un producto deseado, tal como un ácido graso o un derivado de un ácido graso. De la misma manera, la expresión "condiciones en las que se expresa la secuencia polinucleotídica de un vector" significa cualesquier condiciones que permitan que una célula hospedadora sintetice un

polipéptido. Las condiciones adecuadas incluyen, por ejemplo, las condiciones de fermentación. Las condiciones de fermentación pueden comprender muchos parámetros, tales como intervalos de temperatura, niveles de aireación y composición de los medios. Cada una de estas condiciones, de forma individual y en combinación, permite que la célula hospedadora crezca. Los medios de cultivo ilustrativos incluyen caldos o geles. Generalmente, el medio incluye una fuente de carbono que puede ser metabolizada directamente por una célula hospedadora. Además, pueden usarse enzimas en el medio para facilitar la movilización (por ejemplo, la despolimerización de almidón o celulosa a azúcares fermentables) y el metabolismo posterior de la fuente de carbono.

Como se usa en el presente documento, la frase "fuente de carbono" se refiere a un sustrato o compuesto adecuado para usarse como fuente de carbono para el crecimiento de células procarióticas o eucariotas simples. Las fuentes de carbono pueden ser de diferentes formas, incluyendo, pero sin limitación, polímeros, hidratos de carbono, ácidos, alcoholes, aldehídos, cetonas, aminoácidos, péptidos y gases (por ejemplo, CO y CO₂). Los ejemplos de fuentes de carbono incluyen, pero sin limitación, monosacáridos, tales como glucosa, fructosa, manosa, galactosa, xilosa y arabinosa; oligosacáridos, tales como fructo-oligosacárido y galacto-oligosacárido; polisacáridos, tales como almidón, celulosa, pectina y xilano; disacáridos, tales como sacarosa, maltosa y turanosa; material celulósico y variantes tales como metilcelulosa y carboximetilcelulosa de sodio; ésteres de ácidos grasos saturados o insaturados, succinato, lactato y acetato; alcoholes, tales como etanol, metanol y glicerol, o mezclas de los mismos. Además, la fuente de carbono puede ser un producto de la fotosíntesis, tal como glucosa. En determinadas realizaciones preferidas, la fuente de carbono es biomasa. En otras realizaciones preferidas, la fuente de carbono es glucosa. En otras realizaciones preferidas, la fuente de carbono es sacarosa.

Como se usa en el presente documento, El término "biomasa" se refiere a cualquier material biológico del que deriva una fuente de carbono. En algunas realizaciones, una biomasa se procesa en una fuente de carbono, que es adecuada para la bioconversión. En otras realizaciones, la biomasa no requiere procesamiento adicional en una fuente de carbono. La fuente de carbono puede convertirse en un biocombustible. Una fuente ilustrativa de biomasa es materia vegetal o vegetación, tal como de maíz, caña de azúcar o césped de pradera. Otra fuente ilustrativa de biomasa son los productos de desecho metabólicos, tales como materia animal (por ejemplo, estiércol de vaca). Otras fuentes ilustrativas de biomasa incluyen algas y otras plantas marinas. La biomasa también incluye productos de desecho de la industria, agricultura, la silvicultura y los hogares, incluyendo, pero sin limitación, residuos de fermentación, ensilado, paja, madera no procesada, aguas residuales, basura, restos celulósicos urbanos y restos de comida. El término "biomasa" también puede referirse a fuentes de carbono, tales como hidratos de carbono (por ejemplo, monosacáridos, disacáridos o polisacáridos).

Como se usa en el presente documento, el término "clon" normalmente se refiere a una célula o a un grupo de células descendientes de y esencialmente idénticas genéticamente a un único ancestro común, por ejemplo, las bacterias de una colonia bacteriana clonada surgieron de una sola célula bacteriana.

Como se usa en el presente documento, el término "cultivo" típico se refiere a un medio líquido que comprende células viables. En una realización, un cultivo comprende células que se reproducen en un medio de cultivo predeterminado en condiciones controladas, por ejemplo, un cultivo de células hospedadoras recombinantes cultivadas en medios líquidos que comprenden una fuente de carbono y nitrógeno seleccionados.

Por "cultivo" o "cultivo" se entiende el cultivo de una población de células hospedadoras recombinantes en condiciones adecuadas en un medio líquido o sólido. En realizaciones particulares, cultivar se refiere a la bioconversión fermentativa de un sustrato hasta la obtención de un producto final. Los medios de cultivo son bien conocidos, y los componentes individuales de dichos medios de cultivo están disponibles en fuentes comerciales, por ejemplo, con las marcas comerciales Difco™ y BBL™. En un ejemplo no limitante, el medio de nutrientes acuoso es un "medio rico" que comprende fuentes complejas de nitrógeno, sales y carbono, tales como el medio YP, que comprende 10 g/l de peptona y 10 g/l de extracto de levadura de dicho medio.

Para determinar si las condiciones son suficientes para permitir la producción de un producto o la expresión de un polipéptido, una célula hospedadora puede cultivarse, por ejemplo, durante aproximadamente 4, 8, 12, 24, 36, 48, 72 o más horas. Durante y/o después del cultivo, pueden obtenerse y analizarse muestras para determinar si las condiciones permiten la producción o la expresión. Por ejemplo, las células hospedadoras en la muestra o el medio en el que se cultivaron las células hospedadoras pueden someterse a ensayo para detectar la presencia de un producto deseado. Cuando se somete a ensayo la presencia de un ácido graso o un derivado de un ácido graso, pueden usarse ensayos, tales como, pero sin limitación, EM, cromatografía de capa fina (CCF), cromatografía líquida de alto rendimiento (HPLC), cromatografía líquida (CL), CG acoplada a un detector de ionización de llama (DIL), CG-EM y CL-EM. Cuando se somete a ensayo para determinar la expresión de un polipéptido, pueden usarse técnicas tales como, pero sin limitación, transferencia Western y transferencia puntual.

En los métodos de la invención, la producción y el aislamiento de ácidos grasos y derivados de ácidos grasos pueden potenciarse optimizando las condiciones de fermentación. En algunas realizaciones, las condiciones de fermentación se optimizan para aumentar el porcentaje de la fuente de carbono que se convierte en productos hidrocarbonados. Durante los ciclos de vida celular normales, el carbono se usa en funciones celulares, tales como la producción de lípidos, sacáridos, proteínas, ácidos orgánicos y ácidos nucleicos. La reducción de la cantidad de carbono necesaria

para las actividades relacionadas con el crecimiento puede aumentar la eficacia de la conversión de la fuente de carbono en producto. Esto puede conseguirse mediante, por ejemplo, en primer lugar cultivar las células hospedadoras hasta una densidad deseada (por ejemplo, una densidad alcanzada en el pico de la fase logarítmica de crecimiento). En dicho punto, los genes de punto de control de la replicación pueden aprovecharse para detener el crecimiento de las células. Específicamente, pueden usarse mecanismos de detección de quórum (revisado en Camilli *et al.*, *Science* 311: 1113 (2006); Venturi, *FEMS Microbiol. Rev.*, 30: 274-291 (2006); y Reading *et al.*, *FEMS Microbiol. Lett.*, 254: 1-11 (2006)) para activar los genes de punto de control, tales como *p53*, *p21* u otros genes de punto de control.

Los genes que pueden activarse para detener la replicación celular y el crecimiento de *E. coli* incluyen los genes de *umuDC*. La sobreexpresión de los genes de *umuDC* detiene la progresión de la fase estacionaria al crecimiento exponencial (Murli *et al.*, *J. Bacteriol.*, 182: 1127-1135 (2000)). La *UmuC* es una ADN polimerasa que puede realizar la síntesis por translesión sobre las lesiones no codificantes que habitualmente son resultado de la mutagénesis ultravioleta (UV) y química. Los productos génicos de *umuDC* están implicados en el proceso de síntesis por translesión y también sirven como punto de control del daño de secuencia de ADN. Los productos génicos de *umuDC* incluyen *UmuC*, *UmuD*, *umuD'*, *UmuD'2C*, *UmuD'2* y *UmuD2*. Simultáneamente, los genes productores de productos pueden activarse, de manera que se minimiza la necesidad de usar vías de replicación y mantenimiento mientras se prepara un aldehído graso o alcohol graso. También pueden modificarse por ingeniería genética células hospedadoras para que expresen *umuC* y *umuD* de *E. coli* en *pBAD24* bajo el sistema del promotor de *prpBCDE* a través de síntesis *de novo* de este gen con los genes de producción de productos finales apropiados.

La célula hospedadora puede modificarse por ingeniería genética adicionalmente para expresar un celulosoma recombinante, que puede permitir que la célula hospedadora use material celulósico como fuente de carbono. Los celulosomas de ejemplo adecuados para su uso en los métodos de la invención incluyen, por ejemplo, los celulosomas descritos en la Publicación de Solicitud de Patente Internacional WO 2008/100251. La célula hospedadora también puede modificarse por ingeniería genética para asimilar el carbono de manera eficiente y usar materiales celulósicos como fuentes de carbono de acuerdo con los métodos descritos en las Patentes de los EE.UU. 5.000.000; 5.028.539; 5.424.202; 5.482.846; y 5.602.030. Además, la célula hospedadora se puede diseñar genéticamente para expresar una invertasa a fin de poderse usar la sacarosa como fuente de carbono.

En algunas realizaciones de los métodos de fermentación de la invención, la cámara de fermentación encierra una fermentación que experimenta una reducción continua, creando de este modo un entorno reductor estable. El equilibrio de electrones puede mantenerse mediante la liberación de dióxido de carbono (en forma gaseosa). Los esfuerzos para aumentar el equilibrio de NAD/H y NADP/H también pueden facilitar la estabilización del equilibrio de electrones. La disponibilidad del NADPH intracelular también puede potenciarse mediante ingeniería genética de la célula hospedadora para expresar una transhidrogenasa NADH:NADPH. La expresión de una o más transhidrogenasas NADH:NADPH convierte el NADH producido en la glicólisis en NADPH, que puede potenciar la producción de aldehídos grasos y alcoholes grasos.

Para la producción a pequeña escala, las células hospedadoras diseñadas se pueden cultivar en lotes de, por ejemplo, aproximadamente 100 ml, 500 ml, 1 l, 2 l, 5 l o 10 l; fermentarse; e inducirse para expresar una secuencia polinucleotídica deseada, tal como una secuencia polinucleotídica que codifica una PPTasa. Para la producción a gran escala, las células hospedadoras modificadas por ingeniería genética pueden crecer en lotes de aproximadamente 10 l, 100 l, 1000 l, 10.000 l, 100.000 l, 1.000.000 l o más grandes; fermentarse; e inducirse para expresar una secuencia polinucleotídica deseada.

Los ácidos grasos y derivados de los mismos producidos mediante los métodos de la invención generalmente se aíslan de la célula hospedadora. El término "aislado", como se usa en el presente documento con respecto a los productos, tales como ácidos grasos y derivados de los mismos, se refiere a los productos que se separan de los componentes celulares, medios de cultivo celular, o precursores químicos o sintéticos. Los ácidos grasos y derivados de los mismos producidos mediante los métodos descritos en el presente documento pueden ser relativamente inmiscibles en el caldo de fermentación, así como en el citoplasma. Por tanto, los ácidos grasos y sus derivados pueden acumularse en una fase orgánica, ya sea intracelular o extracelularmente. La recogida de los productos en la fase orgánica puede reducir el impacto del ácido graso o derivado de ácido graso en la función celular y puede permitir a la célula hospedadora producir más producto.

En algunas realizaciones, los ácidos grasos y derivados de ácidos grasos producidos mediante los métodos de la invención se purifican. Como se usa en el presente documento, el término "purificar", "purificado/a" o "purificación" significan la eliminación o el aislamiento de una molécula de su entorno mediante, por ejemplo, aislamiento o separación. Las moléculas "esencialmente purificadas" están al menos aproximadamente un 60 % libres (por ejemplo, al menos aproximadamente un 70 % libres, al menos aproximadamente un 75 % libres, al menos aproximadamente un 85 % libres, al menos aproximadamente un 90 % libres, al menos aproximadamente un 95 % libres, al menos aproximadamente un 97 % libres, al menos aproximadamente un 99 % libres) de otros componentes con los que están asociados. Como se usa en el presente documento, estos términos también se refieren a la eliminación de contaminantes de una muestra. Por ejemplo, la eliminación de contaminantes puede dar como resultado un aumento del porcentaje de un aldehído graso o un alcohol graso en una muestra. Por ejemplo, cuando se produce un aldehído graso o un alcohol graso en una célula hospedadora, el aldehído graso o el alcohol graso puede purificarse mediante

la eliminación de las proteínas de la célula hospedadora. Después de la purificación, el porcentaje de un ácido graso o derivado del mismo en la muestra se aumenta.

Como se usa en el presente documento, Los términos "purificar", "purificado/a" y "purificación" son términos relativos que no requieren una pureza absoluta. Por tanto, por ejemplo, cuando se produce un ácido graso o un derivado del mismo en las células hospedadoras, un ácido graso o derivado del mismo purificado es un ácido graso o un derivado del mismo que se separa sustancialmente de otros componentes celulares (por ejemplo, ácidos nucleicos, polipéptidos, lípidos, carbohidratos u otros hidrocarburos).

Adicionalmente, una preparación de ácidos grasos purificados o una preparación de derivados de ácidos grasos purificados es una preparación de ácidos grasos o una preparación de derivados de ácidos grasos en las que el ácido graso o el derivado del mismo está sustancialmente libre de contaminantes, como los que podría haber después de la fermentación. En algunas realizaciones, se purifica un ácido graso o derivado del mismo cuando al menos el 50 % en peso de una muestra está compuesto por el ácido graso o derivado de ácido graso. En otras realizaciones, un ácido graso o un derivado del mismo se purifica cuando al menos aproximadamente el 60 %, por ejemplo, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 92 % o más en peso de una muestra está compuesto por el ácido graso o derivado del mismo. Como alternativa, o además, un ácido graso o derivado del mismo se purifica cuando menos de aproximadamente el 100 %, por ejemplo, menos de aproximadamente el 99 %, menos de aproximadamente el 98 %, menos de aproximadamente el 95 %, menos de aproximadamente el 90 % o menos de aproximadamente el 80 % en peso de una muestra está compuesto por el ácido graso o derivado del mismo. Por tanto, un ácido graso o derivado del mismo purificado puede tener un nivel de pureza limitado por cualquiera de los dos criterios de valoración anteriores. Por ejemplo, un ácido graso o derivado del mismo puede purificarse cuando al menos aproximadamente el 80 %-95 %, al menos aproximadamente el 85 %-99 % o al menos aproximadamente el 90 %-98 % de una muestra está compuesto por el ácido graso o derivado de ácido graso.

El ácido graso o derivado del mismo puede estar presente en el ambiente extracelular, o puede aislarse del ambiente extracelular de la célula hospedadora. En determinadas realizaciones, un ácido graso o derivado del mismo es secretado por la célula hospedadora. En otras realizaciones, un ácido graso o derivado del mismo es transportado al medio extracelular. En otras realizaciones más, el ácido graso o derivado del mismo es transportado pasivamente al medio extracelular. Un ácido graso o derivado del mismo puede aislarse de una célula hospedadora usando métodos conocidos en la técnica, tales como los desvelados en las Publicaciones de Solicitud de Patente Internacional WO 2010/042664 y WO 2010/062480.

Los métodos que se describen en el presente documento pueden dar como resultado la producción de compuestos homogéneos en donde al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 90 % o al menos aproximadamente el 95 %, de los ácidos grasos o derivados de ácidos grasos producidos tendrán longitudes de cadena de carbono que varían en menos de 6 carbonos, menos de 5 carbonos, menos de 4 carbonos, menos de 3 carbonos o menos de aproximadamente 2 carbonos. Como alternativa, o además, los métodos que se describen en el presente documento pueden dar como resultado la producción de compuestos homogéneos en donde menos de aproximadamente el 98 %, menos de aproximadamente el 95 %, menos de aproximadamente el 90 %, menos de aproximadamente el 80 % o menos de aproximadamente el 70 % de los ácidos grasos o derivados de ácidos grasos producidos tendrán longitudes de cadena de carbono que varían en menos de 6 carbonos, menos de 5 carbonos, menos de 4 carbonos, menos de 3 carbonos o menos de aproximadamente 2 carbonos. Por tanto, los ácidos grasos o derivados de ácidos grasos pueden tener un grado de homogeneidad limitado por cualquiera de los dos criterios de valoración anteriores. Por ejemplo, el ácido graso o derivado de ácido graso puede tener un grado de homogeneidad en donde aproximadamente el 70 %-95 %, aproximadamente el 80 %-98 % o aproximadamente el 90 %-95 % de los ácidos grasos o derivados de ácidos grasos producidos tendrán longitudes de cadena de carbono que varían en menos de 6 carbonos, menos de 5 carbonos, menos de 4 carbonos, menos de 3 carbonos o menos de aproximadamente 2 carbonos. Estos compuestos también pueden producirse con un grado de saturación relativamente uniforme.

Como resultado de los métodos de la presente invención, uno o más de entre el título, el rendimiento o la productividad del ácido graso o derivado del mismo producido por la célula hospedadora modificada por ingeniería genética que tiene un nivel de expresión alterado de un polipéptido FadR aumenta con respecto al de la célula hospedadora de tipo silvestre correspondiente.

El término "título" se refiere a la cantidad de ácido graso o derivado de ácido graso producida por unidad de volumen de cultivo de célula hospedadora. En cualquier aspecto de las composiciones y de los métodos descritos en el presente documento, un ácido graso o un derivado de ácido graso tal como una olefina terminal, un aldehído graso, un alcohol graso, un alcano, un éster graso, una cetona o una olefina interna se produce en un título de aproximadamente 25 mg/l, aproximadamente 50 mg/l, aproximadamente 75 mg/l, aproximadamente 100 mg/l, aproximadamente 125 mg/l, aproximadamente 150 mg/l, aproximadamente 175 mg/l, aproximadamente 200 mg/l, aproximadamente 225 mg/l, aproximadamente 250 mg/l, aproximadamente 275 mg/l, aproximadamente 300 mg/l, aproximadamente 325 mg/l, aproximadamente 350 mg/l, aproximadamente 375 mg/l, aproximadamente 400 mg/l, aproximadamente 425 mg/l, aproximadamente 450 mg/l, aproximadamente 475 mg/l, aproximadamente 500 mg/l, aproximadamente 525 mg/l,

aproximadamente 550 mg/l, aproximadamente 575 mg/l, aproximadamente 600 mg/l, aproximadamente 625 mg/l, aproximadamente 650 mg/l, aproximadamente 675 mg/l, aproximadamente 700 mg/l, aproximadamente 725 mg/l, aproximadamente 750 mg/l, aproximadamente 775 mg/l, aproximadamente 800 mg/l, aproximadamente 825 mg/l, aproximadamente 850 mg/l, aproximadamente 875 mg/l, aproximadamente 900 mg/l, aproximadamente 925 mg/l, aproximadamente 950 mg/l, aproximadamente 975 mg/l, aproximadamente 1000 g/l, aproximadamente 1050 mg/l, aproximadamente 1075 mg/l, aproximadamente 1100 mg/l, aproximadamente 1125 mg/l, aproximadamente 1150 mg/l, aproximadamente 1175 mg/l, aproximadamente 1200 mg/l, aproximadamente 1225 mg/l, aproximadamente 1250 mg/l, aproximadamente 1275 mg/l, aproximadamente 1300 mg/l, aproximadamente 1325 mg/l, aproximadamente 1350 mg/l, aproximadamente 1375 mg/l, aproximadamente 1400 mg/l, aproximadamente 1425 mg/l, aproximadamente 1450 mg/l, aproximadamente 1475 mg/l, aproximadamente 1500 mg/l, aproximadamente 1525 mg/l, aproximadamente 1550 mg/l, aproximadamente 1575 mg/l, aproximadamente 1600 mg/l, aproximadamente 1625 mg/l, aproximadamente 1650 mg/l, aproximadamente 1675 mg/l, aproximadamente 1700 mg/l, aproximadamente 1725 mg/l, aproximadamente 1750 mg/l, aproximadamente 1775 mg/l, aproximadamente 1800 mg/l, aproximadamente 1825 mg/l, aproximadamente 1850 mg/l, aproximadamente 1875 mg/l, aproximadamente 1900 mg/l, aproximadamente 1925 mg/l, aproximadamente 1950 mg/l, aproximadamente 1975 mg/l, aproximadamente 2000 mg/l o un intervalo limitado por cualquiera de los dos valores anteriores. En otras realizaciones, un ácido graso o derivado de ácido graso se produce en un título de más de 2000 mg/l, más de 5000 mg/l, más de 10.000 mg/l o más, tal como 50 g/l, 70 g/l, 100 g/l, 120 g/l, 150 g/l o 200 g/l.

Como se usa en el presente documento, el "rendimiento de derivado de ácido graso producido por una célula hospedadora" se refiere a la eficiencia con la que una fuente de carbono de entrada se convierte en producto (es decir, alcohol graso o aldehído graso) en una célula hospedadora. Las células hospedadora modificada por ingeniería genética para producir derivados de ácidos grasos de acuerdo con los métodos de la invención tienen un rendimiento de al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 6 %, al menos un 7 %, al menos un 8 %, al menos un 9 %, al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 %, al menos un 24 %, al menos un 25 %, al menos un 26 %, al menos un 27 %, al menos un 28 %, al menos un 29 % o al menos un 30 % o un intervalo limitado por cualquiera de los dos valores anteriores. En otras realizaciones, se produce un derivado o derivados de ácidos grasos con un rendimiento superior al 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o más. Como alternativa, o además, el rendimiento es de aproximadamente el 30 % o inferior, de aproximadamente el 27 % o inferior, de aproximadamente el 25 % o inferior, o de aproximadamente el 22 % o inferior. Por tanto, el rendimiento puede estar limitado por dos cualquiera de los criterios de valoración anteriores. Por ejemplo, el rendimiento de un derivado o derivados de ácido graso producidos por la célula hospedadora recombinante de acuerdo con los métodos de la invención puede ser del 5 % al 15 %, del 10 % al 25 %, del 10 % al 22 %, del 15 % al 27 %, del 18 % al 22 %, del 20 % al 28 % o del 20 % al 30 %. El rendimiento puede referirse a un derivado de ácido graso particular o a una combinación de derivados de ácidos grasos producidos por un cultivo de células hospedadora recombinantes dado.

En un enfoque, la expresión "productividad del ácido graso o derivado del mismo producido por una célula hospedadora" se refiere a la cantidad de ácido graso o derivado de ácido graso producido por unidad de volumen de cultivo de célula hospedadora por unidad de densidad de cultivo de célula hospedadora. En cualquier aspecto de las composiciones y de los métodos descritos en el presente documento, la productividad de un ácido graso o un derivado de ácido graso tal como una olefina, un aldehído graso, un alcohol graso, un alcano, un éster graso o una cetona producidos por una célula hospedadora modificada por ingeniería genética es de al menos aproximadamente 3 mg/l/DO₆₀₀, al menos aproximadamente 6 mg/l/DO₆₀₀, al menos aproximadamente 9 mg/l/DO₆₀₀, al menos aproximadamente 12 mg/l/DO₆₀₀ o al menos aproximadamente 15 mg/l/DO₆₀₀. Como alternativa, o además, la productividad es de aproximadamente 50 mg/l/DO₆₀₀ o menos, aproximadamente 40 mg/l/DO₆₀₀ o menos, aproximadamente 30 mg/l/DO₆₀₀ o menos o aproximadamente 20 mg/l/DO₆₀₀ o menos. Por tanto, la productividad puede estar limitada por dos cualquiera de los criterios de valoración anteriores. Por ejemplo, la productividad puede ser de aproximadamente 3 a aproximadamente 30 mg/l/DO₆₀₀, de aproximadamente 6 a aproximadamente 20 mg/l/DO₆₀₀ o de aproximadamente 15 a aproximadamente 30 mg/l/DO₆₀₀.

En otro enfoque, el término "productividad" se refiere a la cantidad de un derivado o derivados de ácidos grasos producidos por unidad de volumen de cultivo de células hospedadoras por unidad de tiempo. En cualquier aspecto de las composiciones y de los métodos descritos en el presente documento, la productividad de un derivado o derivados de ácidos grasos producidos por una célula hospedadora recombinante es de al menos 100 mg/l/hora, al menos 200 mg/l/hora, al menos 300 mg/l/hora, al menos 400 mg/l/hora, al menos 500 mg/l/hora, al menos 600 mg/l/hora, al menos 700 mg/l/hora, al menos 800 mg/l/hora, al menos 900 mg/l/hora, al menos 1000 mg/l/hora, al menos 1100 mg/l/hora, al menos 1200 mg/l/hora, al menos 1300 mg/l/hora, al menos 1400 mg/l/hora, al menos 1500 mg/l/hora, al menos 1600 mg/l/hora, al menos 1700 mg/l/hora, al menos 1800 mg/l/hora, al menos 1900 mg/l/hora, al menos 2000 mg/l/hora, al menos 2100 mg/l/hora, al menos 2200 mg/l/hora, al menos 2300 mg/l/hora, al menos 2400 mg/l/hora o al menos 2500 mg/l/hora. Como alternativa, o además, la productividad es de 2500 mg/l/hora o menos, 2000 mg/l/DO₆₀₀ o menos, 1500 mg/l/DO₆₀₀ o menos, 120 mg/l/hora o menos, 1000 mg/l/hora o menos, 800 mg/l/hora o menos, o 600 mg/l/hora o menos. Por tanto, la productividad puede estar limitada por dos cualquiera de los criterios de valoración anteriores. Por ejemplo, la productividad puede ser de 3 a

30 mg/l/hora, de 6 a 20 mg/l/hora o de 15 a 30 mg/l/hora. Por ejemplo, la productividad de un derivado o derivados de ácidos grasos producidos por una célula hospedadora recombinante de acuerdo con los métodos puede ser de 500 mg/l/hora a 2500 mg/l/hora, o de 700 mg/l/hora a 2000 mg/l/hora. La productividad puede referirse a un derivado de ácido graso particular o a una combinación de derivados de ácidos grasos producidos por un cultivo de células hospedadora recombinantes dado.

En los métodos de la invención, la producción y el aislamiento de un ácido graso deseado o derivado del mismo (por ejemplo, acil-CoA, ácidos grasos, olefinas terminales, aldehídos grasos, alcoholes grasos, alcanos, alquenos, ésteres de cera, cetonas y olefinas internas) puede potenciarse alterando la expresión de uno o más genes implicados en la regulación de la producción, degradación y/o secreción de ácido graso, éster graso, alcano, alqueno, olefina, alcohol graso, en la célula hospedadora modificada por ingeniería genética.

Se sabe que FadR modula la expresión y/o actividad de numerosos genes, incluyendo *fabA*, *fabB*, *iclR*, *fadA*, *fadB*, *fadD*, *fadE*, *fadI*, *fadJ*, *fadL*, *fadM*, *uspA*, *aceA*, *aceB* y *aceK*. En algunas realizaciones de los métodos que se describen en el presente documento, la célula hospedadora modificada por ingeniería genética comprende adicionalmente un nivel alterado de expresión de uno o más genes seleccionados entre el grupo que consiste en *fabA*, *fabB*, *iclR*, *fadA*, *fadB*, *fadD*, *fadE*, *fadI*, *fadJ*, *fadL*, *fadM*, *uspA*, *aceA*, *aceB* y *aceK* en comparación con el nivel de expresión del gen o los genes seleccionados en una célula hospedadora de tipo silvestre correspondiente. Los números de registro de ejemplo para polipéptidos codificados por los genes diana de FadR incluyen *fabA* (NP_415474), *fabB* (BAA16180), (NP_418442), *fadA* (YP_026272.1), *fadB* (NP_418288.1), *fadD* (AP_002424), *fadE* (NP_414756.2), *fadI* (NP_416844.1), *fadJ* (NP_416843.1), *fadL* (AAC75404), *fadM* (NP_414977.1), *uspA* (AAC76520), *aceA* (AAC76985.1), *aceB* (AAC76984.1) y *aceK* (AAC76986.1).

En la FIG. 1 se enumeran enzimas y polipéptidos de ejemplo para su uso en la práctica de la invención. Un experto habitual en la materia comprenderá que dependiendo del fin (por ejemplo, el ácido graso deseado o el producto derivado de ácido graso), los genes específicos (o combinaciones de genes) enumerados en la FIG. 1 pueden sobreexpresarse, modificarse, atenuarse o suprimirse en una célula hospedadora que tiene un nivel de expresión alterado de un polipéptido FadR.

En algunas realizaciones, el método comprende modificar la expresión de un gen que codifica una o más de entre una tioesterasa (por ejemplo, *TesA*), una descarboxilasa, una ácido carboxílico reductasa (*CAR*; por ejemplo, *CarB*), una alcohol deshidrogenasa (aldehído reductasa); una aldehído descarboxilasa, una acil-CoA reductasa que forma alcohol graso (*FAR*), una acil ACP reductasa (*AAR*), una éster sintasa, una acil-CoA reductasa (*ACR1*), *OleA*, *OleCD* y *OleBCD*.

En determinadas realizaciones de la invención, la célula hospedadora modificada por ingeniería genética que tiene un nivel de expresión alterado de un polipéptido FadR puede modificarse por ingeniería genética para que comprenda adicionalmente una secuencia polinucleotídica que codifique un polipéptido: (1) que tenga actividad tioesterasa (CE 3.1.2.14), en donde la célula hospedadora modificada por ingeniería genética sintetiza ácidos grasos; (2) que tenga actividad descarboxilasa, en donde la célula hospedadora modificada por ingeniería genética sintetiza olefinas terminales; (3) que tenga actividad ácido carboxílico reductasa, en donde la célula hospedadora modificada por ingeniería genética sintetiza aldehídos grasos; (4) que tenga actividad ácido carboxílico reductasa y alcohol deshidrogenasa (CE 1.1.1.1), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcoholes grasos; (5) que tenga actividad ácido carboxílico reductasa y aldehído descarboxilasa (CE 4.1.99.5), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcanos; (6) que tenga actividad acil-CoA reductasa (CE 1.2.1.50), en donde el microorganismo sintetiza aldehídos grasos; (7) que tenga actividad acil-CoA reductasa (CE 1.2.1.50) y actividad alcohol deshidrogenasa (CE 1.1.1.1), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcoholes grasos; (8) que tenga actividad acil-CoA reductasa (CE 1.2.1.50) y actividad aldehído descarboxilasa (CE 4.1.99.5), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcanos; (9) que tenga actividad acil CoA reductasa que forma alcohol en donde la célula hospedadora modificada por ingeniería genética sintetiza aldehídos grasos y alcoholes grasos; (10) que tenga actividad ácido carboxílico reductasa, en donde la célula hospedadora modificada por ingeniería genética sintetiza aldehídos grasos; (11) que tenga actividad acil ACP reductasa, en donde la célula hospedadora modificada por ingeniería genética sintetiza aldehídos grasos; (12) que tenga actividad acil ACP reductasa y actividad alcohol deshidrogenasa (CE 1.1.1.1), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcoholes grasos; (13) que tenga actividad acil ACP reductasa y actividad aldehído descarboxilasa (CE 4.1.99.5), en donde la célula hospedadora modificada por ingeniería genética sintetiza alcanos; (14) que tenga actividad éster sintasa (CE 3.1.1.67), en donde la célula hospedadora modificada por ingeniería genética sintetiza ésteres grasos; (15) que tenga actividad éster sintasa (CE 3.1.1.67) y (a) actividad ácido carboxílico reductasa, (b) actividad acil-CoA reductasa, (c) actividad acil ACP reductasa o (d) actividad alcohol deshidrogenasa (CE 1.1.1.1), en donde la célula hospedadora modificada por ingeniería genética sintetiza ésteres de cera; (16) que tenga actividad *OleA*, en donde la célula hospedadora modificada por ingeniería genética sintetiza 2-alquil-3-ceto-acil CoA y cetonas; o (17) que tenga actividad *OleCD* u *OleBCD*, en donde la célula hospedadora modificada por ingeniería genética sintetiza olefinas internas.

En algunas realizaciones, el método comprende adicionalmente modificar la expresión de un gen que codifica una ácido graso sintasa en la célula hospedadora. Como se usa en el presente documento, "ácido graso sintasa" significa

cualquier enzima implicada en la biosíntesis de ácidos grasos. En determinadas realizaciones, la modificación de la expresión de un gen que codifica una ácido graso sintasa incluye la expresión de un gen que codifica una ácido graso sintasa en la célula hospedadora y/o el aumento de la expresión o la actividad de una ácido graso sintasa endógena en la célula hospedadora. En realizaciones alternativas, la modificación de la expresión de un gen que codifica una

5 ácido graso sintasa incluye la atenuación de un gen que codifica una ácido graso sintasa en la célula hospedadora y/o la disminución de la expresión o la actividad de una ácido graso sintasa endógena en la célula hospedadora. En algunas realizaciones, la ácido graso sintasa es una tioesterasa (CE 3.1.1.5 o CE 3.1.2.14). En realizaciones particulares, la tioesterasa está codificada por *tesA*, *tesA* sin secuencia líder, *tesB*, *fatB*, *fatB2*, *fatB3*, *fatA* o *fatA1*.

10 En otras realizaciones, la célula hospedadora se modifica genéticamente para expresar un nivel atenuado de una enzima de degradación de ácidos grasos con respecto a una célula hospedadora de tipo silvestre. Como se usa en el presente documento, la expresión "enzima de degradación de ácidos grasos" significa una enzima implicada en la descomposición o la conversión de un ácido graso o derivado de ácido graso en otro producto, tal como, pero sin limitación, una acil-CoA sintasa. En algunas realizaciones, la célula hospedadora se modifica genéticamente para

15 expresar un nivel atenuado de una acil-CoA sintasa con respecto a una célula hospedadora de tipo silvestre. En realizaciones particulares, la célula hospedadora expresa un nivel atenuado de una acil-CoA sintasa codificada por *fadD*, *fadK*, *BH3103*, *yhfl*, *PJI-4354*, *EAV15023*, *fadD1*, *fadD2*, *RPC_4074*, *fadDD35*, *fadDD22*, *faa3p* o el gen que codifica la proteína YP_002028218. En determinadas realizaciones, la célula hospedadora modificada genéticamente comprende una inactivación de uno o más genes que codifican una enzima de degradación de ácidos grasos, tales

20 como los genes de acil-CoA sintasa mencionados anteriormente.

La vía biosintética de ácidos grasos en las células hospedadoras usa los precursores acetil-CoA y malonil-CoA. Las etapas en esta vía son catalizadas por enzimas de las familias de genes de biosíntesis de ácidos grasos (*fab*) y acetil-CoA carboxilasa (*acc*) (véase, por ejemplo, Heath *et al.*, *Prog. Lipid Res.* 40(6): 467-97 (2001)). La acetil-CoA es

25 carboxilada por la acetil-CoA carboxilasa (CE 6.4.1.2), que es una enzima de múltiples subunidades codificada por cuatro genes separados (*accA*, *accB*, *accC* y *accD*) en la mayoría de los procariotas, para formar malonil-CoA. En algunas bacterias, tales como *Corynebacterium glutamicum*, la acetil-CoA carboxilasa consiste en dos subunidades, *AccDA* [YP_225123.1] y *AccBC* [YP_224991], codificadas por *accDA* y *accBC*, respectivamente. Dependiendo del ácido graso deseado o del producto derivado de ácido graso, los genes *fab* y/o *acc* específicos (o combinaciones de

30 los mismos) pueden sobreexpresarse, modificarse, atenuarse o suprimirse en una célula hospedadora modificada por ingeniería genética.

En algunas realizaciones, un complejo acetil-CoA carboxilasa se sobreexpresa en la célula hospedadora modificada por ingeniería genética. En determinadas realizaciones, los genes de subunidades de la acetil-CoA carboxilasa se

35 obtienen de una o más de entre *Corynebacterium glutamicum*, *Escherichia coli*, *Lactococcus lactis*, *Kineococcus radiotolerans*, *Desulfovibrio desulfuricans*, *Erwinia amylovora*, *Rhodospirillum rubrum*, *Vibrio furnissii*, *Stenotrophomonas maltophilia*, *Synechocystis* sp. PCC6803 y *Synechococcus elongatus*.

La biotina proteína ligasa (CE 6.3.4.15) es una enzima que cataliza la unión covalente de biotina a la subunidad de

40 proteína portadora de carboxilo de biotina (BCCP) de acetil-CoA carboxilasa. En algunas realizaciones de la presente invención, una biotina proteína ligasa se expresa o sobreexpresa en la célula hospedadora modificada por ingeniería genética. En determinadas realizaciones, la biotina proteína ligasa es *birA* de *Corynebacterium glutamicum* (YP_225000) o *bplI* de *Saccharomyces cerevisiae* (NP_010140).

La producción de ésteres de ácidos grasos tales como FAME o FAEE en una célula hospedadora puede facilitarse mediante la expresión o sobreexpresión de un éster sintasa (CE 2.3.1.75 o CE 3.1.1.67) en una célula hospedadora

45 modificada por ingeniería genética. En algunas realizaciones, la éster sintasa es ES9 de *Marinobacter hydrocarbonoclasticus* (SEQ ID NO: 2), ES8 de *Marinobacterhydrocar-bonoclasticus* (SEQ ID NO: 3), *AtfA1* de *Alcanivorax borkumensis* SK2 (SEQ ID NO: 4), *AtfA2* de *Alcanivorax borkumensis* SK2 (SEQ ID NO: 5), diacilglicerol

50 O-aciltransferasa de *Marinobacter aquaeolei* VT8 (SEQ ID NO: 6 o SEQ ID NO: 7), una cera sintasa o una cera éster sintasa bifuncional/acil-CoA:diacilglicerol aciltransferasa (*cera-dgaT*).

En determinadas realizaciones, un gen que codifica un polipéptido biosintético de aldehído graso se expresa o sobreexpresa en la célula hospedadora. Se desvelan polipéptidos biosintéticos de aldehídos grasos de ejemplo

55 adecuados para su uso en los métodos de la invención, por ejemplo, en la Publicación de Solicitud de Patente Internacional WO 2010/042664. En realizaciones preferidas, el polipéptido biosintético de aldehído graso tiene actividad de ácido carboxílico reductasa (CE 6.2.1.3 o CE 1.2.1.42), por ejemplo, actividad ácido graso reductasa.

En los métodos de la invención, el polipéptido que tiene actividad ácido carboxílico reductasa no está particularmente

60 limitado. Se desvelan polipéptidos de ejemplo que tienen actividad ácido carboxílico reductasa que son adecuados para su uso en los métodos de la presente invención, por ejemplo, en las Publicaciones de Solicitud de Patente Internacional WO 2010/062480 y WO 2010/042664. En algunas realizaciones, el polipéptido que tiene actividad ácido carboxílico reductasa es *CarB* de *M. smegmatis* (YP_889972) (SEQ ID NO: 8). En otras realizaciones, el polipéptido que tiene actividad ácido carboxílico reductasa es *CarA* [ABK75684] de *M. smegmatis*, *FadD9* [AAK46980] de *M. tuberculosis*, *CAR* [AAR91681] de *Nocardia* sp. NRRL 5646, *CAR* [YP-001070587] de *Mycobacterium* sp. JLS o *CAR* [YP-118225] de *Streptomyces griseus*. Las expresiones "ácido carboxílico reductasa", "CAR", y "polipéptido

65

biosintético de aldehído graso" se usan indistintamente en el presente documento.

En determinadas realizaciones, una tioesterasa y una ácido carboxílico reductasa se expresan o sobreexpresan en la célula hospedadora modificada por ingeniería genética.

En algunas realizaciones, un gen que codifica un polipéptido biosintético de alcohol graso se expresa o sobreexpresa en la célula hospedadora. Se desvelan polipéptidos biosintéticos de alcohol graso de ejemplo adecuados para su uso en los métodos de la invención, por ejemplo, en la Publicación de Solicitud de Patente Internacional WO 2010/062480. En determinadas realizaciones, el polipéptido biosintético de alcohol graso tiene actividad aldehído reductasa o alcohol deshidrogenasa (CE 1.1.1.1). Los ejemplos de polipéptidos biosintéticos de alcohol graso incluyen, pero sin limitación AlrA de *Acenitobacter* sp. M-1 (SEQ ID NO: 9) u homólogos de AlrA y alcohol deshidrogenasas de *E. coli* endógenas tales como YjgB, (AAC77226) (SEQ ID NO: 10), DkgA (NP_417485), DkgB (NP_417473), YdjL (AAC74846), YdjJ (NP_416288), AdhP (NP_415995), YhdH (NP_417719), YahK (NP_414859), YphC (AAC75598), YqhD (446856) y YbbO [AAC73595.1].

Como se usa en el presente documento, la expresión "alcohol deshidrogenasa" es un péptido capaz de catalizar la conversión de un aldehído graso en alcohol (por ejemplo, alcohol graso). Un experto habitual en la materia apreciará que determinadas alcohol deshidrogenasas son capaces de catalizar otras reacciones también. Por ejemplo, determinadas alcohol deshidrogenasas aceptarán otros sustratos además de los aldehídos grasos, y estas alcohol deshidrogenasas inespecíficas también se incluyen en la expresión "alcohol deshidrogenasa". Se desvelan alcohol deshidrogenasas de ejemplo adecuadas para su uso en los métodos de la invención, por ejemplo, en la Publicación de Solicitud de Patente Internacional WO 2010/062480.

En algunas realizaciones, una tioesterasa, una ácido carboxílico reductasa y una alcohol deshidrogenasa se expresan o sobreexpresan en la célula hospedadora modificada por ingeniería genética. En determinadas realizaciones, la tioesterasa es tesA (SEQ ID NO: 11), la ácido carboxílico reductasa es carB (SEQ ID NO: 8) y la alcohol deshidrogenasa es YjgB (SEQ ID NO: 10) o AlrAadp1 (SEQ ID NO: 9).

Las fosfopanteteína transferasas (PPTasas) (CE 2.7.8.7) catalizan la transferencia de 4'-fosfopanteteína de CoA a un sustrato. CAR de *Nocardia* y varios de sus homólogos contienen un supuesto sitio de fijación para la 4'-fosfo-panteteína (PPT) (He *et al.*, *Appl. Environ. Microbiol.*, 70(3): 1874-1881 (2004)). En algunas realizaciones de la invención, una PPTasa se expresa o sobreexpresa en una célula hospedadora modificada por ingeniería genética. En determinadas realizaciones, la PPTasa es EntD de *E. coli* MG1655 (SEQ ID NO: 12).

En algunas realizaciones, una tioesterasa, una ácido carboxílico reductasa, una PPTasa y una alcohol deshidrogenasa se expresan o sobreexpresan en la célula hospedadora modificada por ingeniería genética. En determinadas realizaciones, la tioesterasa es tesA (SEQ ID NO: 11), la ácido carboxílico reductasa es carB (SEQ ID NO: 8), la PPTasa es entD (SEQ ID NO: 12) y la alcohol deshidrogenasa es yjgB (SEQ ID NO: 10) o alrAadp1 (SEQ ID NO: 9).

La divulgación también proporciona un ácido graso o un derivado graso producidos mediante cualquiera de los métodos que se describen en el presente documento. Un ácido graso o derivado del mismo producido mediante cualquiera de los métodos que se describen en el presente documento puede usarse directamente como combustible, aditivos de combustible, materiales de partida para la producción de otros compuestos químicos (por ejemplo, polímeros, tensioactivos, plásticos, tejidos, disolventes, adhesivos, etc.) o aditivos para el cuidado personal. Estos compuestos también pueden usarse como materia prima para reacciones posteriores, por ejemplo, hidrogenación, craqueo catalítico (por ejemplo, a través de hidrogenación, pirolisis o ambos), para preparar otros productos.

En algunos aspectos, la divulgación proporciona una composición de biocombustible que comprende el ácido graso o derivado del mismo producido mediante los métodos que se describen en el presente documento. Como se usa en el presente documento, el término "biocombustible" se refiere a cualquier combustible derivado de biomasa. Los biocombustibles pueden sustituir los combustibles a base de petróleo. Por ejemplo, los biocombustibles incluyen combustibles para el transporte (por ejemplo, gasolina, diésel, combustible para reactores, etc.), combustibles de calefacción y combustibles de generación de electricidad. Los biocombustibles son una fuente de energía renovable. Como se usa en el presente documento, el término "biodiésel" significa un biocombustible que puede ser un sustituto del diésel, que deriva del petróleo. El biodiésel puede usarse en motores diésel de combustión interna, ya sea en forma pura, que se denomina biodiésel "puro", o como una mezcla en cualquier concentración con diésel a base de petróleo. El biodiésel puede incluir ésteres o hidrocarburos, tales como alcoholes. En determinados aspectos, el biocombustible se selecciona entre el grupo que consiste en un biodiésel, un alcohol graso, un éster graso, un triacilglicérido, una gasolina o un combustible para reactores.

Para potenciar el rendimiento de un combustible o motor, se utilizan aditivos de combustible. Por ejemplo, pueden utilizarse aditivos de combustible para alterar el punto de congelación/gelificación, el punto de turbidez, la lubricidad, la viscosidad, la estabilidad oxidativa, la calidad de ignición, el nivel de octanaje y/o el punto de inflamación de un combustible. En Estados Unidos, todos los aditivos de combustibles deben estar registrados en la Agencia de Protección Ambiental (EPA). Los nombres de los aditivos de combustibles y las empresas que venden los aditivos de combustibles están a disposición del público poniéndose en contacto con la EPA o visitando el sitio web de la EPA.

Un experto habitual en la materia apreciará que un biocombustible producido de acuerdo con los métodos que se describen en el presente documento puede mezclarse con uno o más aditivos de combustible para transmitir una calidad deseada.

5 La divulgación también proporciona una composición de tensioactivo o una composición de detergente que comprende un alcohol graso producido mediante cualquiera de los métodos que se describen en el presente documento. Un experto habitual en la materia apreciará que, dependiendo del fin previsto de la composición de tensioactivo o detergente, pueden producirse y usarse diferentes alcoholes grasos. Por ejemplo, cuando los alcoholes grasos que se describen en el presente documento se usan como materia prima para la producción de tensioactivos o detergentes, un experto habitual en la materia apreciará que las características de la materia prima de alcohol graso afectarán a las características de la composición de tensioactivo o detergente producida. Por tanto, las características de la composición de tensioactivo o detergente pueden seleccionarse produciendo alcoholes grasos particulares para su uso como materia prima.

15 Una composición de tensioactivo y/o detergente a base de alcohol graso que se describe en el presente documento puede mezclarse con otros tensioactivos y/o detergentes bien conocidos en la técnica. En algunos aspectos, la mezcla puede incluir al menos aproximadamente el 10 %, al menos aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, al menos aproximadamente el 60 % o un intervalo limitado por cualquiera de los dos valores anteriores, en peso del alcohol graso. En otros ejemplos, puede prepararse una composición de tensioactivo o detergente que incluya al menos aproximadamente el 5 %, al menos aproximadamente el 10 %, al menos aproximadamente el 20 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 % o un intervalo limitado por cualquiera de los dos valores anteriores, en peso de un alcohol graso que incluye una cadena de carbono que tiene 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 o 22 carbonos de longitud. Dichas composiciones de tensioactivo o detergente también pueden incluir al menos un aditivo, tal como una microemulsión o un tensioactivo o detergente de fuentes no microbianas tales como aceites vegetales o petróleo, que puede estar presente en una cantidad de al menos aproximadamente el 5 %, al menos aproximadamente el 10 %, al menos aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 % o un intervalo limitado por cualquiera de los dos valores anteriores, en peso del alcohol graso.

35 Los bioproductos (por ejemplo, ácidos grasos, acil-CoA, hidrocarburos, aldehídos grasos, alcoholes grasos, ésteres grasos, composiciones de tensioactivo y composiciones de biocombustible) producidos de acuerdo con los métodos de la invención pueden distinguirse de los compuestos orgánicos derivados del carbono petroquímico basándose en la huella isotópica del carbono doble o la datación por ^{14}C . Adicionalmente, la fuente específica de carbono de origen biológico (por ejemplo, glucosa frente a glicerol) puede determinarse mediante la huella isotópica del carbono doble (véase, por ejemplo, la Patente de los EE.UU. 7.169.588).

45 La capacidad de distinguir los bioproductos de los compuestos orgánicos a base de petróleo es beneficiosa para rastrear estos materiales en el comercio. Por ejemplo, los compuestos orgánicos o productos químicos que comprenden los perfiles de isótopos de carbono tanto biológicos como derivados del petróleo, pueden distinguirse de los compuestos orgánicos y de los productos químicos hechos solo de materiales a base de petróleo. Por tanto, los materiales preparados de acuerdo con los métodos inventivos pueden seguirse en el comercio basándose en su perfil único de isótopos de carbono.

50 Los bioproductos pueden distinguirse de los compuestos orgánicos a base de petróleo mediante la comparación de la relación de isótopos de carbono estables ($^{13}\text{C}/^{12}\text{C}$) en cada combustible. La relación $^{13}\text{C}/^{12}\text{C}$ en un bioproducto dado es una consecuencia de la relación $^{13}\text{C}/^{12}\text{C}$ en el dióxido de carbono atmosférico en el momento en que se fija el dióxido de carbono. También refleja la vía metabólica exacta. También se producen variaciones regionales. El petróleo, las plantas C_3 (las de hoja ancha), las plantas C_4 (las gramíneas) y los carbonatos marinos muestran diferencias significativas en $^{13}\text{C}/^{12}\text{C}$ y los valores de $\delta^{13}\text{C}$ correspondientes. Además, la materia lipídica de plantas C_3 y C_4 se analiza de manera diferente a la de los materiales derivados de los componentes de hidratos de carbono de las mismas plantas como consecuencia de la vía metabólica.

60 La escala de medición de ^{13}C se definió originalmente por un ajuste a cero con piedra caliza belemnita Pee Dee (PDB, *Pee Dee Belemnite*), en donde los valores se dan en partes por mil desviaciones con respecto a este material. Los valores de " $\delta^{13}\text{C}$ " se expresan en partes por mil (por mil), abreviado, ‰ y se calculan como se indica a continuación:

$$\delta^{13}\text{C} (\text{‰}) = [(^{13}\text{C}/^{12}\text{C})_{\text{muestra}} - (^{13}\text{C}/^{12}\text{C})_{\text{patrón}}] / (^{13}\text{C}/^{12}\text{C})_{\text{patrón}} \times 1000$$

65 En algunos aspectos, un bioproducto producido de acuerdo con los métodos de la invención tiene un $\delta^{13}\text{C}$ de aproximadamente -30 o más, aproximadamente -28 o más, aproximadamente -27 o más, aproximadamente -20 o

más, aproximadamente -18 o más, aproximadamente -15 o más, aproximadamente -13 o más o aproximadamente -10 o más. Como alternativa, o además, un bioproducto tiene un $\delta^{13}\text{C}$ de aproximadamente -4 o menos, aproximadamente -5 o menos, aproximadamente -8 o menos, aproximadamente -10 o menos, aproximadamente -13 o menos, aproximadamente -15 o menos, aproximadamente -18 o menos o aproximadamente -20 o menos. Por tanto, el bioproducto puede tener un $\delta^{13}\text{C}$ limitado por dos cualesquiera de los criterios de valoración anteriores. Por ejemplo, el bioproducto puede tener un $\delta^{13}\text{C}$ de aproximadamente -30 a aproximadamente -15, de aproximadamente -27 a aproximadamente -19, de aproximadamente -25 a aproximadamente -21, de aproximadamente -15 a aproximadamente -5, de aproximadamente -13 a aproximadamente -7, o de aproximadamente -13 a aproximadamente -10. En algunas realizaciones, el bioproducto puede tener un $\delta^{13}\text{C}$ de aproximadamente -10, -11, -12 o -12,3. En otros aspectos, el bioproducto tiene un $\delta^{13}\text{C}$ de aproximadamente -15,4 o más. En otros aspectos más, el bioproducto tiene un $\delta^{13}\text{C}$ de aproximadamente -15,4 a aproximadamente -10,9, o un $\delta^{13}\text{C}$ de aproximadamente -13,92 a aproximadamente -13,84.

Los bioproductos también pueden distinguirse de los compuestos orgánicos a base de petróleo comparando la cantidad de ^{14}C en cada compuesto. Porque ^{14}C tiene una vida media nuclear de 5730 años, los combustibles a base de petróleo que contienen carbono "más antiguo" pueden distinguirse de los bioproductos que contienen carbono "más nuevo" (véase, por ejemplo, Currie, "Source Apportionment of Atmospheric Particles", *Characterization of Environmental Particles*, J. Buffle y H. P. van Leeuwen, Eds., Vol. I de la Serie de Química Analítica Ambiental de la IUPAC, Editorial Lewis, Inc., págs. 3-74 (1992)).

El ^{14}C puede medirse mediante espectrometría de masas con acelerador (EMA), con resultados proporcionados en unidades de "fracción de carbono moderno" (f_m). La f_m se define por los Materiales de Referencia Patrón (MRP) 4990B y 4990C del Instituto Nacional de Estándares y Tecnología (NIST, *National Institute of Standards and Technology*). Como se usa en el presente documento, la "fracción de carbono moderno" o f_m tiene el mismo significado definido por los materiales de referencia estándar (SRM) 4990B y 4990C del Instituto Nacional de Estándares y Tecnología (NIST, *National Institute of Standards and Technology*), conocidos como patrones de ácidos oxálicos HOxI y HOxII, respectivamente. La definición fundamental se refiere a 0,95 veces la relación de isótopos de $^{14}\text{C}/^{12}\text{C}$ de HOxI (denominada AD 1950). Esto es más o menos equivalente a la madera de antes de la revolución industrial corregida para la desintegración. Para la biosfera viva actual (materia vegetal), f_m es aproximadamente 1,1.

En algunos aspectos, un bioproducto producido de acuerdo con los métodos de la invención tiene una $f_m^{14}\text{C}$ de al menos aproximadamente 1, por ejemplo, al menos aproximadamente 1,003, al menos aproximadamente 1,01, al menos aproximadamente 1,04, al menos aproximadamente 1,111, al menos aproximadamente 1,18 o al menos aproximadamente 1,124. Como alternativa, o además, el bioproducto tiene una $f_m^{14}\text{C}$ de aproximadamente 1,130 o menos, por ejemplo, aproximadamente 1,124 o menos, aproximadamente 1,18 o menos, aproximadamente 1,111 o menos o aproximadamente 1,04 o menos. Por tanto, el bioproducto puede tener una $f_m^{14}\text{C}$ limitada por dos cualesquiera de los criterios de valoración anteriores. Por ejemplo, el bioproducto puede tener una $f_m^{14}\text{C}$ de aproximadamente 1,003 a 1,124, una $f_m^{14}\text{C}$ de aproximadamente 1,04 a aproximadamente 1,18 o una $f_m^{14}\text{C}$ de aproximadamente 1,111 a aproximadamente 1,124.

Otra medida de ^{14}C se conoce como el porcentaje de carbono moderno, es decir, pMC. Para un arqueólogo o geólogo que usa dataciones con ^{14}C , AD 1950 es igual a "cero años". Esto también representa un pMC de 100. La "bomba de carbono" en la atmósfera alcanzó casi el doble del nivel normal en 1963 en el pico de ensayos de armas termónucleares. Su distribución dentro de la atmósfera se ha aproximado desde su aparición, mostrando valores superiores a pMC de 100 para plantas y animales que viven desde AD 1950. Ha disminuido gradualmente con el tiempo y el valor actual está cerca de un pMC de 107,5. Esto significa que un material de biomasa nuevo, tal como de maíz, proporcionaría un distintivo de ^{14}C de pMC cercano a 107,5. Los compuestos a base de petróleo tendrán un valor de pMC de cero. La combinación de carbono fósil con carbono actual dará lugar a una dilución del contenido actual de pMC. Suponiendo que un pMC de 107,5 representa el contenido de ^{14}C de los materiales de biomasa actuales y un pMC de 0 representa el contenido de ^{14}C de los productos derivados del petróleo, el valor medido de pMC para ese material reflejará las proporciones de los dos tipos de componentes. Por ejemplo, un material derivado al 100 % de la soja actual tendría un distintivo de radiocarbono de un valor de pMC cercano a 107,5. Si ese material se diluyera al 50 % con productos a base de petróleo, la mezcla resultante tendría un distintivo de radiocarbono de un pMC de aproximadamente 54.

Se deriva un contenido de carbono con base biológica asignando el "100 %" igual a 107,5 pMC y el "0 %" igual a 0 pMC. Por ejemplo, una muestra que mide un pMC de 99 proporcionará un contenido de carbono con base biológica equivalente del 93 %. Este valor se conoce como el resultado medio de carbono con base biológica, y supone que todos los componentes de dentro del material analizado se originaron a partir del material biológico actual o del material a base de petróleo.

En algunos aspectos, un bioproducto producido de acuerdo con los métodos de la invención tiene un pMC de al menos aproximadamente 50, al menos aproximadamente 60, al menos aproximadamente 70, al menos aproximadamente 75, al menos aproximadamente 80, al menos aproximadamente 85, al menos aproximadamente 90, al menos aproximadamente 95, al menos aproximadamente 96, al menos aproximadamente 97 o al menos aproximadamente 98. Como alternativa, o además, el bioproducto tiene un pMC de aproximadamente 100 o menos, aproximadamente

99 o menos, aproximadamente 98 o menos, aproximadamente 96 o menos, aproximadamente 95 o menos, aproximadamente 90 o menos, aproximadamente 85 o menos o aproximadamente 80 o menos. Por tanto, el bioproducto puede tener un pMC limitado por dos cualesquiera de los criterios de valoración anteriores. Por ejemplo, un bioproducto puede tener un pMC de aproximadamente 50 a aproximadamente 100; de aproximadamente 60 a aproximadamente 100; de aproximadamente 70 a aproximadamente 100; de aproximadamente 80 a aproximadamente 100; de aproximadamente 85 a aproximadamente 100; de aproximadamente 87 a aproximadamente 98; o de aproximadamente 90 a aproximadamente 95. En otros aspectos, un bioproducto que se describe en el presente documento tiene un pMC de aproximadamente 90, aproximadamente 91, aproximadamente 92, aproximadamente 93, aproximadamente 94 o aproximadamente 94,2.

Los siguientes ejemplos ilustran adicionalmente la invención.

EJEMPLO 1

Este ejemplo demuestra un método para identificar células hospedadoras modificadas por ingeniería genética que muestran una producción potenciada de ácidos grasos y derivados de los mismos.

ALC310 es una cepa de *E. coli* caracterizada anteriormente que tiene el genotipo MG1655 (Δ fadE::FRT Δ fhuA::FRT fabB[A329V] Δ entD::P_{T5}-entD) que porta el plásmido ALC310 (pCL1920_P_{TRC}_carBopt_13G04_alrA_sthA) (SEQ ID NO: 13) y produce ácidos grasos y derivados de los mismos. Para identificar las cepas que presentan un título o un rendimiento mejorados de los ácidos grasos o derivados de los mismos, se realizó la mutagénesis por transposón de ALC310, seguida de un cribado de alto rendimiento.

El ADN de transposón se preparó mediante la clonación de un fragmento de ADN en el plásmido EZ-Tn5™ pMOD™<R6K ori/MCS> (Epicentre Biotechnologies, Madison, WI). El fragmento de ADN contiene un promotor T5 y un gen de resistencia al cloranfenicol (cat) flanqueado por sitios loxP. El plásmido resultante se denominó p100.38 (SEQ ID NO: 14). El plásmido p100.38 se digirió opcionalmente con la enzima de restricción PshAI, se incubó con la enzima Transposasa EZ-Tn5™ (Epicentre Biotechnologies, Madison, WI) y se electroporó en células ALC310 electrocompetentes según las instrucciones del fabricante. Las colonias resultantes contenían el ADN de transposón insertado aleatoriamente en el cromosoma de ALC310.

Después, los clones de transposón se sometieron a un cribado de alto rendimiento para medir la producción de alcoholes grasos. En resumen, las colonias se cultivaron picando en placas de pocillos profundos que contenían medio Luria-Bertani (LB). Después del crecimiento nocturno, cada cultivo se inoculó en LB fresco. Después de 3 horas de crecimiento, cada cultivo se inoculó en medio FA-2 fresco. Se incluyó espectinomicina (100 µg/ml) en todos los medios para mantener la selección del plásmido 7P36. El medio FA-2 es medio M9 con glucosa al 3 % complementado con antibióticos, citrato de hierro 10 µg/l, tiamina 1 µg/l, tampón Bis-Tris 0,1 M (pH 7,0) y una dilución 1:1000 de la solución de minerales traza descrita en la Tabla 1.

TABLA 1

Solución de minerales traza (filtro esterilizado)

ZnCl₂ • 4H₂O 2 g/l
CaCl₂ • 6H₂O 2 g/l
Na₂MoO₄ • 2H₂O 2 g/l
CuSO₄ • 5H₂O 1,9 g/l
H₃BO₃ 0,5 g/l
HCl concentrado 100 ml/l
c.s. agua Milli-Q

Después de 20 horas de crecimiento en FA-2, los cultivos se extrajeron con acetato de butilo. El extracto en bruto se derivatizó con BSTFA (N,O-bis[trimetilsilil]trifluoroacetamida) y las especies grasas totales (por ejemplo, alcoholes grasos, aldehídos grasos y ácidos grasos) se midieron con CG-DIL como se describe en la Publicación de Solicitud de Patente Internacional WO 2008/119082.

Los clones que produjeron un 15 % más de especies grasas totales que ALC310 se sometieron a una verificación adicional usando una fermentación en matraz de agitación. En resumen, los clones se cultivaron en 2 ml de medio LB complementado con espectinomicina (100 mg/l) a 37 °C. Después del crecimiento nocturno, se transfirieron 100 µl de cultivo a 2 ml de LB fresco complementado con antibióticos. Después de 3 horas de crecimiento, se transfirieron 2 ml de cultivo a un matraz de 125 ml que contenía 18 ml de medio FA-2 complementado con espectinomicina (100 mg/l). Cuando la DO₆₀₀ del cultivo alcanzó 2,5, se añadió 1 mM de IPTG a cada matraz. Después de 20 horas de crecimiento a 37 °C, se extrajo una muestra de 400 µl de cada matraz y las especies grasas totales se extrajeron con 400 µl de acetato de butilo. Los extractos en bruto se analizaron directamente con CG-DIL como se ha descrito anteriormente.

Se identificó un clon de transposón (denominado D288) que presentó títulos aumentados de las especies grasas

totales en comparación con la cepa ALC310 parental (FIG. 2).

Los resultados de este ejemplo demuestran un método para identificar células hospedadoras modificadas por ingeniería genética que presentan una producción potenciada de ácidos grasos y derivados de los mismos en comparación con una célula hospedadora de tipo silvestre correspondiente.

EJEMPLO 2

Este ejemplo demuestra que una célula hospedadora modificada por ingeniería genética con una inserción de transposón en el gen *nhaB* presenta una producción potenciada de ácidos grasos y derivados de los mismos en comparación con la célula hospedadora de tipo silvestre correspondiente.

Se realizó un análisis de secuencia para identificar la ubicación de la inserción del transposón en la cepa D288 identificada en el Ejemplo 1. Para ello, se purificó ADN genómico de un cultivo de 3 ml de LB durante una noche de células D288 usando el kit ZR de ADN fúngico/bacteriano MiniPrep™ (Zymo Research) de acuerdo con las instrucciones del fabricante. El ADN genómico purificado se secuenció hacia afuera del transposón usando los cebadores DG150 (GCAGTTATTGGTGGCCTTAAACGCCTGTGTTGCTACGCCTG) (SEQ ID NO: 15) y DG153 (CCCAGGCTTCCCGGTATCAACAGGACACCAGG) (SEQ ID NO: 16), internos al transposón.

Se determinó que la cepa D288 tenía una inserción de transposón en el gen *nhaB* (FIG. 3).

Los resultados de este ejemplo demuestran que una célula hospedadora de *E. coli* modificada por ingeniería genética con una inserción de transposón en el gen *nhaB* presenta una producción potenciada de ácidos grasos y derivados de los mismos en comparación con una célula hospedadora de *E. coli* de tipo silvestre correspondiente.

EJEMPLO 3

Este ejemplo demuestra que las células hospedadoras modificadas por ingeniería genética que tienen un nivel alterado de producción de FadR presentan una producción potenciada de ácidos grasos y derivados de los mismos.

El gen *nhaB* está próximo al gen que codifica el regulador de la degradación de ácidos grasos, FadR (FIG. 3). Para determinar si la alteración de la expresión de FadR afecta a la producción de ácidos grasos o derivados de los mismos en las células hospedadoras, se clonó y se cribó una biblioteca de expresión de FadR.

Para clonar la biblioteca de expresión, el gen *fadR* de tipo silvestre se amplificó a partir de ADN genómico de *E. coli* MG1655 por PCR usando los cebadores DG191 (SEQ ID NO: 17) y DG192 (SEQ ID NO: 18). Un gen *fadR* mutante que contenía el cambio de aminoácidos S219N también se amplificó a partir de ADN genómico de *E. coli* MG1655 *fadR*[S219N] usando el conjunto de cebadores DG191 y DG192. Los cebadores utilizados en este ejemplo se enumeran en la Tabla 2.

TABLA 2

Cebador	Secuencia	Identificador de secuencia
DG191	ATGGTCATTAAGGCGCAAAGCCCCGG	SEQ ID NO: 17
DG192	GAGACCCACACTACCATCCTCGAGTTATCGCCCCTGA ATGGCTAAATCACCC	SEQ ID NO: 18
SL03	CTCGAGGATGGTAGTGTGGGGTCTCCC	SEQ ID NO: 19
SL23	GAGACCGTTTCTCGAATTAAATATGATACGCTCGAGCT TCGTCTGTTTCTACTGGTATTGGCACAAAC	SEQ ID NO: 20
DG193	TGAAAGATTAAATTTNHHARNDDHDDNWAGGAGNNNN NNNATGGTCATTAAGGCGCAAAGCCCCGG	SEQ ID NO: 21

Se amplificó por PCR un casete génico que codificaba la resistencia a la kanamicina (kan) a partir del plásmido pKD13 usando los cebadores SL03 (SEQ ID NO: 19) y SL23 (SEQ ID NO: 20). Cada casete de *fadR* (es decir, el tipo silvestre y el mutante S219N) se unió por separado con el casete de resistencia a kanamicina usando corte y empalme mediante PCR de extensión de solapamiento (SOE) usando los cebadores SL23 y DG193 (SEQ ID NO: 21). El cebador DG193 contenía nucleótidos degenerados para la generación de variantes de expresión.

El plásmido p100.487 (pCL1920_P_{TRC}_carBopt_13G04_alrA_fabB[A329G]) (SEQ ID NO: 22) se linealizó a través de digestión por restricción con enzimas *Swa*I y *Xho*I. Cada uno de los dos productos *fadR*-kan de PCR SOE se clonó por separado en plásmido linealizado p100.487 usando el sistema INFUSION™ (Clontech, Mountain View, CA) y después los plásmidos se transformaron en células químicamente competentes NEB TURBO™ (New England Biolabs,

Ipswich, MA). Los transformantes se cultivaron en placa en agar LB que contenía kanamicina 50 µg/ml.

Se obtuvieron miles de colonias para *fadR* y *fadR* [S219N]. Las colonias se rasparon de las placas y los plásmidos se aislaron mediante minipreparación de acuerdo con protocolos convencionales. El conjunto de plásmidos resultante se transformó en una cepa de *E. coli* EG149 con un genotipo de MG1655 (Δ fadE::FRT Δ fhuA::FRT fabB[A329V] Δ entD::P_{T5}-entD)) y se seleccionó en placas de LB que contenían espectinomicina 100 µg/ml.

Después, los transformantes se cribaron para la producción de especies grasas totales (por ejemplo, ácidos grasos, aldehídos grasos y alcoholes grasos) usando el procedimiento en pocillo profundo descrito en el Ejemplo 1. Se identificaron numerosas cepas que presentaron una producción potenciada de especies grasas totales en comparación con la cepa ALC487 de control (cepa EG149 que porta plásmido p100.487) (FIG. 4). Las cepas que expresan *fadR* de tipo silvestre o *fadR* [S219N] presentaron una producción potenciada de especies grasas totales en comparación con la cepa ALC487, aunque los títulos más altos se observaron en cepas que expresaban FadR de tipo silvestre (FIG. 4).

A varias de las principales cepas productoras que expresaban FadR de tipo silvestre identificadas en el cribado inicial se les asignaron ID de cepa y se validaron en una fermentación en matraz de agitación. En resumen, cada cepa se sembró en estrías para determinar colonias individuales y se cultivaron tres colonias separadas de cada cepa en tres matraces separados de acuerdo con el protocolo de fermentación en matraces de agitación descrito en el Ejemplo 1. Las especies grasas totales se midieron usando CG-DIL como se describe en el Ejemplo 1. Todas las cepas que expresaban FadR de tipo silvestre presentaron títulos de especies grasas totales más altos en comparación con la cepa ALC487 de control (FIG. 5).

Varias de las principales cepas productoras que expresaban FadR de tipo silvestre después se caracterizaron con el fin de determinar el rendimiento de especies grasas. Para ello, se realizó una fermentación en matraz de agitación como se ha descrito anteriormente, excepto porque (i) la temperatura se mantuvo a 32 °C, (ii) se añadió glucosa adicional después de 18 y 43 horas y (iii) la extracción se realizó a las 68,5 horas. Las especies grasas totales producidas se dividió por la glucosa total consumida para calcular el rendimiento de especies grasas. Todas las cepas que expresaban FadR de tipo silvestre presentaron un mayor rendimiento de especies grasas totales en comparación con la cepa ALC487 de control (FIG. 6).

La cepa D512 se caracterizó adicionalmente evaluando el título de especies grasas totales y el rendimiento después de la fermentación en un biorreactor de 5 l. A una tasa de alimentación de glucosa de 10 g/l/h, la cepa D512 produjo títulos más altos de ácidos grasos y alcoholes grasos en comparación con la cepa ALC487 de control (FIG. 7). Además, el rendimiento total de todas las especies grasas aumentó en la cepa D512 en comparación con la cepa ALC487 (FIG. 7). A una tasa de alimentación de glucosa más alta de 15 g/l/h, la cepa D512 produjo aproximadamente 68,5 g/l de especies grasas totales con un rendimiento de aproximadamente el 20 % (FIG. 7). La cepa D512 produjo un título de especies grasas totales más alto y un rendimiento de 15 g/l/h en comparación con 10 g/l/h (FIG. 7).

Se aisló ADN plasmídico de la cepa D512 y se secuenció de acuerdo con protocolos convencionales. Se determinó que el plásmido obtenido de la cepa D512, denominado pDG109, tenía la secuencia correspondiente a la SEQ ID NO: 23.

Los resultados de este ejemplo demuestran que las células hospedadoras modificadas por ingeniería genética que tienen un nivel de expresión alterado de FadR producen títulos y rendimientos más altos de ácidos grasos y derivados de los mismos en comparación con las células hospedadoras de tipo silvestre correspondientes.

EJEMPLO 4

Este ejemplo demuestra un método para producir títulos altos de ácidos grasos en células hospedadoras modificadas por ingeniería genética que tienen un nivel alterado de expresión de FadR.

La cepa EG149 de *E. coli* utilizada en el Ejemplo 3 sobreexpresa el gen de entD, que codifica una fosfopanteteína transferasa (PPTasa) implicada en la activación de la enzima CarB que cataliza la reducción de los ácidos grasos a aldehídos y alcoholes grasos.

Para evaluar el efecto de la expresión de entD en la producción de ácido graso y alcohol graso en la cepa D512, se generó una variante D512 que contenía una supresión del gen de entD (D512 Δ entD). Las fermentaciones en matraz de agitación se realizaron con la cepa D512 y la cepa D512 Δ entD como se describe en el Ejemplo 1. La cepa D512 produjo títulos altos de alcoholes grasos y comparativamente títulos y rendimientos menores de ácidos grasos (FIG. 7). Por el contrario, la cepa D512 Δ entD produjo títulos y rendimientos altos de ácidos grasos y títulos relativamente bajos de alcoholes grasos (FIG. 8). Los títulos de especies grasas totales fueron similares entre la cepa D512 y la cepa D512 Δ entD (FIG. 8).

Los resultados de este ejemplo demuestran que las células hospedadoras modificadas por ingeniería genética que

tienen un nivel alterado de expresión de FadR producen títulos altos de ácidos grasos cuando se suprime el gen de *entD*.

EJEMPLO 5

Este ejemplo demuestra un método para identificar células hospedadoras modificadas por ingeniería genética que muestran una producción potenciada de ácidos grasos y derivados de los mismos.

Para evaluar adicionalmente el efecto de la expresión alterada de FadR en la producción de ácidos grasos y derivados de los mismos, se prepararon bibliotecas de sitios de unión de ribosomas (RBS) de FadR (S219N) y FadR de tipo silvestre y se cribaron en células hospedadora de *E. coli*.

Una biblioteca de RBS se insertó en dirección 5' del gen *fadR* (S219N) en pDS57 como se indica a continuación. El ADN genómico de una cepa que contenía el alelo de *fadR*(S219N) Moniker stEP005; id: s26z7 se amplificó mediante PCR usando el conjunto de cebadores DG191 (SEQ ID NO: 17) y *fadR* (S219N)_pme319rc (SEQ ID NO: 24). Los cebadores utilizados en este ejemplo se enumeran en la Tabla 3.

TABLA 3

Cebador	Secuencia	Identificador de secuencia
DG191	ATGGTCATTAAGGCGCAAAGCCCGG	SEQ ID NO: 17
<i>fadR</i> (S219N)_pme319rc	CAAAACAGCCAAGCTGGAGACCGTTTTTATCGCC CCTGAATGGCTAAATCACC	SEQ ID NO: 24
377-rbs- <i>fadR</i> (S219N)f	GCCCGAACCCGCAAGTAANHARNDHDDNWAG GARNNNNNNATGGTCATTAAGGCGCAAAGCCC GG	SEQ ID NO: 25
NH246	AAAAACGGTCTCCAGCTTGGCTGTTTTGGCGGAT GAGAGAAGATTTTC	SEQ ID NO: 26
377-3r	TTACTTGCGGGTTCGGGCGC	SEQ ID NO: 27

Después de que se preparase el molde de *fadR* (S219N), el RBS se añadió mediante PCR usando el conjunto de cebadores 377-rbs-*fadR* (S219N)f (SEQ ID NO: 25) y *fadR* (S219N)-pme319rc (SEQ ID NO: 24). El cebador 377-rbs-*fadR* (S219N)f contenía nucleótidos degenerados para introducir variabilidad en la biblioteca de RBS. El RBS- *fadR* (S219N) se ligó con una cadena principal de vector pDS57 (descrito en el Ejemplo 5), usando el kit CLONEZ™ disponible en el mercado de Genscript (Piscataway, NJ) con el conjunto de cebadores NH246 (SEQ ID NO: 26) y 377-3r (SEQ ID NO: 27).

Una biblioteca RBS también se insertó en dirección 5' del gen *fadR* de tipo silvestre en pDS57 usando un protocolo similar, excepto porque el gen *fadR* de tipo silvestre se amplificó mediante PCR usando ADN genómico de DV2 de *E. coli*.

Las construcciones ligadas de pDS57-rbs- *fadR* (S219N) y pDS57-rbs- *fadR* se transformaron por separado en una cepa DAM1 de *E. coli* por electroporación. La cepa DAM1 se produjo como un derivado de la cepa DV2 (MG1655 Δ fadE, Δ fhuA), donde los genes *l* *acl*^q-P_{TTC}-tesA-fadD se integraron en el cromosoma usando el sistema de entrega basado en Tn7 presente en el plásmido pGRG25 (descrito en McKenzie *et al.*, *BMC Microbiology* 6: 39 (2006)). Tras la transformación, las células se recuperaron durante 1 hora a 37 °C, seguido de cultivo en placa agar LB que contenía espectinomicina. Después de la incubación durante una noche a 37 °C, se escogieron colonias individuales para el cribado en 96 placas de pocillos profundos que contenían 300 μ l/pocillo de LB con espectinomicina. Las placas se incubaron en un agitador de 32 °C con 80 % de humedad y se agitaron a 250 RPM durante aproximadamente 5 horas. Después de 5 horas de crecimiento, se transfirieron 30 μ l/pocillo de cultivo de LB a medio FA2 300 μ l/pocillo (2 g/l de nitrógeno) que contenía espectinomicina. Las placas se incubaron de nuevo en un agitador de 32 °C con 80 % de humedad y se agitaron a 250 RPM durante una noche. Se inocularon 30 μ l/pocillo del cultivo durante la noche en medio FA2 300 μ l/pocillo (1 g/l de nitrógeno) que contenía espectinomicina, IPTG 1mM y metanol al 2 %. Una placa de réplica se incubó en un agitador a 32 °C y otra se incubó en un agitador a 37 °C con 80 % de humedad y agitación a 250 RPM durante una noche. La receta para el medio FA2 se enumera en la Tabla 4.

TABLA 4

Reactivo	ml de Reactivo por 1000 ml de FA2
Solución salina 5X	200
Tiamina (10 mg/ml)	0,1
MgSO ₄ 1M	1
CaCl ₂ 1M	0,1
Glucosa al 50 %	60
TM2 (minerales traza sin hierro)	1
Citrato férrico 10 g/l	1
Tampón Bis-Tris 2M	50
NH ₄ Cl*	10
Agua	c.s. para 1000 ml
* eliminado para medio FA2 (nitrógeno 1 g/l)	

Después de aproximadamente 24 horas de incubación, las placas se extrajeron añadiendo HCl 1M 40 µl/pocillo y acetato de butilo 300 µl/pocillo. Las placas se agitaron durante 15 minutos a 2000 RPM y después se centrifugaron durante 10 minutos a 4500 RPM a temperatura ambiente. Se transfirieron 50 µl de la capa orgánica por pocillo a una placa de 96 pocillos poco profunda que contenía BSTFA 50 µl/pocillo (Sigma Aldrich, St. Louis, MO) y los extractos se analizaron mediante CG-DIL.

Se identificaron varios clones de DAM1 de *E. coli* transformados con pDS57-rbs- *fadR* como que producían títulos sustancialmente más altos de FAME y ácidos grasos libres en comparación con DAM1 de *E. coli* de control transformado con pDS57 solo. En general, las variantes de *FadR* produjeron relaciones de C14 a C16 bajas y presentaron títulos generales más altos a 32 °C en comparación con 37 °C.

También se identificaron numerosos clones de DAM1 de *E. coli* transformados con pDS57-rbs-*FadR*(S219N) que como que producían títulos sustancialmente más altos de FAME y ácidos grasos libres en comparación con el DV2 de *E. coli* de control o DAM1 de *E. coli* transformado con pDS57 solo.

Dos de los clones transformados con pDS57-rbs-*FadR*(S219N) identificados en el cribado inicial (designados como P1A4 y P1G7) se caracterizaron adicionalmente en fermentaciones en matraz de agitación. En resumen, cada colonia fue inoculada en 5 ml de LB que contenía espectinomicina y se incubaron a 37 °C con agitación a aproximadamente 200 RPM durante aproximadamente 5 horas. Se transfirieron 1,5 ml del cultivo de LB a 13,5 ml de medio FA2 (2 g/l de nitrógeno) que contenía Tritón X-100 al 0,05 % y espectinomicina en un matraz con deflectores de 125 ml. Los cultivos en matraz se incubaron durante una noche a 32 °C, 80 % de humedad y 250 RPM. Se transfirieron 1,5 ml del cultivo de una noche a un nuevo matraz con deflectores de 125 ml que contenía 13,5 ml de medio FA2 (1 g/l de nitrógeno) que contenía Tritón X-100 al 0,05 %, IPTG 1mM, metanol al 2 % y espectinomicina. Después, los cultivos en matraz se incubaron a 32 °C, 80 % de humedad y 250 RPM. Después de 56 horas de incubación, se tomaron muestras de 500 µl de cada matraz. Se diluyeron 100 µl de cada muestra con 900 µl de agua para medir la DO del cultivo, se diluyeron 100 µl de cada muestra con 900 µl de agua para medir la glucosa restante y se extrajeron 300 µl de cada muestra y se analizaron usando CG-DIL como se ha descrito anteriormente.

Ambas variantes de *FadR* (es decir, P1A4 y P1G7) produjeron títulos y rendimientos más altos de especies grasas totales en comparación con la cepa de control transformada con pDS57 solo en la fermentación en matraz de agitación (FIG. 9).

La producción de especies grasas por P1A4 y P1G7 también se midió en fermentaciones a gran escala. Para ello, se cultivaron células de una solución madre congelada en medio LB durante unas horas y después se transfirieron a un medio definido que consistía en KH₂PO₄ 3 g/l, dihidrato de Na₂HPO₄ 6 g/l, NH₄Cl 2 g/l, MgSO₄ x 7 H₂O 0,24 g/l, glucosa 20 g/l, tampón Bis-Tris 200 mM (pH 7,2), solución de metales traza 1,0 ml/l y tiamina 1,0 mg/l, y se cultivaron durante una noche. La solución de metales traza estaba compuesta por FeCl₃ x 6 H₂O 27 g/l, ZnCl₂ x 4 H₂O 2 g/l, CaCl₂ x 6 H₂O 2 g/l, Na₂MoO₄ x 2 H₂O 2 g/l, CuSO₄ x 5 H₂O 1,9 g/l, H₃BO₃ 0,5 g/l y HCl concentrado 40 ml/l.

Se inocularon 50 ml de cada cultivo durante una noche en 1 litro de medio de producción en un fermentador con control de temperatura, pH, agitación, aireación y oxígeno disuelto. La composición del medio era como se indica a continuación: KH₂PO₄ 1 g/l, (NH₄)₂SO₄ 0,5 g/l, MgSO₄ x 7 H₂O 0,5 g/l, Bacto casaminoácidos 5 g/l, citrato férrico 0,034 g/l, HCl 1M 0,12 ml/l, ZnCl₂ x 4 H₂O 0,02 g/l, CaCl₂ x 2 H₂O 0,02 g/l, Na₂MoO₄ x 2 H₂O 0,02 g/l, CuSO₄ x 5 H₂O 0,019 g/l, H₃BO₃ 0,005 g/l y 1,25 ml/l de una solución vitamínica. La solución vitamínica contenía riboflavina 0,06 g/l, ácido pantoténico 5,40 g/l, niacina 6,0 g/l, piridoxina 1,4 g/l y ácido fólico 0,01 g/l.

Las fermentaciones se realizaron a 32 °C, pH 6,8 y oxígeno disuelto (DO) igual al 25 % de la saturación. El pH se

mantuvo mediante la adición de NH_4OH , que también sirvió como fuente de nitrógeno para el crecimiento celular. Cuando los 5 g/l iniciales de glucosa casi se habían consumido, se suministró al fermentador una alimentación que consistía en glucosa 500 g/l, 1,6 g/l de KH_2PO_4 , $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 3,9 g/l, citrato férrico 0,13 g/l y 30 ml/l de metanol.

En la fase inicial del crecimiento, la producción de FAME se indujo mediante la adición de IPTG 1 mM y 20 ml/l de metanol puro. Después de que la mayor parte del crecimiento celular se completase, la tasa de alimentación se mantuvo a una tasa de 7,5 g de glucosa/l/hora. Tras la inducción, se suministró metanol de forma continua con la alimentación de glucosa. La fermentación continuó durante un período de 3 días y se tomaron muestras en varios puntos temporales para el análisis de especies grasas como se han descrito anteriormente.

P1A4 y P1G7 produjeron títulos más altos de especies grasas totales (FAME y ácidos grasos libres) en comparación con la cepa de control transformada con pDS57 solo en las fermentaciones a gran escala (FIG. 10). Además, P1A4 y P1G7 produjeron rendimientos mayores de especies grasas totales en comparación con la cepa de control en las fermentaciones a gran escala (FIG. 11).

Los resultados de este ejemplo demuestran métodos para producir células hospedadoras modificadas por ingeniería genética que tienen una expresión de FadR alterada y que presenta títulos y rendimientos potenciados de FAME y ácidos grasos en comparación con las células de tipo silvestre correspondientes.

Se ha de interpretar que el uso de los términos "un", "uno/una", "el/la" y las referencias similares usados en el contexto de la presente invención (especialmente en el contexto de las reivindicaciones siguientes) abarcan tanto el singular como el plural, a menos que se indique de otro modo en el presente documento o se contradiga claramente en el contexto. Las expresiones "que comprende/n", "que tiene/n", "que incluye/n" y "que contiene/n" se han de interpretar como expresiones abiertas (es decir, que significan "que incluye/n, pero sin limitación") a menos que se indique lo contrario. La enumeración de intervalos de valores en el presente documento tiene por objeto simplemente servir como método abreviado de hacer referencia de manera individual a cada valor separado que se encuentre dentro del intervalo, a menos que se indique otra cosa en el presente documento, y cada valor separado se incorpora a la memoria descriptiva como si se enumerara de manera individual en el presente documento. Todos los métodos que se describen en el presente documento pueden realizarse en cualquier orden adecuado a menos que se indique otra cosa en el presente documento o se contradiga claramente de otro modo por el contexto. El uso de cualquiera y todos los ejemplos o lenguaje de ejemplo (por ejemplo, "tal como") que se proporcionan en el presente documento, tiene por objeto simplemente iluminar mejor la invención y no plantea una limitación en el alcance de la invención, a menos que se reivindique otra cosa. Ningún lenguaje en la memoria descriptiva ha de interpretarse como una indicación de que cualquier elemento no reivindicado sea esencial para la puesta en práctica de la invención.

LISTADO DE SECUENCIAS

<110> LS9, INC.

<120> MÉTODOS Y COMPOSICIONES PARA LA PRODUCCIÓN MEJORADA DE ÁCIDOS GRASOS Y DERIVADOS DE LOS MISMOS

<130> LS00034 PCT

<140>

<141>

<150> 61/470.989

<151> 01-04-2011

<160> 23

<170> PatentIn versión 3.5

<210> 1

<211> 239

<212> PRT

<213> *Escherichia coli*

<400> 1

```

Met Val Ile Lys Ala Gln Ser Pro Ala Gly Phe Ala Glu Glu Tyr Ile
1           5           10           15

Ile Glu Ser Ile Trp Asn Asn Arg Phe Pro Pro Gly Thr Ile Leu Pro
          20           25           30

Ala Glu Arg Glu Leu Ser Glu Leu Ile Gly Val Thr Arg Thr Thr Leu
          35           40           45

Arg Glu Val Leu Gln Arg Leu Ala Arg Asp Gly Trp Leu Thr Ile Gln
50           55           60

His Gly Lys Pro Thr Lys Val Asn Asn Phe Trp Glu Thr Ser Gly Leu
65           70           75           80

Asn Ile Leu Glu Thr Leu Ala Arg Leu Asp His Glu Ser Val Pro Gln
          85           90           95

Leu Ile Asp Asn Leu Leu Ser Val Arg Thr Asn Ile Ser Thr Ile Phe
          100          105          110

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

Ala Thr Ala Asn Glu Val Ala Asp His Ala Asp Ala Phe Ala Glu Leu
          130          135          140

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

Tyr Gly Leu Ile Leu Asn Gly Met Lys Gly Leu Tyr Thr Arg Ile Gly
          165          170          175

Arg His Tyr Phe Ala Asn Pro Glu Ala Arg Ser Leu Ala Leu Gly Phe
          180          185          190

Tyr His Lys Leu Ser Ala Leu Cys Ser Glu Gly Ala His Asp Gln Val
          195          200          205

Tyr Glu Thr Val Arg Arg Tyr Gly His Glu Ser Gly Glu Ile Trp His
          210          215          220

Arg Met Gln Lys Asn Leu Pro Gly Asp Leu Ala Ile Gln Gly Arg
225          230          235

```

5 <210> 2
 <211> 473
 <212> PRT

<213> *Marinobacter hydrocarbonoclasticus*

<400> 2

```

Met Lys Arg Leu Gly Thr Leu Asp Ala Ser Trp Leu Ala Val Glu Ser
 1          5          10          15

Glu Asp Thr Pro Met His Val Gly Thr Leu Gln Ile Phe Ser Leu Pro
          20          25          30

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

Glu Ala Gly Asp Val Ala Pro Pro Trp Gly Tyr Lys Leu Ala Trp Ser
          50          55          60

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

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

Gly Glu Arg Glu Leu Gly Ile Leu Val Ser Arg Leu His Ser Asn Pro
          100          105          110

Leu Asp Phe Ser Arg Pro Leu Trp Glu Cys His Val Ile Glu Gly Leu
          115          120          125

Glu Asn Asn Arg Phe Ala Leu Tyr Thr Lys Met His His Ser Met Ile
          130          135          140

```

Asp Gly Ile Ser Gly Val Arg Leu Met Gln Arg Val Leu Thr Thr Asp
 145 150 155 160

Pro Glu Arg Cys Asn Met Pro Pro Pro Trp Thr Val Arg Pro His Gln
 165 170 175

Arg Arg Gly Ala Lys Thr Asp Lys Glu Ala Ser Val Pro Ala Ala Val
 180 185 190

Ser Gln Ala Met Asp Ala Leu Lys Leu Gln Ala Asp Met Ala Pro Arg
 195 200 205

Leu Trp Gln Ala Gly Asn Arg Leu Val His Ser Val Arg His Pro Glu
 210 215 220

Asp Gly Leu Thr Ala Pro Phe Thr Gly Pro Val Ser Val Leu Asn His
 225 230 235 240

Arg Val Thr Ala Gln Arg Arg Phe Ala Thr Gln His Tyr Gln Leu Asp
 245 250 255

Arg Leu Lys Asn Leu Ala His Ala Ser Gly Gly Ser Leu Asn Asp Ile
 260 265 270

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

Asn Asn Leu Pro Asp Thr Pro Leu Thr Ala Gly Ile Pro Val Asn Ile
 290 295 300

Arg Pro Ala Asp Asp Glu Gly Thr Gly Thr Gln Ile Ser Phe Met Ile
 305 310 315 320

Ala Ser Leu Ala Thr Asp Glu Ala Asp Pro Leu Asn Arg Leu Gln Gln
 325 330 335

Ile Lys Thr Ser Thr Arg Arg Ala Lys Glu His Leu Gln Lys Leu Pro
 340 345 350

Lys Ser Ala Leu Thr Gln Tyr Thr Met Leu Leu Met Ser Pro Tyr Ile
 355 360 365

Leu Gln Leu Met Ser Gly Leu Gly Gly Arg Met Arg Pro Val Phe Asn
 370 375 380

Val Thr Ile Ser Asn Val Pro Gly Pro Glu Gly Thr Leu Tyr Tyr Glu
 385 390 395 400

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

Gly Ala Leu Asn Ile Thr Cys Leu Ser Tyr Ala Gly Ser Leu Asn Phe
420 425 430

Gly Phe Thr Gly Cys Arg Asp Thr Leu Pro Ser Met Gln Lys Leu Ala
435 440 445

Val Tyr Thr Gly Glu Ala Leu Asp Glu Leu Glu Ser Leu Ile Leu Pro
450 455 460

Pro Lys Lys Arg Ala Arg Thr Arg Lys
465 470

<210> 3

<211> 455

<212> PRT

<213> *Marinobacter hydrocarbonoclasticus*

<400> 3

Met Thr Pro Leu Asn Pro Thr Asp Gln Leu Phe Leu Trp Leu Glu Lys
1 5 10 15

Arg Gln Gln Pro Met His Val Gly Gly Leu Gln Leu Phe Ser Phe Pro
20 25 30

Glu Gly Ala Pro Asp Asp Tyr Val Ala Gln Leu Ala Asp Gln Leu Arg
35 40 45

Gln Lys Thr Glu Val Thr Ala Pro Phe Asn Gln Arg Leu Ser Tyr Arg
50 55 60

Leu Gly Gln Pro Val Trp Val Glu Asp Glu His Leu Asp Leu Glu His
65 70 75 80

His Phe Arg Phe Glu Ala Leu Pro Thr Pro Gly Arg Ile Arg Glu Leu
85 90 95

Leu Ser Phe Val Ser Ala Glu His Ser His Leu Met Asp Arg Glu Arg
100 105 110

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

Ala Leu Tyr Thr Lys Val His His Ser Leu Val Asp Gly Val Ser Ala
130 135 140

5

10

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

Val Ser Ile Asp Met Asp Arg Leu Ala Leu Asn Met Thr Leu Thr Ser
 405 410 415

Tyr Asn Asp Gln Val Glu Phe Gly Leu Ile Gly Cys Arg Arg Thr Leu
 420 425 430

Pro Ser Leu Gln Arg Met Leu Asp Tyr Leu Glu Gln Gly Leu Ala Glu
 435 440 445

Leu Glu Leu Asn Ala Gly Leu
 450 455

<210> 4

<211> 457

5 <212> PRT

<213> *Alcanivorax borkumensis*

<400> 4

Met Lys Ala Leu Ser Pro Val Asp Gln Leu Phe Leu Trp Leu Glu Lys
 1 5 10 15

Arg Gln Gln Pro Met His Val Gly Gly Leu Gln Leu Phe Ser Phe Pro
 20 25 30

Glu Gly Ala Gly Pro Lys Tyr Val Ser Glu Leu Ala Gln Gln Met Arg
 35 40 45

Asp Tyr Cys His Pro Val Ala Pro Phe Asn Gln Arg Leu Thr Arg Arg
 50 55 60

Leu Gly Gln Tyr Tyr Trp Thr Arg Asp Lys Gln Phe Asp Ile Asp His
 65 70 75 80

His Phe Arg His Glu Ala Leu Pro Lys Pro Gly Arg Ile Arg Glu Leu
 85 90 95

Leu Ser Leu Val Ser Ala Glu His Ser Asn Leu Leu Asp Arg Glu Arg
 100 105 110

Pro Met Trp Glu Ala His Leu Ile Glu Gly Ile Arg Gly Arg Gln Phe
 115 120 125

Ala Leu Tyr Tyr Lys Ile His His Ser Val Met Asp Gly Ile Ser Ala
 130 135 140

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

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

Ser Tyr Val Asp Gln Met Glu Phe Gly Leu Ile Ala Cys Arg Arg Thr
420 425 430

Leu Pro Ser Met Gln Arg Leu Leu Asp Tyr Leu Glu Gln Ser Ile Arg
435 440 445

Glu Leu Glu Ile Gly Ala Gly Ile Lys
450 455

<210> 5
<211> 451
<212> PRT
<213> *Alcanivorax borkumensis*
<400> 5

Met Ala Arg Lys Leu Ser Ile Met Asp Ser Gly Trp Leu Met Met Glu
1 5 10 15

Thr Arg Glu Thr Pro Met His Val Gly Gly Leu Ala Leu Phe Ala Ile
20 25 30

Pro Glu Gly Ala Pro Glu Asp Tyr Val Glu Ser Ile Tyr Arg Tyr Leu
35 40 45

Val Asp Val Asp Ser Ile Cys Arg Pro Phe Asn Gln Lys Ile Gln Ser
50 55 60

His Leu Pro Leu Tyr Leu Asp Ala Thr Trp Val Glu Asp Lys Asn Phe
65 70 75 80

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

Val Arg Glu Leu Leu Ala Leu Val Ser Arg Leu His Ala Gln Arg Leu
100 105 110

Asp Pro Ser Arg Pro Leu Trp Glu Ser Tyr Leu Ile Glu Gly Leu Glu
115 120 125

Gly Asn Arg Phe Ala Leu Tyr Thr Lys Met His His Ser Met Val Asp
130 135 140

Gly Val Ala Gly Met His Leu Met Gln Ser Arg Leu Ala Thr Cys Ala
145 150 155 160

Glu Asp Arg Leu Pro Ala Pro Trp Ser Gly Glu Trp Asp Ala Glu Lys
165 170 175

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

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

Leu Glu Pro
450

<210> 6
<211> 455
<212> PRT
<213> *Marinobacter aquaeolei*

<400> 6

Met Thr Pro Leu Asn Pro Thr Asp Gln Leu Phe Leu Trp Leu Glu Lys
1 5 10 15

Arg Gln Gln Pro Met His Val Gly Gly Leu Gln Leu Phe Ser Phe Pro
20 25 30

Glu Gly Ala Pro Asp Asp Tyr Val Ala Gln Leu Ala Asp Gln Leu Arg
35 40 45

Gln Lys Thr Glu Val Thr Ala Pro Phe Asn Gln Arg Leu Ser Tyr Arg
50 55 60

Leu Gly Gln Pro Val Trp Val Glu Asp Glu His Leu Asp Leu Glu His
65 70 75 80

His Phe Arg Phe Glu Ala Leu Pro Thr Pro Gly Arg Ile Arg Glu Leu
85 90 95

Leu Ser Phe Val Ser Ala Glu His Ser His Leu Met Asp Arg Glu Arg
100 105 110

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

Ala Leu Tyr Thr Lys Val His His Ser Leu Val Asp Gly Val Ser Ala
130 135 140

Met Arg Met Ala Thr Arg Met Leu Ser Glu Asn Pro Asp Glu His Gly
145 150 155 160

Met Pro Pro Ile Trp Asp Leu Pro Cys Leu Ser Arg Asp Arg Gly Glu
165 170 175

Ser Asp Gly His Ser Leu Trp Arg Ser Val Thr His Leu Leu Gly Leu
180 185 190

5

10

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

Lys Thr Ile Asn Gln Ala Arg Lys Asp Pro Ala Tyr Asp Ser Ile Phe
 210 215 220

His Ala Pro Arg Cys Met Leu Asn Gln Lys Ile Thr Gly Ser Arg Arg
 225 230 235 240

Phe Ala Ala Gln Ser Trp Cys Leu Lys Arg Ile Arg Ala Val Cys Glu
 245 250 255

Ala Tyr Gly Thr Thr Val Asn Asp Val Val Thr Ala Met Cys Ala Ala
 260 265 270

Ala Leu Arg Thr Tyr Leu Met Asn Gln Asp Ala Leu Pro Glu Lys Pro
 275 280 285

Leu Val Ala Phe Val Pro Val Ser Leu Arg Arg Asp Asp Ser Ser Gly
 290 295 300

Gly Asn Gln Val Gly Val Ile Leu Ala Ser Leu His Thr Asp Val Gln
 305 310 315 320

Glu Ala Gly Glu Arg Leu Leu Lys Ile His His Gly Met Glu Glu Ala
 325 330 335

Lys Gln Arg Tyr Arg His Met Ser Pro Glu Glu Ile Val Asn Tyr Thr
 340 345 350

Ala Leu Thr Leu Ala Pro Ala Ala Phe His Leu Leu Thr Gly Leu Ala
 355 360 365

Pro Lys Trp Gln Thr Phe Asn Val Val Ile Ser Asn Val Pro Gly Pro
 370 375 380

Ser Arg Pro Leu Tyr Trp Asn Gly Ala Lys Leu Glu Gly Met Tyr Pro
 385 390 395 400

Val Ser Ile Asp Met Asp Arg Leu Ala Leu Asn Met Thr Leu Thr Ser
 405 410 415

Tyr Asn Asp Gln Val Glu Phe Gly Leu Ile Gly Cys Arg Arg Thr Leu
 420 425 430

Pro Ser Leu Gln Arg Met Leu Asp Tyr Leu Glu Gln Gly Leu Ala Glu
 435 440 445

Leu Glu Leu Asn Ala Gly Leu
450 455

5 <210> 7
<211> 473
<212> PRT
<213> *Marinobacter aquaeolei*

<400> 7

```

Met Lys Arg Leu Gly Thr Leu Asp Ala Ser Trp Leu Ala Val Glu Ser
1      5      10      15

Glu Asp Thr Pro Met His Val Gly Thr Leu Gln Ile Phe Ser Leu Pro
20      25      30

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

Glu Ala Gly Asp Val Ala Pro Pro Trp Gly Tyr Lys Leu Ala Trp Ser
50      55      60

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

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

Gly Glu Arg Glu Leu Gly Ile Leu Val Ser Arg Leu His Ser Asn Pro
100     105     110

Leu Asp Phe Ser Arg Pro Leu Trp Glu Cys His Val Ile Glu Gly Leu
115     120     125

Glu Asn Asn Arg Phe Ala Leu Tyr Thr Lys Met His His Ser Met Ile
130     135     140

Asp Gly Ile Ser Gly Val Arg Leu Met Gln Arg Val Leu Thr Thr Asp
145     150     155     160

Pro Glu Arg Cys Asn Met Pro Pro Pro Trp Thr Val Arg Pro His Gln
165     170     175

Arg Arg Gly Ala Lys Thr Asp Lys Glu Ala Ser Val Pro Ala Ala Val
180     185     190

Ser Gln Ala Met Asp Ala Leu Lys Leu Gln Ala Asp Met Ala Pro Arg
195     200     205

```

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

Pro Lys Lys Arg Ala Arg Thr Arg Lys
465 470

<210> 8
<211> 1173
<212> PRT
<213> *Mycobacterium smegmatis*

<400> 8

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

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

Arg Gly Glu Leu Leu Val Arg Ser Gln Thr Leu Thr Pro Gly Tyr Tyr
 465 470 475 480

Lys Arg Pro Glu Val Thr Ala Ser Val Phe Asp Arg Asp Gly Tyr Tyr
 485 490 495

His Thr Gly Asp Val Met Ala Glu Thr Ala Pro Asp His Leu Val Tyr
 500 505 510

Val Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val
 515 520 525

Ala Val Ala Asn Leu Glu Ala Val Phe Ser Gly Ala Ala Leu Val Arg
 530 535 540

Gln Ile Phe Val Tyr Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val
 545 550 555 560

Val Val Pro Thr Pro Glu Ala Leu Glu Gln Tyr Asp Pro Ala Ala Leu
 565 570 575

Lys Ala Ala Leu Ala Asp Ser Leu Gln Arg Thr Ala Arg Asp Ala Glu
 580 585 590

Leu Gln Ser Tyr Glu Val Pro Ala Asp Phe Ile Val Glu Thr Glu Pro
 595 600 605

Phe Ser Ala Ala Asn Gly Leu Leu Ser Gly Val Gly Lys Leu Leu Arg
 610 615 620

Pro Asn Leu Lys Asp Arg Tyr Gly Gln Arg Leu Glu Gln Met Tyr Ala
 625 630 635 640

Asp Ile Ala Ala Thr Gln Ala Asn Gln Leu Arg Glu Leu Arg Arg Ala
 645 650 655

Ala Ala Thr Gln Pro Val Ile Asp Thr Leu Thr Gln Ala Ala Ala Thr
 660 665 670

Ile Leu Gly Thr Gly Ser Glu Val Ala Ser Asp Ala His Phe Thr Asp
 675 680 685

Leu Gly Gly Asp Ser Leu Ser Ala Leu Thr Leu Ser Asn Leu Leu Ser
 690 695 700

Asp Phe Phe Gly Phe Glu Val Pro Val Gly Thr Ile Val Asn Pro Ala

705		710		715		720									
Thr	Asn	Leu	Ala	Gln	Leu	Ala	Gln	His	Ile	Glu	Ala	Gln	Arg	Thr	Ala
				725					730					735	
Gly	Asp	Arg	Arg	Pro	Ser	Phe	Thr	Thr	Val	His	Gly	Ala	Asp	Ala	Thr
			740					745					750		
Glu	Ile	Arg	Ala	Ser	Glu	Leu	Thr	Leu	Asp	Lys	Phe	Ile	Asp	Ala	Glu
		755					760					765			
Thr	Leu	Arg	Ala	Ala	Pro	Gly	Leu	Pro	Lys	Val	Thr	Thr	Glu	Pro	Arg
	770					775						780			
Thr	Val	Leu	Leu	Ser	Gly	Ala	Asn	Gly	Trp	Leu	Gly	Arg	Phe	Leu	Thr
785					790					795					800
Leu	Gln	Trp	Leu	Glu	Arg	Leu	Ala	Pro	Val	Gly	Gly	Thr	Leu	Ile	Thr
			805						810					815	
Ile	Val	Arg	Gly	Arg	Asp	Asp	Ala	Ala	Ala	Arg	Ala	Arg	Leu	Thr	Gln
			820					825					830		
Ala	Tyr	Asp	Thr	Asp	Pro	Glu	Leu	Ser	Arg	Arg	Phe	Ala	Glu	Leu	Ala
		835					840					845			
Asp	Arg	His	Leu	Arg	Val	Val	Ala	Gly	Asp	Ile	Gly	Asp	Pro	Asn	Leu
	850					855					860				
Gly	Leu	Thr	Pro	Glu	Ile	Trp	His	Arg	Leu	Ala	Ala	Glu	Val	Asp	Leu
865					870					875				880	
Val	Val	His	Pro	Ala	Ala	Leu	Val	Asn	His	Val	Leu	Pro	Tyr	Arg	Gln
				885					890					895	
Leu	Phe	Gly	Pro	Asn	Val	Val	Gly	Thr	Ala	Glu	Val	Ile	Lys	Leu	Ala
			900					905					910		
Leu	Thr	Glu	Arg	Ile	Lys	Pro	Val	Thr	Tyr	Leu	Ser	Thr	Val	Ser	Val
		915					920					925			
Ala	Met	Gly	Ile	Pro	Asp	Phe	Glu	Glu	Asp	Gly	Asp	Ile	Arg	Thr	Val
	930					935					940				
Ser	Pro	Val	Arg	Pro	Leu	Asp	Gly	Gly	Tyr	Ala	Asn	Gly	Tyr	Gly	Asn
945					950					955					960

Ser Lys Trp Ala Gly Glu Val Leu Leu Arg Glu Ala His Asp Leu Cys
965 970 975

Gly Leu Pro Val Ala Thr Phe Arg Ser Asp Met Ile Leu Ala His Pro
980 985 990

Arg Tyr Arg Gly Gln Val Asn Val Pro Asp Met Phe Thr Arg Leu Le
995 1000 1005

Leu Ser Leu Leu Ile Thr Gly Val Ala Pro Arg Ser Phe Tyr Ile
1010 1015 1020

Gly Asp Gly Glu Arg Pro Arg Ala His Tyr Pro Gly Leu Thr Val
1025 1030 1035

Asp Phe Val Ala Glu Ala Val Thr Thr Leu Gly Ala Gln Gln Arg
1040 1045 1050

Glu Gly Tyr Val Ser Tyr Asp Val Met Asn Pro His Asp Asp Gly
1055 1060 1065

Ile Ser Leu Asp Val Phe Val Asp Trp Leu Ile Arg Ala Gly His
1070 1075 1080

Pro Ile Asp Arg Val Asp Asp Tyr Asp Asp Trp Val Arg Arg Phe
1085 1090 1095

Glu Thr Ala Leu Thr Ala Leu Pro Glu Lys Arg Arg Ala Gln Thr
1100 1105 1110

Val Leu Pro Leu Leu His Ala Phe Arg Ala Pro Gln Ala Pro Leu
1115 1120 1125

Arg Gly Ala Pro Glu Pro Thr Glu Val Phe His Ala Ala Val Arg
1130 1135 1140

Thr Ala Lys Val Gly Pro Gly Asp Ile Pro His Leu Asp Glu Ala
1145 1150 1155

Leu Ile Asp Lys Tyr Ile Arg Asp Leu Arg Glu Phe Gly Leu Ile
1160 1165 1170

<210> 9
<211> 341
<212> PRT
<213> *Acinetobacter baylyi*

<400> 9

Met Ala Thr Thr Asn Val Ile His Ala Tyr Ala Ala Met Gln Ala Gly

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

Asn Gly Thr Phe His Phe Leu Gly Val Val Met Glu Pro Ile Pro Val
260 265 270

Pro Val Gly Ala Leu Leu Gly Gly Ala Lys Ser Leu Thr Ala Ser Pro
275 280 285

Thr Gly Ser Pro Ala Ala Leu Arg Lys Leu Leu Glu Phe Ala Ala Arg
290 295 300

Lys Asn Ile Ala Pro Gln Ile Glu Met Tyr Pro Met Ser Glu Leu Asn
305 310 315 320

Glu Ala Ile Glu Arg Leu His Ser Gly Gln Ala Arg Tyr Arg Ile Val
325 330 335

Leu Lys Ala Asp Phe
340

<210> 10
<211> 339
<212> PRT
<213> *Escherichia coli*

<400> 10

Met Ser Met Ile Lys Ser Tyr Ala Ala Lys Glu Ala Gly Gly Glu Leu
1 5 10 15

Glu Val Tyr Glu Tyr Asp Pro Gly Glu Leu Arg Pro Gln Asp Val Glu
20 25 30

Val Gln Val Asp Tyr Cys Gly Ile Cys His Ser Asp Leu Ser Met Ile
35 40 45

Asp Asn Glu Trp Gly Phe Ser Gln Tyr Pro Leu Val Ala Gly His Glu
50 55 60

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

Leu Gln Val Gly Gln Arg Val Gly Ile Gly Trp Thr Ala Arg Ser Cys
85 90 95

Gly His Cys Asp Ala Cys Ile Ser Gly Asn Gln Ile Asn Cys Glu Gln
100 105 110

Gly Ala Val Pro Thr Ile Met Asn Arg Gly Gly Phe Ala Glu Lys Leu
115 120 125

Arg Ala Asp Trp Gln Trp Val Ile Pro Leu Pro Glu Asn Ile Asp Ile

5

10

130	135	140
Glu Ser Ala Gly Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Lys Pro		
145	150	155 160
Leu Leu Met His His Ile Thr Ala Thr Ser Arg Val Gly Val Ile Gly		
	165	170 175
Ile Gly Gly Leu Gly His Ile Ala Ile Lys Leu Leu His Ala Met Gly		
	180	185 190
Cys Glu Val Thr Ala Phe Ser Ser Asn Pro Ala Lys Glu Gln Glu Val		
	195	200 205
Leu Ala Met Gly Ala Asp Lys Val Val Asn Ser Arg Asp Pro Gln Ala		
	210	215 220
Leu Lys Ala Leu Ala Gly Gln Phe Asp Leu Ile Ile Asn Thr Val Asn		
	225	230 235 240
Val Ser Leu Asp Trp Gln Pro Tyr Phe Glu Ala Leu Thr Tyr Gly Gly		
	245	250 255
Asn Phe His Thr Val Gly Ala Val Leu Thr Pro Leu Ser Val Pro Ala		
	260	265 270
Phe Thr Leu Ile Ala Gly Asp Arg Ser Val Ser Gly Ser Ala Thr Gly		
	275	280 285
Thr Pro Tyr Glu Leu Arg Lys Leu Met Arg Phe Ala Ala Arg Ser Lys		
	290	295 300
Val Ala Pro Thr Thr Glu Leu Phe Pro Met Ser Lys Ile Asn Asp Ala		
	305	310 315 320
Ile Gln His Val Arg Asp Gly Lys Ala Arg Tyr Arg Val Val Leu Lys		
	325	330 335
Ala Asp Phe		

<210> 11
 <211> 208
 <212> PRT
 <213> *Escherichia coli*

<400> 11

ES 2 808 287 T9

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

5

<210> 12
 <211> 209
 <212> PRT
 <213> *Escherichia coli*

10

<400> 12

```

Met Val Asp Met Lys Thr Thr His Thr Ser Leu Pro Phe Ala Gly His
1          5          10          15

Thr Leu His Phe Val Glu Phe Asp Pro Ala Asn Phe Cys Glu Gln Asp
          20          25          30

Leu Leu Trp Leu Pro His Tyr Ala Gln Leu Gln His Ala Gly Arg Lys
          35          40          45

Arg Lys Thr Glu His Leu Ala Gly Arg Ile Ala Ala Val Tyr Ala Leu
          50          55          60

Arg Glu Tyr Gly Tyr Lys Cys Val Pro Ala Ile Gly Glu Leu Arg Gln
65          70          75          80

Pro Val Trp Pro Ala Glu Val Tyr Gly Ser Ile Ser His Cys Gly Thr
          85          90          95

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

Glu Ile Phe Ser Val Gln Thr Ala Arg Glu Leu Thr Asp Asn Ile Ile
          115          120          125

Thr Pro Ala Glu His Glu Arg Leu Ala Asp Cys Gly Leu Ala Phe Ser
          130          135          140

Leu Ala Leu Thr Leu Ala Phe Ser Ala Lys Glu Ser Ala Phe Lys Ala
145          150          155          160

Ser Glu Ile Gln Thr Asp Ala Gly Phe Leu Asp Tyr Gln Ile Ile Ser
          165          170          175

Trp Asn Lys Gln Gln Val Ile Ile His Arg Glu Asn Glu Met Phe Ala
          180          185          190

Val His Trp Gln Ile Lys Glu Lys Ile Val Ile Thr Leu Cys Gln His
          195          200          205

```

Asp

- 5 <210> 13
 <211> 12397
 <212> ADN
 <213> Secuencia artificial
- 10 <220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Polinucleótido sintético"

ES 2 808 287 T9

<400> 13

cactatacca attgagatgg gctagtcaat gataattact agtccttttc ctttgagttg 60

tgggtatctg taaattctgc tagacctttg ctggaaaact tgtaaattct gctagaccct 120

ctgtaaattc cgctagacct ttgtgtgttt tttttgttta tattcaagtg gttataattt	180
atagaataaa gaaagaataa aaaaagataa aaagaataga tcccagccct gtgtataact	240
cactacttta gtcagttccg cagtattaca aaaggatgtc gcaaacgctg tttgctcctc	300
tacaaaacag accttaaaac cctaaaggcg tcggcatccg cttacagaca agctgtgacc	360
gtctccggga gctgcatgtg tcagaggttt tcaccgtcat caccgaaacg cgcgaggcag	420
cagatcaatt cgcgcgcgaa ggcgaagcgg catgcattta cgttgacacc atcgaatggt	480
gcaaaacctt tcgcggtatg gcatgatagc gcccggaaga gagtcaattc aggggtggtga	540
atgtgaaacc agtaacgtta tacgatgtcg cagagtatgc cgggtgtctct tatcagaccg	600
tttcccgcgt ggtgaaccag gccagccacg tttctgcgaa aacgcgggaa aaagtggaag	660
cggcgatggc ggagctgaat tacattccca accgcgtggc acaacaactg gcgggcaaac	720
agtcgttget gattggcgtt gccacctcca gtctggccct gcacgcgcgc tcgcaaattg	780
tcgcggcgat taaatctcgc gccgatcaac tgggtgccag cgtggtggtg tcgatggtag	840
aacgaagcgg cgtcgaagcc tgtaagcgg cgggtgcacaa tcttctcgcg caacgcgtca	900
gtgggctgat cattaactat ccgctggatg accaggatgc cattgctgtg gaagctgcct	960
gcactaatgt tccggcggtta tttcttgatg tctctgacca gacacccatc aacagtatta	1020
ttttctccca tgaagacggt acgcgactgg gcgtggagca tctggtcgca ttgggtcacc	1080
agcaaatcgc gctgttagcg ggcccattaa gttctgtctc ggcgcgctctg cgtctggtg	1140
gctggcataa atatctcact cgcaatcaaa ttcagccgat agcggaacgg gaaggcgact	1200
ggagtgccat gtccggtttt caacaaacca tgcaaagtct gaatgagggc atcgttccca	1260
ctgcgatgct ggttgccaac gatcagatgg cgctgggcgc aatgcgcgcc attaccgagt	1320
cggggtgcg cgttggtgcg gatattctcg tagtgggata cgacgatacc gaagacagct	1380
catgttatat cccgcggtta accaccatca aacaggattt tcgcctgctg gggcaaacca	1440
gcgtggaccg cttgctgcaa ctctctcagg gccaggcggg gaagggcaat cagctgttgc	1500
ccgtctcact ggtgaaaaga aaaaccaccc tggcgcccaa tacgcaaacc gcctctcccc	1560
gcgcgttggc cgattcatta atgcagctgg cagcacaggt tccccgactg gaaagcgggc	1620
agtgagcgca acgcaattaa tgtaagttag cgcgaattga tctggtttga cagcttatca	1680
tcgactgcac ggtgcaccaa tgcttctggc gtcaggcagc catcggaagc tgtggtatgg	1740
ctgtgcaggc cgtaaatcac tgcataattc gtgtcgctca aggcgcactc ccgttctgga	1800
taatgttttt tgcgccgaca tcataacggg tctggcaaat attctgaaat gagctgttga	1860
caattaatca tccggctcgt ataatgtgtg gaattgtgag cggataacaa tttcacacag	1920
gaaacagcgc cgctgagaaa aagcgaagcg gcaactgctct ttaacaattt atcagacaat	1980

ctgtgtgggc	actcgaccgg	aattatcgat	taactttatt	attaaaaatt	aaagaggtat	2040
atattaatgt	atcgattaaa	taaggaggaa	taaaccatga	cgagcgatgt	tcacgacgcg	2100
accgacggcg	ttaccgagac	tgacttgat	gatgagcaga	gcactcgtcg	tattgcagaa	2160
ctgtacgcaa	cggaccaga	gttcgcagca	gcagctcctc	tgccggccgt	tgctgatgog	2220
gcgcacaaac	cgggcctgcg	tctggcggaa	atcctgcaga	ccctgttcac	cggctacggc	2280
gatcgtccgg	cgctgggcta	tcgtgcacgt	gagctggcga	cggacgaagg	cggctcgtacg	2340
gtcacgcgtc	tgctgccgcg	cttcgatacc	ctgacctatg	cacaggtgtg	gagccgtgtt	2400
caagcagtgg	ctgcagcgtt	gcgtcacaat	ttcgcacaaac	cgatttacc	gggcgacgog	2460
gtcgcgacta	tgggctttgc	gagcccggac	tatttgacgc	tggatctggt	gtgcgcgtat	2520
ctgggcctgg	tcagcgttcc	tttgacgcat	aacgctccgg	tgtctcgcct	ggccccgatt	2580
ctggccgagg	tggaaaccgog	tattctgacg	gtgagcgcag	aatacctgga	cctggcgggt	2640
gaatccgtcc	gtgatgtgaa	ctccgtcagc	cagctgggtg	ttttcgacca	tcacccggaa	2700
gtggacgatc	accgtgacgc	actggctcgc	gcacgcgagc	agctggccgg	caaaggtatc	2760
gcagttacga	ccctggatgc	gatcgcagac	gaaggcgcag	gtttgccggc	tgagccgatt	2820
tacacggcgg	atcacgatca	gcgtctggcc	atgattctgt	ataccagcgg	ctctacgggt	2880
gtcccgaaag	gcgcgatgta	caccgaagcg	atggtggctc	gcctgtggac	tatgagcttt	2940
atcacggggcg	acccgacccc	ggttatcaac	gtgaacttca	tgccgctgaa	ccatctgggc	3000
ggtcgtatcc	cgattagcac	cgcctgcag	aatggcggta	ccagctactt	cgttccggaa	3060
agcgacatga	gcacgctgtt	tgaggatctg	gccctggctc	gccctaccga	actgggtctg	3120
gtgccgcgtg	ttgcggacat	gctgtaccag	catcatctgg	cgaccgtgga	tcgcctggtg	3180
accagggcg	cggacgaact	gactgcggaa	aagcaggccg	gtgcggaact	gcgtgaacag	3240
gtcttggggcg	gtcgtgttat	caccggtttt	gtttccaccg	cgcctgtggc	ggcagagatg	3300
cgtgcttttc	tggatatcac	cttgggtgca	cacatcgttg	acggttacgg	tctgaccgaa	3360
accggtgcgg	tcaccctgta	tgggtgtgatt	gttcgtcctc	cggtcattga	ttacaagctg	3420
atcgatgtgc	cggagctggg	ttacttctcc	accgacaaac	cgtaccgcgg	tggcgagctg	3480
ctggttcgta	gccaaacgtt	gactccgggt	tactacaagc	gccagaagt	caccgcgtcc	3540
gttttcgatc	gcgacggcta	ttaccacacc	ggcgacgtga	tggcagaaac	cgcgccagac	3600
cacctggtgt	atgtggaccg	cgcacaacat	gttctgaagc	tggcgcaagg	tgaatttgtc	3660
gccgtggcta	acctggaggc	cgttttcagc	ggcgtgctc	tggtcgccca	gattttcgtg	3720
tatggtaaca	gcgagcgcag	ctttctgttg	gctgttggtg	tcctacccc	ggaggcgtg	3780
gagcaatacg	accctgccgc	attgaaagca	gccctggcgg	attcgctgca	gcgtacggcg	3840
cgtgatgccg	agctgcagag	ctatgaagtg	cggcgggact	tcattgttga	gactgagcct	3900

tttagcgctg	cgaacgggtct	gctgagcggg	gttggcaagt	tgctgcgctcc	gaatttgaag	3960
gatcgctacg	gtcagcggtt	ggagcagatg	tacgcggaca	tcgcgggtac	gcaggcgaac	4020
caattgcggtg	aactgcgcgg	tgctgcggtc	actcaaccgg	tgatcgacac	gctgacgcaa	4080
gctgcggcga	ccatcctggg	taccggcagc	gaggttgcaa	gcgacgcaca	ctttactgat	4140
ttgggcggtg	attctctgag	cgcgctgacg	ttgagcaact	tgctgtctga	cttctttggc	4200
tttgaagtcc	cggttggcac	gattgttaac	ccagcgacta	atctggcaca	gctggcgcaa	4260
catatcgagg	cgcagcgcac	ggcgggtgac	cgccgtccat	cctttacgac	ggccacgggt	4320
gcggatgcta	cggaaatccg	tgcaagcgaa	ctgactctgg	acaaattcat	cgacgctgag	4380
actctgcgcg	cagcacctgg	tttgccgaag	gttacgactg	agccgcgtac	ggcctctgtg	4440
agcgggtgcca	atggttgggt	gggcccgttc	ctgaccctgc	agtggctgga	acgtttggca	4500
ccggttggcg	gtaccctgat	caccattgtg	cgcggtcgtg	acgatgcagc	ggcacgtgca	4560
cgtttgactc	aggcttacga	tacggaccca	gagctgtccc	gccgcttcgc	tgagttggcg	4620
gatcgccact	tgcggtgtgt	ggcaggtgat	atcggcgatc	cgaatctggg	cctgaccccg	4680
gagatttggc	accgtctggc	agcagaggtc	gatctggtcg	ttcatccagc	ggccctggtc	4740
aaccacgtcc	tgccgtaccg	ccagctgttt	ggcccgaaatg	ttgttggcac	cgccgaagtt	4800
atcaagttgg	ctctgaccga	gcgcacaaag	cctgttacct	acctgtccac	ggttagcgtc	4860
gcgatgggta	ttcctgattt	tgaggaggac	ggtgacattc	gtaccgtcag	cccgggtcgt	4920
ccgctggatg	gtggctatgc	aaatggctat	ggcaacagca	agtgggctgg	cgaggtgctg	4980
ctgcgcgagg	cacatgacct	gtgtggcctg	ccggttgcca	cgtttcgtag	cgacatgatt	5040
ctggcccacc	cgcgctaccg	tggccaagtg	aatgtgccgg	acatgttcac	ccgtctgctg	5100
ctgtccctgc	tgatcacggg	tgtggcaccg	cgttccttct	acattgggtga	tggcgagcgt	5160
ccgcgtgcac	actaccggg	cctgaccgtc	gattttgttg	cggagcggt	tactaccctg	5220
ggtgctcagc	aacgtgagg	ttatgtctcg	tatgacgtta	tgaatccgca	cgatgacgggt	5280
attagcttgg	atgtctttgt	ggactggctg	attcgtgcgg	gccaccaat	tgaccgtgtt	5340
gacgactatg	atgactgggt	gcgtcgtttt	gaaaccgcgt	tgaccgcctt	gccggagaaa	5400
cgtcgtgcgc	agaccgttct	gccgctgctg	catgccttct	gcgcgccaca	ggcgccgttg	5460
cgtggcgccc	ctgaaccgac	cgaagtgttt	catgcagcgg	tgcgtaaccg	taaagtcggt	5520
ccgggtgata	ttccgcacct	ggatgaagcc	ctgatcgaca	agtacatccg	tgacctgcgc	5580
gagttcggtc	tgatttagaa	ttccataatt	gctgttagga	gatatatatg	gcggacacgt	5640
tattgattct	gggtgatagc	ctgagcgccg	ggtatcgaat	gtctgccagc	gcggcctggc	5700
ctgccttgtt	gaatgataag	tggcagagta	aaacgtcggg	agttaatgcc	agcatcagcg	5760

gcgacacctc	gcaacaagga	ctggcgcgcc	ttccggetct	gctgaaacag	catcagccgc	5820
gttgggtgct	ggttgaactg	ggcggctgtg	acggtttgcg	tggttttcag	ccacagcaaa	5880
cogagcaaac	gctgcgccag	atthttgcagg	atgtcaaagc	cgccaacgct	cttccattgt	5940
taatgcaaat	acgtctgcct	tacaactatg	gtcgtcgta	taatgaagcc	tttagcgcca	6000
tttaccocaa	actcgccaaa	gagtttgatg	ttccgctgct	gccctttttt	atggaagagg	6060
tctgcctcaa	gccacaatgg	atgcaggatg	acggatttca	tcccaacgc	gacgcccagc	6120
cgtttattgc	cgactggatg	gcgaagcagt	tgcagccttt	aaccaatcat	gactcataag	6180
cttctaagga	aataatagga	gattgaaaat	ggcaacaact	aatgtgattc	atgcttatgc	6240
tgcaatgcag	gcagggtgaag	cactcgtgcc	ttattcgttt	gatgcaggcg	aactgcaacc	6300
acatcagggt	gaagttaaag	tcgaatattg	tgggctgtgc	cattccgatg	tctcgggtact	6360
caacaacgaa	tggcattctt	cggtttatcc	agtcgtggca	ggtcatgaag	tgattggtac	6420
gattacccaa	ctgggaagtg	aagccaaagg	actaaaaatt	ggtcaacgtg	ttggtattgg	6480
ctggacggca	gaaagctgtc	aggcctgtga	ccaatgcata	agtggtcagc	aggtattgtg	6540
cacgggcgaa	aataccgcaa	ctattattgg	tcattgctgg	ggctttgcag	ataaggttcg	6600
tgcaggctgg	caatgggtca	ttcccctgcc	cgacgaactc	gatccgacca	gtgctgggtcc	6660
tttgctgtgt	ggcggaatca	cagtatttga	tccaatttta	aaacatcaga	ttcaggctat	6720
tcattcatgtt	gctgtgattg	gtatcgggtg	tttgggacat	atggccatca	agctacttaa	6780
agcatggggc	tgtgaaatta	ctgcgtttag	ttcaaataca	aacaaaaccg	atgagctcaa	6840
agctatgggg	gccgatcacg	tggtcaatag	cogtgatgat	gccgaaatta	aatcgcaaca	6900
gggtaaatth	gatttactgc	tgagtacagt	taatgtgcct	ttaaactgga	atgcgtatct	6960
aaacacactg	gcacccaatg	gcactttcca	ttttttgggc	gtggtgatgg	aaccaatccc	7020
tgtacctgtc	ggtgcgctgc	taggaggtgc	caaatacgta	acagcatcac	caactggctc	7080
gcctgctgcc	ttacgtaagc	tgctcgaatt	tgcggcacgt	aagaatatcg	cacctcaaat	7140
cagatgttat	cctatgtcgg	agctgaatga	ggccatcgaa	cgcttacatt	cgggtcaagc	7200
acgttatcgg	attgtactta	aagccgattt	ttaacctagg	gataatagag	gttaagagcg	7260
gccagatgcc	acattcctac	gattacgatg	ccatagtaat	aggttccggc	cccggcggcg	7320
aaggcgctgc	aatgggcctg	gttaagcaag	gtgcgcgcgt	cgcagttatc	gagcgttatc	7380
aaaatgttgg	cggcggttgc	accactggg	gcaccatccc	gtcgaaagct	ctccgtcacg	7440
cogtcagccg	cattatagaa	ttcaatcaaa	accactttta	cagcgaccat	tcccgaactgc	7500
tccgctcttc	ttttgccgat	atccttaacc	atgccgataa	cgtgattaat	caacaaacgc	7560
gcatgcgtca	gggatttttac	gaacgtaatc	actgtgaaat	attgcaggga	aacgctcgct	7620
ttgttgacga	gcatacgttg	gcgctggatt	gcccgacgg	cagcgttgaa	acactaaccg	7680

ctgaaaaatt	tggtattgcc	tgcggtcttc	gtccatatca	tccaacagat	gttgatttca	7740
cccatccacg	catttacgac	agcgactcaa	ttctcagcat	gcaccacgaa	ccgcgccatg	7800
tacttatcta	tggtgctgga	gtgatcggct	gtgaatatgc	gtcgatcttc	cgcggtatgg	7860
atgtaaaagt	ggatctgatc	aacacccgcg	atcgcttgc	ggcatttctc	gatcaagaga	7920
tgtcagattc	tctctcctat	cacttctgga	acagtggcgt	agtgattcgt	cacaacgaag	7980
agtacgagaa	gatcgaaggc	tgtgacgatg	gtgtgatcat	gcattctgaag	tggggtaaaa	8040
aactgaaagc	tgactgcctg	ctctatgcc	acggtcgcac	cggtaatacc	gattcgcctg	8100
cgttacagaa	cattgggcta	gaaactgaca	gccgcggaca	gctgaaggtc	aacagcatgt	8160
atcagaccgc	acagccacac	gtttacgcgg	tgggcgacgt	gattgggtat	ccgagcctgg	8220
cgtcggcggc	ctatgaccag	gggcgcattg	ccgcgcaggc	gctggtaaaa	ggcgaagcca	8280
ccgcacatct	gattgaagat	atccctaccg	gtatttacac	catcccgga	atcagctctg	8340
tgggcaaaac	cgaacagcag	ctgaccgcaa	tgaagtgcc	atatgaagtg	ggccgcgccc	8400
agtttaaaaca	tctggcacgc	gcacaaatcg	tggcatgaa	cgtgggcacg	ctgaaaattt	8460
tgttccatcg	ggaacaaaa	gagattctgg	gtattcactg	ctttggcgag	cgcgctgccg	8520
aaattattca	tatcggtcag	gcgattatgg	aacagaaagg	tggcggcaac	actattgagt	8580
acttcgtcaa	caccaccttt	aactaccoga	cgatggcgga	agcctatcgg	gtagctgcgt	8640
taaacggttt	aaacgcctg	ttttaaaact	tatcgaaatg	gccatccatt	cttggtttaa	8700
acggtctcca	gcttggtcgt	tttgccggat	gagagaagat	tttcagcctg	atacagatta	8760
aatcagaacg	cagaagcgg	ctgataaaac	agaatttgcc	tggcggcagt	agcgcggtgg	8820
tcccacctga	cccatgccg	aactcagaag	tgaacgccg	tagcgccgat	ggtagtgtgg	8880
ggtctcccca	tgcgagagta	gggaactgcc	aggcatcaaa	taaaacgaaa	ggctcagtcg	8940
aaagactggg	cctttcgttt	tatctgttgt	ttgtcgggtga	acgctctcct	gacgcctgat	9000
gcggtatttt	ctccttacgc	atctgtgcgg	tatttcacac	cgcatatgg	gcactctcag	9060
tacaatctgc	tctgatgccg	catagttaag	ccagccccga	caccgcgcaa	caccgcctga	9120
cgagcttagt	aaagccctcg	ctagatttta	atgcggatgt	tgcgattact	tgcgcaacta	9180
ttgcgataac	aagaaaaagc	cagcctttca	tgatatact	cccaatttgt	gtagggctta	9240
ttatgcacgc	ttaaaaataa	taaaagcaga	cttgacctga	tagtttggct	gtgagcaatt	9300
atgtgcttag	tgcatctaac	gcttgagtta	agccgcgccg	cgaagcggcg	tgggcttgaa	9360
cgaattgtta	gacattattt	gccgactacc	ttggtgatct	cgcctttcac	gtagtggaca	9420
aattcttcca	actgatctgc	gcgcgaggcc	aagcgatctt	cttcttgtcc	aagataagcc	9480
tgtctagctt	caagtatgac	gggctgatac	tgggccggca	ggcgctccat	tgccagtcg	9540

gcagcgacat ccttcggcgc gattttgccg gttactgcgc tgtaccaaact gcgggacaac	9600
gtaagcacta catttcgctc atcgccagcc cagtcggggc gcgagttcca tagcgttaag	9660
gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag ttcctccgcc	9720
gctggacctt ccaaggcaac gctatgttct cttgcttttg tcagcaagat agccagatca	9780
atgtcgatcg tggctggctc gaagatacct gcaagaatgt cattgcgctg ccattctcca	9840
aattgcagtt cgcgcttagc tggataacgc cacggaatga tgtcgtcgtg cacaacaatg	9900
gtgactttta cagcgccggg aatctcgctc tctccagggg aagccgaagt ttccaaaagg	9960
tcgttgatca aagctcgccg cgttgtttca tcaagcctta cggtcaccgt aaccagcaaa	10020
tcaatatcac tgtgtggctt caggccgcca tccactgcgg agccgtacaa atgtacggcc	10080
agcaacgctg gttcgagatg gcgctcgatg acgccaacta cctctgatag ttgagtcgat	10140
acttcggcga tcaccgcttc cctcatgatg tttaactttg ttttagggcg actgccctgc	10200
tgcgtaacat cgttgctgct ccataacatc aaacatcgac ccacggcgta acgcgcttgc	10260
tgcttgatg cccgagggat agactgtacc ccaaaaaaac agtcataaca agccatgaaa	10320
accgccactg cgcgcttacc accgctgcgt tcgggtcaagg ttctggacca gttgcgtgag	10380
cgcatacgtt acttgcatta cagcttacga accgaacagg cttatgtcca ctgggttcgt	10440
gccttcacac gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgag	10500
gcattttctg cctggctggc gaacgagcgc aagggttcgg tctccacgca tcgtcaggca	10560
ttggcggcct tgctgttctt ctacggcaag gtgctgtgca cggatctgcc ctggcttcag	10620
gagatcggaa gacctcggcc gtgcgggcgc ttgcccgtgg tgctgacccc ggatgaagtg	10680
gttcgcatcc tcggttttct ggaaggcgag catcgtttgt tcgccagct tctgtatgga	10740
acgggcatgc ggatcagtga gggtttgcaa ctgcgggtca aggatctgga tttcgatcac	10800
ggcacgatca tcgtgcggga gggcaagggc tccaaggatc gggccttgat gttaccgag	10860
agcttggcac ccagcctgcg cgagcagggg aattaattcc cacgggtttt gctgcccgc	10920
aacgggctgt tctgggtgtg ctagtgtgtt atcagaatcg cagatccggc ttcagccggt	10980
ttgccggctg aaagcgctat ttcttcaga attgccatga ttttttccc acgggaggcg	11040
tcactggctc ccgtgttgtc ggcagctttg attcgataag cagcatcgcc tgtttcaggc	11100
tgtctatgtg tgactgttga gctgtaacaa gttgtctcag gtgttcaatt tcatgttcta	11160
gttgctttgt tttactggtt tcacctgttc tattaggtgt tacatgctgt tcatctgtta	11220
cattgtcgat ctgttcattg tgaacagctt tgaatgcacc aaaaactcgt aaaagctctg	11280
atgtatctat cttttttaca ccgttttcat ctgtgcatat ggacagtttt ccctttgata	11340
tgtaacgggtg aacagttgtt ctacttttgt ttgttagtgt tgatgcttca ctgatagata	11400
caagagccat aagaacctca gatccttccg tatttagcca gtatgttctc tagtgtggtt	11460

cgttggtttt gcgtgagcca tgagaacgaa ccattgagat catacttact ttgcatgtca 11520
 ctcaaaaatt ttgcctcaaa actggtgagc tgaatttttg cagttaaagc atcgtgtagt 11580
 gtttttctta gtccgttatg taggtaggaa tctgatgtaa tgggtgttggt tattttgtca 11640
 ccattcattt ttatctggtt gttctcaagt tccggttacga gatccatttg tctatctagt 11700
 tcaacttgga aaatcaacgt atcagtcggg cggcctcgct tatcaaccac caatttcata 11760
 ttgctgtaag tgttttaaate tttacttatt gggtttcaaaa cccattgggtt aagcctttta 11820
 aactcatggt agttattttc aagcattaac atgaacttaa attcatcaag gctaattctct 11880
 atatttgcct tgtgagtttt cttttgtgtt agttctttta ataaccactc ataaatcctc 11940
 atagagtatt tgttttcaaa agacttaaca tgttccagat tatattttat gaattttttt 12000
 aactggaaaa gataaggcaa tatctcttca ctaaaaacta attctaattt ttcgcttgag 12060
 aacttggcat agtttgtcca ctggaaaatc tcaaagcctt taaccaaagg attcctgatt 12120
 tccacagttc tcgtcatcag ctctctggtt gcttttagcta atacaccata agcatttttc 12180
 ctactgatgt tcatcatctg agcgtattgg ttataagtga acgataccgt ccgttctttc 12240
 cttgtagggt tttcaatcgt ggggttgagt agtgccacac agcataaaat tagcttggtt 12300
 tcatgctccg ttaagtcata gcgactaatc gctagttcat ttgctttgaa aacaactaat 12360
 tcagacatac atctcaattg gtctaggtga ttttaat 12397

<210> 14
 <211> 3227
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Polinucleótido sintético"

<400> 14

gacgaaaggg cctcgtgata cgcctatttt tatagggttaa tgtcatgata ataatggttt 60
 cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatttt 120
 tctaaataca ttcaaataatg tatccgctca tgagacaata accctgataa atgcttcaat 180
 aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt 240
 ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 300
 ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 360
 tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc 420
 tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatac 480
 actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 540

gcatgacagt	aagagaatta	tgcagtgtctg	ccataaccat	gagtgataac	actgcggcca	600
acttacttct	gacaacgata	ggaggaccga	aggagctaac	cgcttttttg	cacaacatgg	660
gggatcatgt	aactcgcctt	gatcgttggg	aaccggagct	gaatgaagcc	ataccaaacg	720
acgagcgtga	caccacgatg	cctgtagcaa	tggcaacaac	gttgcgcaaa	ctattaactg	780
gcgaactact	tactctagct	tcccggcaac	aattaataga	ctggatggag	gcggataaag	840
ttgcaggacc	acttctgcgc	tggcccttc	cggctggctg	gtttattgct	gataaatctg	900
gagccggtga	gcgtgggtct	cgcggtatca	ttgcagcact	ggggccagat	ggtaagccct	960
cccgtatcgt	agttatctac	acgacgggga	gtcaggcaac	tatggatgaa	cgaaatagac	1020
agatcgtctga	gatagggtgcc	tcactgatta	agcattggta	actgtcagac	caagtttact	1080
catatatact	ttagattgat	ttaaaacttc	atttttaatt	tgtgcatccg	aagatcagca	1140
gttcaacctg	ttgatagtac	gtactaagct	ctcatgtttc	acgtactaag	ctctcatgtt	1200
taacgtacta	agctctcatg	tttaacgaac	taaaccctca	tggctaacgt	actaagctct	1260
catggctaac	gtactaagct	ctcatgtttg	aacaataaaa	ttaatataaa	tcagcaactt	1320
aaatagcctc	taaggtttta	agttttataa	gaaaaaaaaa	aatatataag	gctttttaaag	1380
ctagctttta	aggtttcacc	atgttctttc	ctgcgttata	ccctgattct	gtggataacc	1440
gtattaccgc	ctttgagtga	gctgataccg	ctcgccgcag	ccgaacgacc	gagcgcagcg	1500
agtcagttag	cgaggaagcg	gaagagcgcc	caatacgcaa	accgcctctc	cccgcgcgtt	1560
ggccgattca	ttaagacagc	tgtctcttat	acacatctca	accctgaagc	tcttggtggc	1620
tagtgcgtag	tcgttggtgaa	gctttccgct	gtttctgcat	tcttacgttt	taggatgcat	1680
atggcggccg	cataacttgc	tatagcatac	attatacgaa	gttatctaga	gttgcatgcc	1740
tgcaggtccg	cttattatca	cttattcagg	cgtagcaacc	aggcgtttta	gggcaccaat	1800
aactgcctta	aaaaaattac	gccccgcctt	gccactcacc	gcagtactgt	tgtaattcat	1860
taagcattct	gccgacatgg	aagccatcac	aaacggcatg	atgaacctga	atcgccagcg	1920
gcatcagcac	cttgtcgcct	tgcgtataat	atttgcccat	ggtgaaaacg	ggggcgaaga	1980
agttgtccat	attggccacg	tttaaatcaa	aactggtgaa	actcaccacg	ggattggctg	2040
agacgaaaaa	catattctca	ataaaccttc	tagggaaata	ggccaggttt	tcaccgtaac	2100
acgccacatc	ttgcgaatat	atgtgtagaa	actgccggaa	atcgtcgtgg	tattcactcc	2160
agagcgatga	aaacgtttca	gtttgctcat	ggaaaacggt	gtaacaaggg	tgaacactat	2220
cccatatcac	cagctcaccg	tctttcattg	ccatacggaa	ttccggatga	gcattcatca	2280
ggcgggcaag	aatgtgaata	aaggccggat	aaaacttgtg	cttatttttc	tttacggctc	2340
ttaaaaaggc	cgtaatatcc	agctgaacgg	tctggttata	ggtacattga	gcaactgact	2400
gaaatgcctc	aaaatgttct	ttacgatgcc	attgggatat	atcaacgggtg	gtatatccag	2460

tgattttttt	ctccatttta	gcttccttag	ctcctgaaaa	tctcgataac	tcaaaaaata	2520
cgcccggtag	tgatcttatt	tcattatggg	gaaagttgga	acctcttacg	tgccgatcaa	2580
cgtctcattt	tcgcaaaaag	ttggcccagg	gcttcccggg	atcaacaggg	acaccaggat	2640
ttattttatt	tgcgaaagtga	tcttcctgca	caggtattta	ttcgactcta	gataacttcg	2700
tatagcatat	attatacgaa	ggtatggatc	cagcttatcg	ataccgtcaa	acaaatcata	2760
aaaaatttat	ttgcttttcag	gaaaattttt	ctgtataata	gattcaattg	cgatgacgac	2820
gaacacgcat	taaggagggtg	aagagctcga	attcgagcca	atatgcgaga	acacccgaga	2880
aaattcatcg	atgatgggtg	agatgtgtat	aagagacagc	tgtcgtaata	gcgaagaggc	2940
ccgcaccgat	cgcccttccc	aacagttgcg	cagcctgaat	ggcgaatggc	gcctgatgcg	3000
gtattttctc	cttacgcata	tgtgoggtat	ttcacaccgc	atatggtgca	ctctcagtag	3060
aatctgctct	gatgccgcat	agttaagcca	gccccgacac	ccgccaacac	ccgctgacgc	3120
gccctgacgg	gcttgtctgc	tcccggcatc	cgcttacaga	caagctgtga	ccgtctccgg	3180
gagctgcatg	tgtcagaggt	tttcaccgtc	atcaccgaaa	cgcgcgga		3227

<210> 15
 <211> 40
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

<400> 15
 gcagttattg gtgcccttaa acgcctggtt gctacgcctg 40

<210> 16
 <211> 34
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

<400> 16
 cccagggtt cccggtatca acaggacac cagg 34

<210> 17
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

<400> 17
 atggtcatta aggcgcaaag cccgg 25

<210> 18

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

5

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

10

<400> 18
 gagacccac actaccatcc tcgagttatc gccctgaat ggctaaatca ccc 53

15

<210> 19
 <211> 27
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

25

<400> 19
 ctcgaggatg gtagtgtggg gtctccc 27

30

<210> 20
 <211> 69
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

<400> 20
 gagaccgttt ctccaattta aatatgatac gctcgagctt cgtctgtttc tactgggtatt 60
 ggcacaaac 69

40

<210> 21
 <211> 65
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <221> fuente
 <223> /nota= "Descripción de la secuencia artificial: Cebador sintético"

50

<220>
 <221> base_modificada
 <222> (16)..(16)
 <223> a, c, t, g, desconocida u otra

55

<220>
 <221> base_modificada
 <222> (21)..(21)
 <223> a, c, t, g, desconocida u otra

60

<220>
 <221> base_modificada
 <222> (34)..(40)
 <223> a, c, t, g, desconocida u otra

<400> 21

tgaaagatta aatttnhhar nddhddnwag gagnnnnnnn atggtcatta aggcgcaaag 60

cccgg 65

5

<210> 22

<211> 12206

<212> ADN

<213> Secuencia artificial

10

<220>

<221> fuente

<223> /nota= "Descripción de la secuencia artificial: Polinucleótido sintético"

15

<400> 22

cactatacca attgagatgg gctagtcaat gataattact agtcottttc ctttgagttg 60

tgggtatctg taaattctgc tagacctttg ctggaaaact tgtaaattct gctagacct 120

ctgtaaattc cgctagacct ttgtgtgttt tttttgttta tattcaagtg gttataattt 180

atagaataaa gaaagaataa aaaaagataa aaagaataga tcccagccct gtgtataact 240

cactacttta gtcagttccg cagtattaca aaaggatgtc gcaaacgctg tttgctcctc 300

tacaaaacag accttaaaac cctaaaggcg tcggcatccg cttacagaca agctgtgacc 360

gtctccggga gctgcatgtg tcagagggtt tcacogtcat caccgaaacg cgcgaggcag 420

cagatcaatt cgcgcgcgaa ggcgaagcgg catgcattta cgttgacacc atcgaatggt 480

gcaaaacctt tcgcggtatg gcatgatagc gcccggaaga gagtcaattc aggggtggtga 540

atgtgaaacc agtaacgtta tacgatgtcg cagagtatgc cgggtgtctct tatcagaccg 600

tttcccgcgt ggtgaaccag gccagccacg tttctgcgaa aacgcgggaa aaagtgggaag 660

cggcgatggc ggagctgaat tacattccca accgcgtggc acaacaactg gcggggcaaac 720

agtcgttgct gattggcggt gccacctcca gtctggccct gcacgcgccg tcgcaaattg 780

tcgcggcgat taaatctcgc gccgatcaac tgggtgccag cgtggtggtg tcgatggtag 840

aacgaagcgg	cgtcgaagcc	tgtaaagcgg	cgggtgcacaa	tcttctcgcg	caacgcgtca	900
gtgggctgat	cattaactat	ccgctggatg	accaggatgc	cattgctgtg	gaagctgcct	960
gcactaatgt	tccggcggtta	tttcttgatg	tctctgacca	gacacccatc	aacagtatta	1020
ttttctccca	tgaagacggt	acgcgactgg	gcgtggagca	tctggtcgca	ttgggtcacc	1080
agcaaatacgc	gctgttagcg	ggcccatata	gttctgtctc	ggcgcgtctg	cgtctggctg	1140
gctggcataa	atatctcact	cgcaatcaaa	ttcagccgat	agcggaacgg	gaaggcgact	1200
ggagtgccat	gtccggtttt	caacaaacca	tgcaaagtct	gaatgagggc	atcgttccca	1260
ctgcgatgct	ggttgccaac	gatcagatgg	cgtcgggcgc	aatgcgcgcc	attaccgagt	1320
ccgggctgcg	cgttggtgcg	gatctctcgg	tagtgggata	cgacgatacc	gaagacagct	1380
catgttatat	cccgcggtta	accaccatca	aacaggattt	tgcctgctg	gggcaaacca	1440
gcgtggaccg	cttgcctgca	ctctctcagg	gccaggcggg	gaagggcaat	cagctgttgc	1500
ccgtctcact	ggtgaaaaga	aaaaccaccc	tggcgcccaa	tacgcaaacc	gcctctcccc	1560
gcgcgttggc	cgattcatta	atgcagctgg	cacgacaggt	ttcccgactg	gaaagcgggc	1620
agtgagcgca	acgcaattaa	tgtaagttag	cgcgaattga	tctggtttga	cagcttatca	1680
tgcactgcac	ggtgcaccaa	tgcttctggc	gtcaggcagc	catcggaagc	tgtggtatgg	1740
ctgtgcaggt	cgtaaatac	tgcataattc	gtgtcgcctc	aggcgcactc	ccgttctgga	1800
taatgttttt	tgcgcgcgaca	tcataacggg	tctggcaaat	attctgaaat	gagctgttga	1860
caattaatca	tccggctcgt	ataatgtgtg	gaattgtgag	cggataacaa	ttcacacag	1920
gaaacagcgc	cgtgagaaa	aagcgaagcg	gcactgctct	ttaacaattt	atcagacaat	1980
ctgtgtgggc	actcgaccgg	aattatcgat	taactttatt	attaaaaatt	aaagaggtat	2040
atattaatgt	atcgattaaa	taaggaggaa	taaaccatga	cgagcgatgt	tcacgacgcg	2100
accgacggcg	ttaccgagac	tgactggat	gatgagcaga	gcactcgtcg	tattgcagaa	2160
ctgtacgcaa	cggacccaga	gttcgcagca	gcagctcctc	tgccggccgt	tgctgatgcg	2220
gcgcacaaac	cgggcctgcg	tctggcggaa	atcctgcaga	ccctgttcac	cggctacggc	2280
gatcgtccgg	cgtcgggcta	tctgtcacgt	gagctggcga	cggacgaagg	cggctgtacg	2340
gtcacgcgtc	tgctgccgcg	cttcgatacc	ctgacctatg	cacaggtgtg	gagccgtgtt	2400
caagcagtgg	ctgcagcgtt	gcgtcacaa	ttcgcacaa	cgatttacc	gggcgacgcg	2460
gtcgcgacta	tccgctttgc	gagcccgac	tatttgacgc	tggatctggg	gtgcgcgtat	2520
ctgggcctgg	tcagcgttcc	tttgacgat	aacgctccgg	tgtctcgct	ggccccgatt	2580
ctggccgagg	tggaaccgcg	tattctgacg	gtgagcgag	aatacctgga	cctggcggtt	2640
gaatccgtcc	gtgatgtgaa	ctccgtcagc	cagctggttg	ttttcgacca	tcacccggaa	2700

gtggacgatc	accgtgacgc	actggctcgc	gcacgcgagc	agctggccgg	caaaggtatc	2760
gcagttacga	ccctggatgc	gatcgagac	gaaggcgcag	gtttgccggc	tgagccgatt	2820
tacacggcgg	atcacgatca	gcgtctggcc	atgattctgt	ataccagcgg	ctctacgggt	2880
gctccgaaaag	gcgcgatgta	caccgaagcg	atggtggctc	gcctgtggac	tatgagcttt	2940
atcacgggcg	acccgacccc	ggttatcaac	gtgaacttca	tgccgctgaa	ccatctgggc	3000
ggtcgtatcc	cgattagcac	cgccgtgcag	aatggcggta	ccagctactt	cgttccggaa	3060
agcgacatga	gcacgctggt	tgaggatctg	gccctggtcc	gccctaccga	actgggtctg	3120
gtgccgcgtg	ttgcggacat	gctgtaccag	catcatctgg	cgaccgtgga	tcgcctggtg	3180
accagggcg	cggacgaact	gactgcggaa	aagcaggccg	gtgcggaact	gcgtgaacag	3240
gtcttgggcg	gtcgtgttat	caccggtttt	gtttccaccg	cgccgttggc	ggcagagatg	3300
cgtgcttttc	tggatatcac	cttgggtgca	cacatcgttg	acggttacgg	tctgaccgaa	3360
accggtgcgg	tcaccctgta	tggtgtgatt	gttcgtcctc	cggtcattga	ttacaagctg	3420
atcgatgtgc	cggagctggg	ttacttctcc	accgacaaac	cgtaccgcgg	tggcgagctg	3480
ctggttcgta	gccaaacggt	gactccgggt	tactacaagc	gccagaagt	caccgcgtcc	3540
gttttcgatc	gcgacggcta	ttaccacacc	ggcgacgtga	tggcagaaac	cgcgccagac	3600
cacctggtgt	atgtggaccg	cgcacaacat	gttctgaagc	tggcgcaagg	tgaatttgtc	3660
gccgtggcta	acctggaggc	cgttttcagc	ggcgtgctc	tggtccgcca	gattttcgtg	3720
tatggtaaca	gcgagcgcag	ctttctgttg	gctgttggtg	tccttaccct	ggaggcgtg	3780
gagcaatacg	accctgccgc	attgaaagca	gccctggcgg	attcgctgca	gcgtacggcg	3840
cgtgatgccg	agctgcagag	ctatgaagtg	ccggcggact	tcattgttga	gactgagcct	3900
tttagcgtg	cgaacggctc	gctgagcgg	gttggcaagt	tgctgcgtcc	gaatttgaag	3960
gatcgctacg	gtcagcgttt	ggagcagatg	tacgcggaca	tcgcggctac	gcaggcgaac	4020
caattgcgtg	aactgcgcgc	tgctgcggct	actcaaccgg	tgatcgacac	gctgacgcaa	4080
gctgcggcga	ccatcctggg	taccggcagc	gaggttgcaa	gcgacgcaca	ctttactgat	4140
ttgggcggtg	attctctgag	cgcgtgacg	ttgagcaact	tgctgtctga	cttctttggc	4200
tttgaagtcc	cggttggcac	gattgttaac	ccagcgacta	atctggcaca	gctggcgcaa	4260
catatcgagg	cgcagcgcac	ggcgggtgac	cgcctgcat	cctttacgac	ggtccacggg	4320
gcggatgcta	cggaaatccg	tgcaagcgaa	ctgactctgg	acaaattcat	cgacgctgag	4380
actctgcgcg	cagcacctgg	tttgccgaag	gttacgactg	agccgcgtac	ggtcctgttg	4440
agcggtgcca	atggttggtt	gggcccgttc	ctgaccctgc	agtggctgga	acgtttggca	4500
ccggttggcg	gtaccctgat	caccattgtg	cgcggtcgtg	acgatgcagc	ggcacgtgca	4560
cgtttgactc	aggettacga	tacggaccca	gagctgtccc	gccgcttcgc	tgagttggcg	4620

gategccact	tgcgtgtggt	ggcaggtgat	atcggcgatc	cgaatctggg	cctgaccccc	4680
gagatttggc	accgtctggc	agcagaggtc	gatctggtcg	ttcatccagc	ggccctggtc	4740
aaccacgtcc	tgccgtaccg	ccagctgttt	ggcccgaaatg	ttgttggcac	cgccgaagtt	4800
atcaagttgg	ctctgaccga	gcgcatcaag	cctgttacct	acctgtccac	ggttagcgtc	4860
gcgatgggta	ttcctgattt	tgaggaggac	ggtgacattc	gtaccgtcag	cccggttcgt	4920
ccgctggatg	gtggctatgc	aaatggctat	ggcaacagca	agtgggctgg	cgaggtgctg	4980
ctgcgcgagg	cacatgacct	gtgtggcctg	ccggttgcca	cgtttcgtag	cgacatgatt	5040
ctggcccacc	cgcgctaccg	tggccaagtg	aatgtgccgg	acatgttcac	ccgtctgctg	5100
ctgtccctgc	tgatcacggg	tgtggcaccg	cgttccttct	acattggtga	tggcgagcgt	5160
ccgctgcac	actaccggg	cctgaccgtc	gattttgttg	cggaagcgg	tactaccctg	5220
ggtgctcagc	aacgtgagg	ttatgtctcg	tatgacgtta	tgaatccgca	cgatgacgg	5280
attagcttgg	atgtctttgt	ggactggctg	attcgtgcgg	gccacccaat	tgaccgtgtt	5340
gacgactatg	atgactgggt	gcgtcgtttt	gaaaccgcgt	tgaccgcctt	gccggagaaa	5400
cgctgtgcgc	agaccgttct	gccgctgctg	catgcctttc	gcgcgccaca	ggcgccgttg	5460
cgtggcgccc	ctgaaccgac	cgaagtgttt	catgcagcgg	tgcgtaccgc	taaagtgcgt	5520
ccgggtgata	ttccgcacct	ggatgaagcc	ctgatcgaca	agtacatccg	tgacctgcgc	5580
gagttcggtc	tgatttagaa	ttcttttagcg	ttaagaagga	gatatatatg	gcggacacgt	5640
tattgattct	gggtgatagc	ctgagcgccg	ggtatcgaat	gtctgccagc	gcggcctggc	5700
ctgccttggt	gaatgataag	tggcagagta	aaacgtcgg	agttaatgcc	agcatcagcg	5760
gcgacacctc	gcaacaagga	ctggcgcgcc	ttccggtctc	gctgaaacag	catcagccgc	5820
gttgggtgct	ggttgaactg	ggcggcaatg	acggtttgcg	tggttttcag	ccacagcaaa	5880
ccgagcaaac	gctgcgccag	atgttcagag	atgtcaaagc	cgccaacgct	gaaccattgt	5940
taatgcaaat	acgtctgcct	tacaactatg	gtcgtcggtta	taatgaagcc	tttagcgcca	6000
tttaccctaa	actcgccaaa	gagtttgatg	ttccgctgct	gccctttttt	atggaagagg	6060
tctgcctcaa	gccacaatgg	atgcaggatg	acggatttca	tcccaaccgc	gacgcccagc	6120
cgtttattgc	cgactggatg	gcgaagcagt	tgacgccttt	agtaaatacat	gactcataag	6180
cttctaagga	aataatagga	gattgaaaat	ggcaacaact	aatgtgattc	atgcttatgc	6240
tgcaatgcag	gcagggtgaag	cactcgtgcc	ttattcgttt	gatgcaggcg	aactgcaacc	6300
acatcagggt	gaagttaaag	tcgaatattg	tggtgtgtgc	cattccgatg	tctcggtact	6360
caacaacgaa	tggcattctt	cggtttatcc	agtcgtggca	ggtcatgaag	tgattggtac	6420
gattaccctaa	ctgggaagtg	aagccaaagg	actaaaaatt	ggtcaacgtg	ttggtattgg	6480

ctggacggca	gaaagctgtc	aggcctgtga	ccaatgcac	agtggtcagc	aggtattgtg	6540
cacgggcgaa	aataccgcaa	ctattattgg	tcatgctggg	ggctttgcag	ataaggttcg	6600
tgcaggctgg	caatgggtca	ttcccctgcc	cgacgaactc	gatccgacca	gtgctgggtcc	6660
tttgctgtgt	ggcggaatca	cagtatttga	tccaatttta	aaacatcaga	ttcaggctat	6720
tcatcatgtt	gctgtgattg	gtatcgggtg	tttgggacat	atggccatca	agctacttaa	6780
agcatggggc	tgtgaaatta	ctgcgtttag	ttcaaatacca	aacaaaaccg	atgagctcaa	6840
agctatgggg	gccgatcacg	tgggtcaatag	ccgtgatgat	gccgaaatta	aatcgcaaca	6900
gggtaaattt	gatttactgc	tgagtacagt	taatgtgcct	ttaaactgga	atgcgtatct	6960
aaacacactg	gcacccaatg	gcactttcca	ttttttgggc	gtgggtgatgg	aaccaatccc	7020
tgtacctgtc	ggtgcgctgc	taggaggtgc	caaataccta	acagcatcac	caactggctc	7080
gcctgctgcc	ttacgtaagc	tgctcgaatt	tgcggcacgt	aagaatatcg	cacctcaa	7140
cgagatgtat	cctatgtcgg	agctgaatga	ggccatcgaa	cgcttacatt	cgggtcaagc	7200
acgttatcgg	attgtactta	aagccgattt	ttaacctagg	atcagtatct	ggtaggagat	7260
cacggatgaa	acgtgcagtg	attactggcc	tgggcattgt	ttccagcatc	ggtaataa	7320
agcaggaagt	cctggcatct	ctgcgtgaag	gacgttcagg	gatcactttc	tctcaggagc	7380
tgaaggattc	cggcatgcgt	agccacgtct	ggggcaacgt	aaaactggat	accactggcc	7440
tcattgaccg	caaagttgtg	cgctttatga	gcgacgcac	catttatgca	ttcctttcta	7500
tggagcaggc	aatcgctgat	gcgggcctct	ctccggaagc	ttaccagaat	aacccgcgcg	7560
ttggcctgat	tgcaggttcc	ggcggcggtc	ccccgcgttt	ccagggtgtc	ggcgtgacg	7620
caatgcgcgg	ccgcgcgggc	ctgaaagcgg	ttggcccgta	tgtgggtcacc	aaagcgatgg	7680
catccggcgt	ttctgcctgc	ctcgccaccc	cgtttaaaat	tcattggcgtt	aactactcca	7740
tcagctccgc	gtgtgcgact	tcgcacact	gtatcggtaa	cgcagtagag	cagatccaac	7800
tgggcaaaaca	ggacatcgtg	tttgctggcg	gcggcgaaga	gctgtgctgg	gaaatggctt	7860
gcgaattcga	cgcaatgggt	gcgctgtcta	ctaaatacaa	cgacaccccg	gaaaaagcct	7920
cccgacttta	cgacgctcac	cgtgacggtt	tcgttatcgc	tggcggcggc	ggtatggtag	7980
tggttgaaga	gctggaacac	gcgctggcgc	gtgggtgctca	catctatgct	gaaatcggtg	8040
gctacggcgc	aacctctgat	ggtgcagaca	tggttgctcc	gtctggcgaa	ggcgcagtac	8100
gctgcatgaa	gatggcgatg	catggcgttg	ataccccaat	cgattacctg	aactcccacg	8160
gtacttcgac	tccggttggc	gacgtgaaag	agctggcagc	tatccgtgaa	gtgttcggcg	8220
ataagagccc	ggcgatttct	gcaaccaaag	ggatgaccgg	tcactctctg	ggcgtgctg	8280
gcgtacagga	agctatctac	tctctgctga	tgtcggaaaca	cggctttatc	gccccgagca	8340
tcaacattga	agagctggac	gagcaggctg	cgggtctgaa	catcgtgacc	gaaacgaccg	8400

atcgcgaaact gaccaccggt atgtctaaca gcttcgggctt cggcggcacc aacgccacgc	8460
tggtaatgcg caagctgaaa gattaaattt aaatgctatg tctcgagaaa cggctctccag	8520
cttggtctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa atcagaacgc	8580
agaagcggtc tgataaaaca gaatttgcct ggcgggcagta gcggggtggt cccacctgac	8640
cccatgccga actcagaagt gaaacgcgct agcgccgatg gtagtggtggg gtctcccat	8700
gcgagagtag ggaactgcc a ggcatacaat aaaacgaaag gctcagtcga aagactgggc	8760
ctttcgtttt atctgttggt tgtcggtgaa cgctctcctg acgcctgatg cggatatttc	8820
tccttacgca tctgtgcggt atttcacacc gcatatggtg cactctcagt acaatctgct	8880
ctgatgccgc atagttaagc cagccccgac acccgccaac acccgctgac gagcttagta	8940
aagccctcgc tagattttta tgcggatggt gcgattactt cgccaactat tgcgataaca	9000
agaaaaagcc agcctttcat gatatatctc ccaatttgtg tagggcttat tatgcacgct	9060
taaaaataat aaaagcagac ttgacctgat agtttggtctg tgagcaatta tgtgcttagt	9120
gcatctaacg cttgagttaa gccgcggcgc gaagcggcgt cggcttgaac gaattgttag	9180
acattatttg ccgactacct tggatgatct gcctttcacg tagtggacaa attcttccaa	9240
ctgatctgcg cgcgaggcca agcgatcttc ttcttgtcca agataagcct gtctagcttc	9300
aagtatgacg ggctgatact gggccggcag gcgctccatt gccagtcgg cagcgacatc	9360
cttcggcgcg attttgccg ttactgcgct gtaccaaagc cgggacaacg taagcactac	9420
atttcgctca tcgccagccc agtcgggccc cgagttccat agcgttaagg ttccatttag	9480
cgcctcaaat agatcctgtt caggaaccgg atcaaagagt tcctccgccc ctggacctac	9540
caaggcaacg ctatgttctc ttgcttttgt cagcaagata gccagatcaa tgtcgatcgt	9600
ggctggctcg aagataacct caagaatgtc attgcgctgc cattctccaa attgcagttc	9660
gcgcttagct ggataacgcc acggaatgat gtcgctcgtgc acaacaatgg tgacttctac	9720
agcgcggaga atctcgctct ctcaggggga agccgaagtt tccaaaaggc cgttgatcaa	9780
agctcgccgc gttgtttcat caagccttac ggtcacccgta accagcaaat caatatcact	9840
gtgtggcttc aggcggccat ccactgcgga gccgtacaaa tgtacggcca gcaacgctcg	9900
ttcgagatgg cgctcgatga cgccaactac ctctgatagt tgagtcgata cttcggcgat	9960
caccgcttcc ctcatgatgt ttaactttgt tttagggcga ctgcctgct gcgtaacatc	10020
gttgctgctc cataacatca aacatcgacc cagggcgtaa cgcgcttgct gcttggatgc	10080
ccgaggcata gactgtaccc caaaaaaaca gtcataacaa gccatgaaaa ccgccactgc	10140
gccgttacca ccgctgcgtt cgggtcaagg tctggaccag ttgcgtgagc gcatacgcta	10200
cttgcatcac agcttacgaa ccgaacaggc ttatgtccac tgggttcgtg ccttcatccg	10260

tttccacggt	gtgcgtcacc	eggcaacctt	gggcagcagc	gaagtcgagg	catttctgtc	10320
ctggctggcg	aacgagcgca	aggtttcggt	ctccacgcat	cgtcaggcat	tggcggcctt	10380
gctgtttcttc	tacggcaagg	tgtgtgtcac	ggatctgccc	tggcttcagg	agatcggaag	10440
acctcggecg	tcgcggcgct	tgcgggtggt	gctgaccccg	gatgaagtgg	ttcgcatcct	10500
cggttttctg	gaaggcgagc	atcgtttggt	cggccagctt	ctgtatggaa	cgggcatgcg	10560
gatcagttag	ggtttgcaac	tgcgggtcaa	ggatctggat	ttcgatcacg	gcacgatcat	10620
cgtgcgggag	ggcaagggct	ccaaggatcg	ggccttgatg	ttacccgaga	gcttggcacc	10680
cagcctgcgc	gagcagggga	attaattccc	acgggttttg	ctgcccga	acgggctgtt	10740
ctggtgttgc	tagtttggtt	tcagaatcgc	agatccggct	tcagccgggt	tgcgggtga	10800
aagcgctatt	tcttcagaa	ttgccatgat	ttttcccca	cgggagggct	cactggctcc	10860
cgtgttgtcg	gcagctttga	ttcgataagc	agcatcgctt	gtttcaggct	gtctatgtgt	10920
gactgttgag	ctgtaacaag	ttgtctcagg	tgttcaattt	catgttctag	ttgctttgtt	10980
ttactgggtt	cacctgttct	attaggtgtt	acatgctgtt	catctgttac	attgtcgatc	11040
tgttcatggt	gaacagcttt	gaatgcacca	aaaactcgta	aaagctctga	tgtatctatc	11100
ttttttacac	cgttttcatc	tgtgcatatg	gacagttttc	cctttgatat	gtaacgggtga	11160
acagttgttc	tacttttggt	tgttagtctt	gatgcttcac	tgatagatac	aagagccata	11220
agaacctcag	atccttccgt	atttagccag	tatgttctct	agtgtgggtc	gttgtttttg	11280
cgtgagccat	gagaacgaac	cattgagatc	atacttactt	tgcattgtcac	tcaaaaattt	11340
tgcctcaaaa	ctggtgagct	gaatttttgc	agttaaagca	tcgtgtagtg	tttttcttag	11400
tccgttatgt	aggtaggaat	ctgatgtaat	ggttgttggt	attttgtcac	cattcatttt	11460
tatctgggtg	ttctcaagtt	cggttacgag	atccatttgt	ctatctagtt	caacttgga	11520
aatcaacgta	tcagtcgggc	ggcctcgctt	atcaaccacc	aatttcatat	tgtgtgaagt	11580
gtttaaatct	ttacttattg	gtttcaaaac	ccattgggtt	agccttttaa	actcatggta	11640
gttattttca	agcattaaca	tgaacttaaa	ttcatcaagg	ctaactctta	tatttgcctt	11700
gtgagttttc	ttttgtgtta	gttcttttaa	taaccactca	taaatcctca	tagagtattt	11760
gttttcaaaa	gacttaacat	gttccagatt	atattttatg	aattttttta	actggaaaag	11820
ataaggcaat	atctcttcac	taaaaactaa	ttctaatttt	tcgcttgaga	acttggcata	11880
gtttgtccac	tggaaaatct	caaagccttt	aaccaaagga	ttcctgattt	ccacagttct	11940
cgtcatcagc	tctctgggtg	ctttagctaa	tacaccataa	gcattttccc	tactgatgtt	12000
catcatctga	gcgtattggt	tataagtga	cgataccgtc	cgttctttcc	ttgtagggtt	12060
ttcaatcgtg	gggttgagta	gtgccacaca	gcataaaatt	agcttgggtt	catgctccgt	12120
taagtcatag	cgactaatcg	ctagttcatt	tgctttgaaa	acaactaatt	cagacataca	12180

tctcaattgg tctaggtgat tttaat

12206

5 <210> 23
<211> 14893
<212> ADN
<213> Secuencia artificial

10 <220>
<221> fuente
<223> /nota= "Descripción de la secuencia artificial: Polinucleótido sintético"

<400> 23

cactatacca attgagatgg gctagtcaat gataattact agtccttttc ctttgagttg	60
tgggtatctg taaattctgc tagacctttg ctggaaaact tgtaaattct gctagacctt	120
ctgtaaattc cgctagacct ttgtgtgttt tttttgttta tattcaagtg gttataattt	180
atagaataaa gaaagaataa aaaaagataa aaagaataga tcccagccct gtgtataact	240
cactacttta gtcagttccg cagtattaca aaaggatgtc gcaaacgctg tttgctcctc	300
tacaaaacag accttaaaac cctaaaggcg tcggcatccg cttacagaca agctgtgacc	360
gtctccggga gctgcatgtg tcagaggttt tcaccgtcat caccgaaacg cgcgaggcag	420
cagatcaatt cgcgcgcgaa ggcgaagcgg catgcattta cgttgacacc atcgaatggt	480
gcaaaacctt tcgcggtatg gcatgatagc gcccggaaga gagtcaattc aggggtggtga	540
atgtgaaacc agtaacgtta tacgatgtcg cagagtatgc cgggtgtctct tatcagaccg	600
tttcccgcgt ggtgaaccag gccagccacg tttctgcgaa aacgcgggaa aaagtggaag	660
cggcgatggc ggagctgaat tacattccca accgcgtggc acaacaactg gcgggcaaac	720
agtcgttgct gattggcggt gccacctcca gtctggccct gcacgcgcgc tcgcaaattg	780
tcgcggcgat taaatctcgc gccgatcaac tgggtgccag cgtgggtggtg tcgatggtag	840
aacgaagcgg cgtcgaagcc tgtaaagcgg cgggtgcacaa tcttctcgcg caacgcgtca	900
gtgggctgat cattaactat ccgctggatg accaggatgc cattgctgtg gaagctgcct	960
gcactaatgt tccggcggtta tttcttgatg tctctgacca gacaccatc aacagtatta	1020
ttttctccca tgaagacggt acgcgactgg gcgtggagca tctggctcgca ttgggtcacc	1080
agcaaatcgc gctgttagcg ggcccattaa gttctgtctc ggcgcgctctg cgtctggctg	1140
gctggcataa atatctcact cgcaatcaaa ttcagccgat agcgggaacgg gaaggcgact	1200
ggagtgccat gtccggtttt caacaaacca tgcaaagtct gaatgagggc atcgttccca	1260
ctgcgatgct ggttgccaac gatcagatgg cgctgggcgc aatgcgcgcc attaccgagt	1320
ccgggctgcg cgttggtgcg gatatctcgg tagtgggata cgacgatacc gaagacagct	1380
catgttatat cccgcggtta accaccatca aacaggattt tcgcctgctg gggcaaacca	1440

gcgtggaccg	cttgcctgcaa	ctctctcagg	gccaggcggt	gaagggcaat	cagctgttgc	1500
ccgtctcact	ggtgaaaaga	aaaaccaccc	tggcgcccaa	tacgcaaacc	gcctctcccc	1560
gcgcgttggc	cgattcatta	atgcagctgg	cacgacaggt	ttcccgactg	gaaagcgggc	1620
agtgagcgca	acgcaattaa	tgtaagttag	cgcgaattga	tctggtttga	cagcttatca	1680
tgcactgcac	ggtgcaccaa	tgcttctggc	gtcaggcagc	catcggaagc	tgtggtatgg	1740
ctgtgcaggt	cgtaaatcac	tgcataattc	gtgtcgctca	aggcgcactc	ccgttctgga	1800
taatgttttt	tgcccgaca	tcataacggt	tctggcaaatt	attctgaaat	gagctgttga	1860
caattaatca	tccggctcgt	ataatgtgtg	gaattgtgag	cggataacaa	tttcacacag	1920
gaaacagcgc	cgctgagaaa	aagcgaagcg	gcactgctct	ttaacaattt	atcagacaat	1980
ctgtgtgggc	actcgaccgg	aattatcgat	taactttatt	attaaaaatt	aaagaggtat	2040
atattaatgt	atcgattaaa	taaggaggaa	taaaccatga	cgagcgatgt	tcacgacgcg	2100
accgacggcg	ttaccgagac	tgcactggat	gatgagcaga	gcactcgctc	tattgcagaa	2160
ctgtacgcaa	cggaccacga	gttcgcagca	gcagctcctc	tgcggccgt	tgtcgatgcg	2220
gcgcacaaac	cgggcctcgc	tctggcgga	atcctgcaga	ccctgttcac	cggctacggc	2280
gatcgctccg	cgctgggcta	tctgtcacgt	gagctggcga	cggacgaagg	cggctcgtag	2340
gtcacgcgtc	tgtgcccgcg	cttcgatacc	ctgacctatg	cacaggtgtg	gagccgtgtt	2400
caagcagtgg	ctgcagcgtt	gcgtcaccaat	ttcgcacaa	cgattttacc	gggcgacgcg	2460
gtcgcgacta	tccgctttgc	gagcccgac	tatttgacgc	tggatctggg	gtgcgcgtat	2520
ctgggcctgg	tcagcgttcc	tttgcagcat	aacgctccgg	tgtctgcct	ggccccgatt	2580
ctggccgagg	tggaaccgcg	tattctgacg	gtgagcgag	aataacctga	cctggcggtt	2640
gaatccgtcc	gtgatgtgaa	ctccgtcagc	cagctgggtg	ttttcgacca	tcacccgaa	2700
gtggacgatc	accgtgacgc	actggctcgc	gcacgcgagc	agctggccgg	caaaggtatc	2760
gcagttacga	ccctggatgc	gatcgacagc	gaaggcgag	gtttgcgggc	tgagccgatt	2820
tacacggcgg	atcacgatca	gcgtctggcc	atgattctgt	ataccagcgg	ctctacgggt	2880
gctccgaaag	gcgcgatgta	cacogaagcg	atggtggctc	gcctgtggac	tatgagcttt	2940
atcacggggc	acccgacccc	ggttatcaac	gtgaacttca	tgcgcgtgaa	ccatctgggc	3000
ggtcgtatcc	cgattagcac	cgcogtgag	aatggcggtg	ccagctactt	cgttccggaa	3060
agcgacatga	gcacgctgtt	tgaggatctg	gccctggctc	gccctaccga	actgggtctg	3120
gtgccgcgtg	ttgcggacat	gctgtaccag	catcatctgg	cgaacgtgga	tcgcctgggt	3180
accagggcg	cggacgaact	gactgcggaa	aagcaggccg	gtgcggaact	gcgtgaacag	3240
gtcttggggc	gtcgtgttat	caccggtttt	gtttccaccg	cgcggttggc	ggcagagatg	3300
cgtgcttttc	tggatatcac	cttgggtgca	cacatcgttg	acggttacgg	tctgaccgaa	3360

accggtgctg	tcacccgtga	tgggtgtgatt	gttcgtcctc	cggtcattga	ttacaagctg	3420
atcgatgtgc	cggagctggg	ttactttctcc	accgacaaac	cgtacccgcg	tggcgagctg	3480
ctggttcgta	gccaaacgtt	gactccgggt	tactacaagc	gccagaagt	caccgcgtcc	3540
gttttcgatc	gcgacggcta	ttaccacacc	ggcgacgtga	tggcagaaac	cgcgccagac	3600
cacctgggtg	atgtggaccg	cgcacaacat	gttctgaagc	tggcgcaagg	tgaatttgct	3660
gccgtggcta	acctggaggc	cgttttcagc	ggcgctgtct	tggcccgcca	gattttcgtg	3720
tatggtaaca	gcgagcgag	ctttctgttg	gctgttggtg	tcctacccc	ggaggcgctg	3780
gagcaatacg	accctgccgc	attgaaagca	gccctggcgg	attcgctgca	gcgtacggcg	3840
cgtgatgccg	agctgcagag	ctatgaagtg	ccggcgact	tcattgttga	gactgagcct	3900
tttagcgctg	cgaacggctc	gctgagcggg	gttggcaagt	tgctgcgtcc	gaatttgaag	3960
gatcgctacg	gtcagcggtt	ggagcagatg	tacgcggaca	tcgcggctac	gcaggcgaac	4020
caattgcgtg	aactgcgccg	tgctgcggct	actcaaccgg	tgatcgacac	gctgacgcaa	4080
gctgcggcga	ccatcctggg	taccggcagc	gaggttgcaa	gcgacgcaca	ctttactgat	4140
ttgggcggtg	attctctgag	cgcgctgacg	ttgagcaact	tgctgtctga	cttctttggc	4200
tttgaagtcc	cggttggcac	gattgttaac	ccagcgacta	atctggcaca	gctggcgcaa	4260
catatcgagg	cgcagcgcac	ggcgggtgac	cgcctccat	cctttacgac	ggtccacggt	4320
gcggatgcta	cggaaatccg	tgcaagcgaa	ctgactctgg	acaaattcat	cgacgctgag	4380
actctgcgcg	cagcacctgg	tttgccgaag	gttacgactg	agccgcgtac	ggtcctgttg	4440
agcggtgcca	atggttggtt	gggccgcttc	ctgaccctgc	agtggctgga	acgtttggca	4500
ccggttggcg	gtaccctgat	caccattgtg	cgcggtcgtg	acgatgcagc	ggcacgtgca	4560
cgtttgactc	aggcttacga	tacggacca	gagctgtccc	gccgcttcgc	tgagttggcg	4620
gatcgccact	tgctgtgtgt	ggcaggtgat	atcggcgatc	cgaatctggg	cctgaccccc	4680
gagatttggc	accgtctggc	agcagaggtc	gatctggctg	ttcatccagc	ggccctggtc	4740
aaccacgtcc	tgccgtaccg	ccagctgttt	ggtccgaatg	ttgttggcac	cgccgaagtt	4800
atcaagttgg	ctctgaccga	gcgcatcaag	cctgttacct	acctgtccac	ggttagcgct	4860
gcgatgggta	ttcctgattt	tgaggaggac	ggtgacattc	gtaccgtcag	cccggttcgt	4920
ccgctggatg	gtggctatgc	aaatggctat	ggcaacagca	agtgggctgg	cgaggtgctg	4980
ctgcgcgagg	cacatgacct	gtgtggcctg	ccggttgcca	cgtttcgtag	cgacatgatt	5040
ctggcccacc	cgcgtaccg	tggccaagtg	aatgtgccgg	acatgttcac	ccgtctgctg	5100
ctgtccctgc	tgatcacggg	tgtggcaccg	cgttccttct	acattggtga	tggcgagcgt	5160
ccgcgtgcac	actaccggg	cctgaccgtc	gattttgttg	cggaagcggg	tactaccctg	5220

ggtgctcagc aacgtgaggg ttatgtctcg tatgacgtta tgaatccgca cgatgacggg 5280
 attagcttgg atgtctttgt ggactggctg attcgtgcgg gccacccaat tgaccgtggt 5340
 gacgactatg atgactgggt gcgtcgtttt gaaaccgcgt tgaccgcctt gccggagaaa 5400
 cgctcgtgcgc agaccgttct gccgctgctg catgcctttc gcgcgccaca ggcgccgttg 5460
 cgtggcgccc ctgaaccgac cgaagtgttt catgcagcgg tgcgtaccgc taaagtccgt 5520
 ccgggtgata ttccgcacct ggatgaagcc ctgatcgaca agtacatccg tgacctgcgc 5580
 gagttcgggtc tgatttagaa ttcttttagcg ttaagaagga gatatatatg goggacacgt 5640
 tattgattct ggggtgatagc ctgagcgccg ggtatcgaat gtctgccagc ggggcctggc 5700
 ctgccttggt gaatgataag tggcagagta aaacgtcggg agttaatgcc agcatcagcg 5760
 ggcacacctc gcaacaagga ctggcgcgcc ttccggctct gctgaaacag catcagccgc 5820
 gttgggtgct ggttgaactg ggccgcaatg acggtttgcg tggttttcag ccacagcaaa 5880
 ccgagcaaac gctgcgccag attttgagg atgtcaaagc cgcacaacgt gaaccattgt 5940
 taatgcaaat acgtctgcct tacaactatg gtcgtcgta taatgaagcc tttagcgcca 6000
 tttaccccaa actcgccaaa gagtttgatg ttccgctgct gccctttttt atggaagagg 6060
 tctgcctcaa gccacaatgg atgcaggatg acggatttca tcccaaccgc gacgccagc 6120
 cgtttattgc cgactggatg gcgaagcagt tgcagccttt agtaaatcat gactcataag 6180
 cttctaagga aataatagga gattgaaaat ggcaacaact aatgtgattc atgcttatgc 6240
 tgcaatgcag gcagggtgaag cactcgtgcc ttattcgttt gatgcaggcg aactgcaacc 6300
 acatcagggt gaagttaaag tcgaatattg tgggctgtgc cattccgatg tctcggtact 6360
 caacaacgaa tggcattctt cggtttatcc agtcgtggca ggtcatgaag tgattggtac 6420
 gattacccaa ctgggaagtg aagccaaagg actaaaaatt ggtcaacgtg ttggtattgg 6480
 ctggacggca gaaagctgtc aggcctgtga ccaatgcac agtggtcagc aggtattgtg 6540
 cacgggcgaa aataccgcaa ctattattgg tcatgctggt ggctttgcag ataaggttcg 6600
 tgcaggctgg caatgggtca ttccctgcc cgacgaactc gatccgacca gtgctggtcc 6660
 tttgctgtgt ggcggaatca cagtatttga tccaatttta aaacatcaga ttcaggctat 6720
 tcatcatggt gctgtgattg gtatcgggtg tttgggacat atggccatca agctacttaa 6780
 agcatggggc tgtgaaatta ctgcgttttag ttcaaatcca aacaaaaccg atgagctcaa 6840
 agctatgggg gccgatcacg tggatcaatag ccgtgatgat gccgaaatta aatcgcaaca 6900
 gggtaaatat gatttactgc tgagtacagt taatgtgcct ttaaactgga atgcgtatct 6960
 aaacacactg gcacccaatg gcactttcca ttttttgggc gtggtgatgg aaccaatccc 7020
 tgtacctgtc ggtgcgctgc taggaggtgc caaatcgcta acagcatcac caactggctc 7080
 gcctgctgcc ttacgtaagc tgctcgaatt tgcggcacgt aagaatatcg cacctcaaat 7140

cgagatgtat	cctatgtcgg	agctgaatga	ggccatcgaa	cgcttacatt	cggttcaagc	7200
acgttatcgg	attgtactta	aagccgattt	ttaacctagg	atcagtatct	ggtaggagat	7260
cacggatgaa	acgtgcagt	attactggcc	tgggcattgt	ttccagcatc	ggtaataacc	7320
agcaggaagt	cctggcatct	ctgcgtgaag	gacgttcagg	gatcactttc	tctcaggagc	7380
tgaaggattc	cggcattcgt	agccacgtct	ggggcaacgt	aaaactggat	accactggcc	7440
tcattgaccg	caaagttgtg	cgctttatga	gcgacgcac	catttatgca	ttcctttcta	7500
tggagcaggc	aatcgctgat	gcgggcctct	ctccggaagc	ttaccagaat	aaccgcgcgc	7560
ttggcctgat	tgcaggttcc	ggcggcggct	ccccgcgttt	ccaggtgttc	ggcgctgacg	7620
caatgcgcgc	ccgcgcgcgc	ctgaaagcgc	ttggcccgtg	tgtgggtcacc	aaagcgatgg	7680
catccggcgt	ttctgcctgc	ctcgccaccc	cgttttaaata	tcattggcgtt	aactactcca	7740
tcagctccgc	gtgtgcgact	tccgcacact	gtatcggtaa	cgacagtagag	cagatccaac	7800
tgggcaaaca	ggacatcgtg	tttgctggcg	gcggcgaaga	gctgtgctgg	gaaatggctt	7860
gcgaattcga	cgcaatgggt	gcgctgtcta	ctaaatacaa	cgacaccccg	gaaaaagcct	7920
cccgtactta	cgacgctcac	cgtgacggtt	tcgttatcgc	tggcggcggc	ggtatggtag	7980
tggttgaaga	gctggaacac	gcgctggcgc	gtggtgctca	catctatgct	gaaatcggtg	8040
gctacggcgc	aacctctgat	ggtgcagaca	tggttgctcc	gtctggcgaa	ggcgacgtac	8100
gctgcatgaa	gatggcgatg	catggcgttg	ataccccaat	cgattacctg	aactcccacg	8160
gtacttcgac	tccggttggc	gacgtgaaag	agctggcagc	tatccgtgaa	gtgttcggcg	8220
ataagagccc	ggcgatttct	gcaaccaaag	ggatgaccgg	tactctctg	ggcgctgctg	8280
gcgtacagga	agctatctac	tctctgctga	tgctggaaca	cggtttatc	gccccgagca	8340
tcaacattga	agagctggac	gagcaggctg	cggtctgaa	catcgtgacc	gaaacgaccg	8400
atcgcgaaact	gaccaccgtt	atgtctaaca	gcttcggctt	cggcggcacc	aacgccacgc	8460
tggtaatgcg	caagctgaaa	gattaaattt	ctaaaattca	gttaggaggt	atttgatggt	8520
cattaaggcg	caaagcccgc	cggttttcgc	ggaagagtac	attattgaaa	gtatctggaa	8580
taaccgcttc	cctcccggga	ctattttgcc	cgcagaacgt	gaactttcag	aattaattgg	8640
cgtaacgcgt	actacgttac	gtgaagtgtt	acagcgtctg	gcacgagatg	gctggttgac	8700
cattcaacat	ggcaagccga	cgaaggtgaa	taattttctg	gaaacttccg	gtttaaatat	8760
ccttgaaaca	ctggcgcgac	tggatcacga	aagtgtgccg	cagcttattg	ataatttgct	8820
gtcgggtcgt	accaatattt	ccactatttt	tattcgcacc	gcgtttcgtc	agcatcccga	8880
taaagcgcag	gaagtgctgg	ctaccgctaa	tgaagtggcc	gatcacgccg	atgcctttgc	8940
cgagctggat	tacaacatat	tccgcggcct	ggcgtttgct	tccggcaacc	cgatttacgg	9000

tctgattctt	aacgggatga	aagggctgta	tacgcgtatt	ggtcgtcact	atttcgccaa	9060
tccggaagcg	cgcagtctgg	cgctgggctt	ctaccacaaa	ctgtcggcgt	tgtgcagtga	9120
agggcgcgac	gatcaggtgt	acgaaacagt	gcgtcgctat	gggcatgaga	gtggcgagat	9180
ttggcaccgg	atgcagaaaa	atctgccggg	tgatttagcc	attcaggggc	gataactcga	9240
ggatggtagt	gtggggtctc	cccatgcgag	agtagggaac	tgccaggcat	caaataaaac	9300
gaaaggctca	gtcgaaagac	tgggcctttc	gttttatctg	ttgtttgtcg	gtgaacgctc	9360
tcctgagtag	gacaaatccg	ccgggagcgg	atttgaacgt	tgcgaagcaa	cggcccggag	9420
ggtggcgggc	aggacgcccg	ccataaactg	ccaggcatca	aattaagcag	aaggccatcc	9480
tgacggatgg	ccttttttgcg	tggccagtgc	caagcttgca	tgcagattgc	agcattacac	9540
gtcttgagcg	attgtgtagg	ctggagctgc	ttcgaagttc	ctatactttc	tagagaatag	9600
gaacttcgga	ataggaactt	caagatcccc	ttattagaag	aactcgtcaa	gaaggcgata	9660
gaaggcgatg	cgctgcgaat	cgggagcggc	gataccgtaa	agcacgagga	agcggtcagc	9720
ccattcgccg	ccaagctctt	cagcaatatc	acgggtagcc	aacgctatgt	cctgatagcg	9780
gtccgccaca	cccagccggc	cacagtcgat	gaatccagaa	aagcggccat	ttccaccat	9840
gatattcggc	aagcaggcat	cgccatgggt	cacgacgaga	tcctcgccgt	cgggcatgcg	9900
cgccttgagc	ctggcgaaca	gttcggctgg	cgcgagcccc	tgatgctctt	cgtccagatc	9960
atcctgatcg	acaagaccgg	cttccatccg	agtaoigtgt	cgctcgatgc	gatgtttcgc	10020
ttggtggtcg	aatgggcagg	tagccggatc	aagcgtatgc	agccgcgcga	ttgcatcagc	10080
catgatggat	actttctcgg	caggagcaag	gtgagatgac	aggagatcct	gccccggcac	10140
ttcgcccaat	agcagccagt	cccttcccgc	ttcagtgaca	acgtcgagca	cagctgcgca	10200
aggaacgccc	gtcgtggcca	gccacgatag	ccgcgtgcc	tcgtcctgca	gttcattcag	10260
ggcaccggac	aggtcggctc	tgacaaaaag	aaccgggcgc	ccctgcgctg	acagccggaa	10320
cacggcggca	tcagagcagc	cgattgtctg	ttgtgcccag	tcatagccga	atagcctctc	10380
cacccaagcg	gccggagaaac	ctgcgtgcaa	tccatcttgt	tcaatcatgc	gaaacgatcc	10440
tcacctgtc	tcttgatcag	atcttgatcc	cctgcgccat	cagatccttg	gcggaagaa	10500
agccatccag	tttactttgc	agggcttccc	aaccttacca	gagggcgccc	cagctggcaa	10560
ttccggttcg	cttgcgtgcc	ataaaaccgc	ccagtctagc	tatcgccatg	taagcccact	10620
gcaagctacc	tgcctttctc	ttgcgcttgc	gttttccctt	gtccagatag	cccagtagct	10680
gacattcatc	cggggtcagc	accgtttctg	cggactggct	ttctacgtgt	tccgcttcct	10740
ttagcagccc	ttgcgccctg	agtgccttgc	gcagcgtgag	cttcaaaagc	gctctgaagt	10800
tcctatactt	tctagagaat	aggaacttcg	aactgcaggt	cgacggatcc	ccggaattaa	10860
ttctcatgtt	tgacagctta	tcaactgatca	gtgaattaat	ggcgatgacg	catcctcacg	10920

ataatatccg ggtaggcgca atcactttcg tctctactcc gttacaaagc gaggcctgggt	10980
atttcccggc ctttctgtta tccgaaatcc actgaaagca cagcggctgg ctgaggagat	11040
aaataataaa cgaggggctg tatgcacaaa gcatcttctg ttgagttaag aacgagtatc	11100
gagatggcac atagccttgc tcaaattgga atcaggtttg tgccaatacc agtagaaaca	11160
gacgaagctc gagcgtatca tatttaaatt cgagaaacgg tctccagctt ggctgttttg	11220
gcggatgaga gaagattttc agcctgatac agattaaatc agaacgcaga agcggctctga	11280
taaaacagaa tttgcctggc ggcagtagcg cgggtggctcc acctgacccc atgccgaact	11340
cagaagtga aacgcctagc gccgatggta gtgtggggctc tccccatgcg agagtaggga	11400
actgccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt tcgttttatc	11460
tgttgtttgt cgggtgaacgc tctcctgacg cctgatgcgg tattttctcc ttacgcatct	11520
gtgcggtatt tcacaccgca tatggtgcac tctcagtaca atctgctctg atgccgcata	11580
gttaagccag ccccgacacc cgccaacacc cgctgacgag cttagtaaag ccctcgctag	11640
attttaatgc ggatgttgcg attacttcgc caactattgc gataacaaga aaaagccagc	11700
ctttcatgat atatctccca atttgtgtag ggcttattat gcacgcttaa aaataataaa	11760
agcagacttg acctgatagt ttggctgtga gcaattatgt gcttagtgca tctaacgctt	11820
gagttaagcc gcgcgcgaa ggcgcgtcgg cttgaacgaa ttgtagaca ttatttgccg	11880
actacottgg tgatctcgcc ttccacgtag tggacaaatt cttccaactg atctgcgcgc	11940
gaggccaagc gatcttcttc ttgtccaaga taagcctgtc tagcttcaag tatgacgggc	12000
tgatactggg ccggcaggcg ctccattgcc cagtcggcag cgacatcctt cggcgcgatt	12060
ttgcgggta ctgcgctgta ccaaagtcgg gacaacgtaa gcactacatt tcgctcatcg	12120
ccagcccagt cgggcggcga gttccatagc gttaaggttt catttagcgc ctcaaataga	12180
tcctgttcag gaaccggatc aaagagttcc tccgcgcgtg gacctaccaa ggcaacgcta	12240
tgttctcttg cttttgtcag caagatagcc agatcaatgt cgatcgtggc tggctcgaag	12300
atacctgcaa gaatgtcatt gcgctgccat tctccaaatt gcagttcgcg cttagctgga	12360
taacgccacg gaatgatgtc gtcgtgcaca acaatggtga cttctacagc gcggagaatc	12420
tcgctctctc caggggaagc cgaagtttcc aaaaggctgt tgatcaaagc tcgccgcgtt	12480
gtttcatcaa gccttacggt caccgtaacc agcaaataca tatcactgtg tggcttcagg	12540
ccgccatcca ctgcggagcc gtacaaatgt acggccagca acgtcgggtc gagatggcgc	12600
tcgatgacgc caactacctc tgatagttga gtcgatactt cggcgatcac cgcttccctc	12660
atgatgttta actttgtttt agggcgactg ccctgctgcg taacatcgtt gctgctccat	12720
aacatcaaac atcgaccac ggcgtaacgc gcttgctgct tggatgcccg aggcatagac	12780

tgtaccccaa	aaaaacagtc	ataacaagcc	atgaaaaccg	ccactgcgcc	gttaccaccg	12840
ctgcgttcgg	tcaaggttct	ggaccagttg	cgtgagcgca	tacgctactt	gcattacagc	12900
ttacgaaccg	aacaggctta	tgtccactgg	gttcgtgcct	tcatccgttt	ccacgggtgtg	12960
cgtcaccocg	caaccttggg	cagcagcgaa	gtcgaggcat	ttctgtcctg	gctggcgaa	13020
gagcgcaagg	tttcgggtctc	cacgcacgt	caggcatttg	cggccttgc	gttcttctac	13080
ggcaagggtg	tgtgcacgga	tctgccctgg	cttcaggaga	tcggaagacc	tggcccgctg	13140
cggcgcttgc	cgggtgtgct	gaccccggt	gaagtgggtc	gcacccctgg	ttttctggaa	13200
ggcgagcctc	gtttgttgc	ccagcttctg	tatggaacgg	gcaccccggt	cagtgggggt	13260
ttgcaactgc	gggtcaagga	tctggatttc	gatcacggca	cgcacacgt	gccccggggc	13320
aagggtccca	aggatcgggc	cttgatgtta	cccgagagct	tggcaccacg	cctgcgcgag	13380
caggggaatt	aattcccacg	ggttttgctg	cccgcaaacc	ggctgttctg	gtgttgctag	13440
tttgttatca	gaatcgaga	tccggcttca	gccggtttgc	cggctgaaag	cgtatttct	13500
tccagaattg	ccatgatttt	ttcccacgg	gaggcgctac	tggctcccg	gttgtcgga	13560
gctttgatcc	gataagcagc	atgcctgtt	tcaggctgtc	tatgtgtgac	tgttgagctg	13620
taacaagttg	tctcaggtgt	tcaatttcat	gttctagttg	ctttgtttta	ctggtttcac	13680
ctgttctatt	aggtgttaca	tgtgttcat	ctgttacatt	gtcgatctgt	tcatggtgaa	13740
cagctttgaa	tgcacaaaaa	actogtaaaa	gctctgatgt	atctatcttt	tttacaccgt	13800
tttcatctgt	gcacatggac	agttttccct	ttgatatgta	acgggtgaaca	gttgttctac	13860
ttttgtttgt	tagtcttgat	gcttcaactga	tagatacaag	agccataaga	acctcagatc	13920
cttccgtatt	tagccagtat	gttctctagt	gtggttcgtt	gtttttgcgt	gagccatgag	13980
aacgaaccat	tgagatcata	cttactttgc	atgtcactca	aaaattttgc	ctcaaaactg	14040
gtgagctgaa	tttttgagct	taaagcatcg	tgtagtgttt	ttcttagtcc	gttatgtagg	14100
taggaatctg	atgtaatggt	tgttggtatt	ttgtcaccat	tcatttttat	ctggttgttc	14160
tcaagttcgg	ttacgagatc	catttgtcta	tctagttcaa	cttggaataa	caacgtatca	14220
gtcggggcgg	ctcgttatac	aaccaccaat	ttcatattgc	tgtagtgtt	taaatcttta	14280
cttattgggt	tcaaaaccca	ttgggttaagc	cttttaaact	catggtagtt	attttcaagc	14340
attaacatga	acttaaattc	atcaaggcta	atctctatat	ttgccttctg	agttttcttt	14400
tgtgttagtt	cttttaataa	ccactcataa	atcctcatag	agtatttgtt	ttcaaaagac	14460
ttaacatggt	ccagattata	ttttatgaat	ttttttaact	ggaaaagata	aggcaatata	14520
tcttactaa	aaactaattc	taatttttcg	cttgagaact	tggcatagtt	tgtccactgg	14580
aaaatctcaa	agcctttaac	caaaggattc	ctgatttcca	cagttctcgt	catcagctct	14640
ctggttgctt	tagctaatac	accataagca	ttttccctac	tgtgttcat	catctgagcg	14700

ES 2 808 287 T9

tattggttat aagtgaacga taccgtccgt tctttccttg tagggttttc aatcgtggg	14760
ttgagtagtg ccacacagca taaaattagc ttggtttcat gctccgttaa gtcatagcga	14820
ctaategeta gttcatttgc ttgaaaaca actaattcag acatacatct caattggtct	14880
aggtgatttt aat	14893

REIVINDICACIONES

1. Un método para producir un ácido graso o derivado del mismo, que comprende;

- 5 (a) proporcionar una célula hospedadora bacteriana modificada por ingeniería genética que comprende un 5 promotor heterólogo y/o un sitio de unión a ribosomas unido operativamente a una secuencia polinucleotídica que codifica un polipéptido FadR, provocando dicho promotor y/o sitio de unión a ribosomas la sobreexpresión del polipéptido FadR en dicha célula,
- 10 (b) cultivar la célula hospedadora bacteriana modificada por ingeniería genética en un medio de cultivo en condiciones que permitan la producción de un ácido graso o un derivado del mismo, y
- (c) aislar el ácido graso o derivado del mismo de la célula hospedadora bacteriana modificada por ingeniería genética, en donde el ácido graso o el derivado del mismo es un ácido graso, una acil-ACP, una acil-CoA, un aldehído graso, un alcohol de cadena corta, un alcohol de cadena larga, un alcohol graso, un hidrocarburo o un éster.
- 15 2. El método de la reivindicación 1, en donde el polipéptido FadR es codificado por un gen *FadR* obtenido de *Escherichia*, *Salmonella*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Cronobacter*, *Yersinia*, *Serratia*, *Erwinia*, *Pectobacterium*, *Photobacterium*, *Edwardsiella*, *Shewanella* o *Vibrio*.
- 20 3. El método de la reivindicación 1 o 2, en donde el polipéptido FadR es un polipéptido FadR de tipo silvestre o un polipéptido FadR mutante.
4. El método de la reivindicación 3, en donde el polipéptido FadR comprende una mutación en un resto de aminoácido que corresponde al aminoácido 219 de la SEQ ID NO: 1, preferentemente en donde la mutación es una sustitución del
- 25 resto de aminoácido correspondiente al aminoácido 219 de la SEQ ID NO: 1 con un resto de asparagina.
5. El método de una cualquiera de las reivindicaciones 1-4, en donde el polipéptido FadR comprende una secuencia de aminoácidos que tiene al menos un 80 % de identidad con la secuencia de aminoácidos de la SEQ ID NO: 1.
- 30 6. El método de una cualquiera de las reivindicaciones 1-5, en donde la expresión de uno o más genes seleccionados entre el grupo que consiste en *fabA*, *fabB*, *iclR*, *fadA*, *fadB*, *fadD*, *fadE*, *fadI*, *fadJ*, *fadL*, *fadM*, *uspA*, *aceA*, *aceB* y *aceK* está atenuado en la célula hospedadora bacteriana modificada por ingeniería genética.
7. El método de una cualquiera de las reivindicaciones 1-6, en donde el ácido graso o derivado del mismo es un éster, y el éster es una cera, un éster graso o un éster de ácido graso.
- 35 8. El método de la reivindicación 7, en donde el ácido graso o derivado del mismo es un éster de ácido graso, y el éster de ácido graso es un éster metílico de ácido graso (FAME) o un éster etílico de ácido graso (FAEE).
- 40 9. El método de la reivindicación 7 u 8, en donde una éster sintasa se sobreexpresa en la célula hospedadora modificada por ingeniería genética, preferentemente en donde la éster sintasa es ES9 de *Marinobacter hydrocarbonoclasticus* (SEQ ID NO: 2), ES8 de *Marinobacter hydrocarbonoclasticus* (SEQ ID NO: 3), AtfAI de *Alcanivorax borkumensis* SK2 (SEQ ID NO: 4), AtfA2 de *Alcanivorax borkumensis* SK2 (SEQ ID NO: 5), O-acilglicerol diacetiltransferasa de *Marinobacter aquaeolei* VT8 (SEQ ID NO: 6 o SEQ ID NO: 7), una cera sintasa o un éster de
- 45 cera bifuncional sintasa/acil-CoA:aciltransferasa de diacilglicerol (cera-dgaT).
10. El método de una cualquiera de las reivindicaciones 7-9, en donde un complejo acetil-CoA carboxilasa se sobreexpresa en la célula hospedadora modificada por ingeniería genética, preferentemente en donde el complejo acetil-CoA carboxilasa está codificado por dos o más genes de la subunidad acetil-CoA carboxilasa obtenidos de uno
- 50 o más de entre *Corynebacterium glutamicum*, *Escherichia coli*, *Lactococcus lactis*, *Kineococcus radiotolerans*, *Desulfovibrio desulfuricans*, *Erwinia amylovora*, *Rhodospirillum rubrum*, *Vibrio furnissii*, *Stenotrophomonas maltophilia*, *Synechocystis* sp. PCC6803 o *Synechococcus elongatus*.
11. El método de la reivindicación 10, en donde una biotina proteína ligasa se sobreexpresa en la célula hospedadora modificada por ingeniería genética.
- 55 12. El método de la reivindicación 7, en donde el ácido graso o derivado del mismo es un aldehído graso o un alcohol graso y en donde una ácido carboxílico reductasa y una tioesterasa se sobreexpresan en la célula hospedadora modificada por ingeniería genética, preferentemente en donde una alcohol deshidrogenasa se sobreexpresa en la
- 60 célula hospedadora modificada por ingeniería genética.
13. El método de la reivindicación 12, en donde la ácido carboxílico reductasa es carB (SEQ ID NO: 8), la tioesterasa es tesA (SEQ ID NO: 11) y la alcohol deshidrogenasa es YjgB (SEQ ID NO: 10) o AlrAadpl (SEQ ID NO: 9).
- 65 14. El método de la reivindicación 12 o 13, en donde una fosfopanteteinil transferasa (PPTasa) se sobreexpresa en la célula hospedadora modificada por ingeniería genética, preferentemente en donde la PPTasa es EntD de *E. coli*

MG1655 (SEQ ID NO: 12).

15. El método de la reivindicación 14, en donde la célula hospedadora se selecciona entre el grupo que consiste en una célula de *E. coli*, una célula de *Pantoea citrea*, una célula de *Bacillus lentus*, una célula de *Bacillus brevis*, una
- 5 célula de *Bacillus stearothermophilus*, una célula de *Bacillus lichen formis*, una célula de *Bacillus alkalophilus*, una célula de *Bacillus coagulans*, una célula de *Bacillus circulans*, una célula de *Bacillus pumilis*, una célula de *Bacillus thuringiensis*, una célula de *Bacillus clausii*, una célula de *Bacillus megaterium*, una célula de *Bacillus subtilis*, una célula de *Bacillus amyloliquefaciens*, una célula de *Trichoderma koningii*, una célula de *Trichoderma viride*, una célula de *Trichoderma reesei*, una célula de *Trichoderma longibrachiatum*, una célula de *Aspergillus awamori*, una célula de
- 10 *Aspergillus fumigates*, una célula de *Aspergillus foetidus*, una célula de *Aspergillus nidulans*, una célula de *Aspergillus niger*, una célula de *Aspergillus oryzae*, una célula de *Humicola insolens*, una célula de *Humicola lanuginosa*, una célula de *Rhodococcus opacus*, una célula de *Rhizomucor miehei*, una célula de *Mucor miehei*, una célula de *Streptomyces lividans*, una célula de *Streptomyces murinus*, una célula de *Actinomyces*, una célula de *Avicennia*
- 15 *Botryococcuse braunii*, una célula de *Chlamydomonas reinhardtii*, una célula de *Dunaliella salina*, un *Synechococcus* Sp. célula PCC 7002, un *Synechococcus* Sp. célula PCC 7942, un *Synechocystis* Sp. célula PCC 6803, una célula de *Thermosynechococcus elongates BP-1*, una célula de *Chlorobium tepidum*, una célula de *Chlorojlexus auranticus*, una célula de *Cromatium vinosum*, una célula de *Rhodospirillum rubrum*, una célula de *Rhodobacter capsulatus*, una célula de *Rhodopseudomonas palustris*, una célula de *Clostridium ljungdahlii*, una célula de *Clostridium thermocellum*,
- 20 una célula de *Penicillium chrysogenum*, una célula de *Pichia pastoris*, una célula de *Saccharomyces cerevisiae*, una célula de *Schizosaccharomyces pombe*, una célula de *Pseudomonas fluorescens* y una célula de *Zymomonas mobilis*.

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
Por favor obsérvese que dependiendo del fin de los genes pueden sobreexpresarse, modificarse, atenuarse o suprimirse						
1. Aumento de la producción de ácido graso/Aumento de la producción de producto.						
<i>aumento de acil-ACP o acil-CoA</i> <i>reducción del catabolismo de derivados e intermedios</i> <i>reducción de la inhibición por retroalimentación</i> <i>atenuación de otras vías que consumen ácidos grasos</i>						
	accA	Acetil-CoA carboxilasa, subunidad A (carboxiltransferasa alfa)	AAC73296, NP_414727	6.4.1.2	aumento de la producción de Malonil-CoA	<i>Escherichia coli, Lactococos</i>
	accB	Acetil-CoA carboxilasa, subunidad B (BCCP: portador de carboxilo de biotina)	NP_417721	6.4.1.2	aumento de la producción de Malonil-CoA	<i>Escherichia coli, Lactococos</i>
	accC	Acetil-CoA carboxilasa, subunidad C (biotina carboxilasa)	NP_417722	6.4.1.2, 6.3.4.14	aumento de la producción de Malonil-CoA	<i>Escherichia coli, Lactococos</i>
	accD	Acetil-CoA carboxilasa, subunidad D (carboxiltransferasa beta)	NP_416819	6.4.1.2	aumento de la producción de Malonil-CoA	<i>Escherichia coli, Lactococos</i>
	accDAI (dtsR1)	Acetil-CoA carboxilasa, subunidades fusionadas D y A (carboxiltransferasa)	YP_224991	6.4.1.2	aumento de la producción de Malonil-CoA	<i>Corynebacterium glutamicum</i>
	accBC	Acetil-CoA carboxilasa, subunidades fusionadas B y C (BCCP: proteína portadora de carboxilo de biotina y biotina carboxilasa)	YP_224999	6.4.1.2, 6.3.4.14	aumento de la producción de Malonil-CoA	<i>Corynebacterium glutamicum</i>

FIG. 1A

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	acc1	Acetil-CoA carboxilasa (carboxiltransferasa, BCCP: proteína portadora de carboxilo de biotina y biotina carboxilasa)	NP_014413	6.4.1.2, 6.3.4.14	aumento de la producción de Malonil-CoA	<i>Saccharomyces cerevisiae</i>
	birA	biotina proteína ligasa	YP_225000	6.3.4.15	aumento de la producción de Malonil-CoA	<i>Corynebacterium glutamicum</i>
	bpl1	biotina proteína ligasa	NP_010140	6.3.4.15	aumento de la producción de Malonil-CoA	<i>Saccharomyces cerevisiae</i>
	aceE	piruvato deshidrogenasa, subunidad E1	NP_414656, AAC73226	1.2.4.1	aumento de la producción de acetil- CoA	<i>Escherichia coli</i>
	aceF	piruvato deshidrogenasa, subunidad E2	NP_414657	2.3.1.12	aumento de la producción de acetil- CoA	<i>Escherichia coli</i>
	ackA	acetato cinasa	AAC75356, NP_416799	2.7.2.1	aumento de la producción de acetil- CoA	<i>Escherichia coli</i>
	ackB	acetato cinasa AckB	BAB81430	2.7.2.1	aumento de la producción de acetil- CoA	<i>Escherichia coli</i>
	acpP	proteína portadora de acilo	AAC74178	NINGUNO	aumento de la producción de acil- ACP graso/CoA	<i>Escherichia coli</i>
	fadD	acil-CoA sintasa	AP_002424	2.3.1.86, 6.2.1.3	aumento de la producción de ácido graso	<i>Escherichia coli</i> W3110
	adhE	alcohol deshidrogenasa	CAA47743	1.1.1.1, 1.2.1.10	aumento de la producción de acetil- CoA	<i>Escherichia coli</i> W3111
	fabA	beta-hidroxidecanoil tioéster deshidratasa/isomerasa	NP_415474	4.2.1.60	aumento de la producción de acil- ACP graso/CoA	<i>E. coli</i> K12

FIG. 1B

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	fabB	3-oxoacil-[proteína-portadora-de-acilo] sintasa I	BAA16180	CE: 2.3.1.41	aumento de la producción de acil-ACP graso/CoA	<i>Escherichia coli</i>
	fabD	[proteína-portadora-de-acilo] S-maloniltransferasa	AAC74176	2.3.1.39	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12
	fabF	3-oxoacil-[proteína-portadora-de-acilo] sintasa II	AAC74179	2.3.1.179	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12
	fabG	3-oxoacil-[proteína-portadora-de-acilo] reductasa	AAC74177	1.1.1.100	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12
	fabH	3-oxoacil-[proteína-portadora-de-acilo] sintasa III	AAC74175	2.3.1.180	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12
	fabI	enoiI-[proteína-portadora-de-acilo] reductasa	NP_415804	1.3.1.9	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12
	fabR	Represor de la transcripción	NP_418398	NINGUNO	modulación de la producción de ácidos grasos insaturados	<i>E. coli</i> K12
	fabV	enoiI-[proteína-portadora-de-acilo] reductasa	YP_001217283	1.3.1.9	aumento de la producción de acil-ACP graso/CoA	<i>Vibrio cholerae</i>
	fabZ	(3R)-hidroximiristol acil proteína portadora deshidratasa	NP_414722	4.2.1.-	aumento de la producción de acil-ACP graso/CoA	<i>E. coli</i> K12

FIG. 1C

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	fadE	acil-CoA deshidrogenasa	AAC73325	1.3.99.3, 1.3.99.-	reducción de la degradación de ácidos grasos	<i>E. coli</i> K13
	frdA	fumarato reductasa	NP_418578	1.3.1.6	aumento de la producción de acetil-CoA	<i>Escherichia coli</i>
	GST, gshB	Glutación sintasa	P04425	θ.3.2.3	aumento de Acil-CoA	<i>E. coli</i> K12
	gpsA	sn-glicerol 3-fosfato deshidrogenasa biosintética	AAC76632, NP_418065	CE: 1.1.1.94	aumento de la producción de acetil-CoA	<i>E. coli</i> K12
	ldhA	lactato deshidrogenasa	AAC74462, NP_415898	CE: 1.1.1.27, 1.1.1.28	aumento de la producción de acetil-CoA	<i>E. coli</i> K12
	panD	aspartato 1-descarboxilasa	BAB96708	4.1.1.11	aumento de Acil-CoA	<i>Escherichia coli</i> W3110
	panK también conocido como coaA	pantotenato cinasa	AAC76952	2.7.1.33	aumento de la producción de acetil-CoA	<i>E. coli</i>
	panK también conocido como coaA, R106K	pantotenato cinasa	AAC76952	2.7.1.33	aumento de la producción de acetil-CoA	<i>E. coli</i>
	pdh	Piruvato deshidrogenasa	BAB34380, AAC73226, NP_415392	1.2.4.1	aumento de la producción de acetil-CoA	
	pflB	formiato acetiltransferasa (piruvato formiato liasa)	AAC73989, P09373	CE: 2.3.1.54	aumento de la producción de acetil-CoA	
	plsB	aciltransferasa	AAC77011	2.3.1.15	reducción de los límites de la agrupación de Acil-CoA	<i>E. coli</i> K12
	poxB	piruvato oxidasa	AAC73958, NP_415392	1.2.2.2	aumento de la producción de acetil-CoA	
	pta	fosfotransacetilasa	AAC75357, NP_416800	2.3.1.8	aumento de la producción de acetil-CoA	

FIG. 1D

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	udhA	nucleótido piridínico transhidrogenasa	CAA46822	1.6.1.1	conversión de NADH en NADPH o viceversa	
	atoB	acetil-CoA acetiltransferasa	NP_416728	2.3.1.9	síntesis de acetoacetil-CoA	<i>E. coli</i>
	atoC	proteína reguladora de la transcripción	NP_416724	NINGUNO	síntesis de acetoacetil-CoA	<i>E. coli</i>
	yqeF	acetil-CoA acetiltransferasa	NP_417321	2.3.1.9	síntesis de acetoacetil-CoA	<i>E. coli</i>
	fadB	3-hidroxiacil-CoA condensada epimerasa/delta(3)-cis-delta(2)-trans-enoil-CoA isomerasa/enoil-CoA hidratasa y 3-hidroxiacetil-CoA deshidrogenasa	AP_003956	4.2.1.17, 5.1.2.3, 5.3.3.8, 1.1.1.35	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>
	fadJ	3-hidroxiacil-CoA deshidrogenasa; K01692 enoil-CoA hidratasa; K01782 3-hidroxiacil-CoA epimerasa	AAC75401	1.1.1.35, 4.2.1.17, 5.1.2.3	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>
	fadA	3-cetoacil-CoA tiolasa	BAE77458	2.3.1.16	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>
	fadI	beta-cetoacil-CoA tiolasa	AAC75402	2.3.1.16	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>
	fadR	proteína reguladora de la transcripción	NP_415705	NINGUNO	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>

FIG. 1E

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	ydiO	acil-coA deshidrogenasa	YP_852786	1.3.99.-	Bloqueo e inversión de la degradación de ácidos grasos	<i>E. coli</i>
	ter	transenol-CoA reductasa	G5EU90	1.3.1.8	biosíntesis inversa de ácidos grasos	<i>Euglena gracialis</i>
2. Control de la estructura						
2A. Control de la longitud de la cadena						
2	tesA	tioesterasa	P0ADA1	3.1.2.-, 3.1.1.5	Longitud de cadena C18	
	tesA sin secuencia líder	tioesterasa	AAC73596, NP_415027	3.1.2.-, 3.1.1.5	C18:1	<i>E. coli</i>
	tesA sin líder	tioesterasa	P0ADA1	3.1.2.-, 3.1.1.5	Longitud de cadena <C18	<i>E. coli</i>
	fatB1 (umbellularia)	tioesterasa	041635	3.1.2.14	C12:0	<i>Umbellularia californica</i>
	fatB2 (umbellularia) SUPRIMIR umbelluria)	tioesterasa	AAC49269	3.1.2.14	C8:0-C10:0	<i>Cuphea hookeriana</i>
	fatB3	tioesterasa	AAC72881	3.1.2.14	C14:0-C16:0	<i>Cuphea hookeriana</i>
	fatB (canela)	tioesterasa	039473	3.1.2.14	C14:0	<i>Cinnamomum camphora</i>
	fatB [M141T] *	tioesterasa	CAA85388	3.1.2.14	C16:1	<i>Arabidopsis thaliana</i>
	fatA1 (Helianthus)	tioesterasa	AAL79361	3.1.2.14	C18:1	<i>Helianthus annuus</i>

FIG. 1F

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	Atfata (ARABIDOPSIS FATA ACIL-ACP TIOESTERASA)	tioesterasa	NP_189147, NP_193041	3.1.2.14	C18:1	<i>Arabidopsis thaliana</i>
	fatA	tioesterasa	CAC39106	3.1.2.14	C18:1	<i>Brassica juncea</i>
	fatA (cuphea)	tioesterasa	AAC72883	3.1.2.14	C18:1	<i>Cuphea hookeriana</i>
	tes	tioesterasa	YP_130990	3.1.2.14		<i>Photobacterium profundum</i>
	tesB	tioesterasa	NP_414986	3.1.2.14		<i>E. coli</i>
	fadM	tioesterasa	NP_414977	3.1.2.14		<i>E. coli</i>
	yciA	tioesterasa	NP_415769	3.1.2.14		<i>E. coli</i>
	ybgC	tioesterasa	NP_415264	3.1.2.14		<i>E. coli</i>
<u>2B. Control de ramificación</u>						
	atenuación de FabH					
	expresión de FabH de <i>S. glaucescens</i> o <i>S. coelicolor</i> e inactivación de FabH endógeno				aumento de derivados de ácidos grasos de cadena ramificada	

FIG. 1G

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	expresión de FabH de <i>B. subtilis</i> e inactivación de FabH endógeno					
	bdk - E3 - subunidad de deshidrolipoil deshidrogenasa			CE 1.2.4.4		
	bkd - E1 - subunidad alfa/beta	subunidades de descarboxilasa de complejo de a-cetoácido deshidrogenasa de cadena ramificada		CE 1.2.4.4		
	bkd - E2 - subunidad de dihidrolipoil transacilasa			CE 1.2.4.4		
	bkdA1	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	NP_628006	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>
	bkdB1	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	NP_628005	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>
	bkdC1	dihidrolipoil transacilasa (E2)	NP_628004	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>

FIG. 1H

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	bkdA2	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	NP_733618	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>
	bkdB2	subunidad b de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	NP_628019	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>
	bkdC2	dihidrolipoil transacetilasa (E2)	NP_628018	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces coelicolor</i>
	bkdA	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	<u>BAC72074</u>	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>
	bkdB	subunidad b de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	<u>BAC72075</u>	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>
	bkdC	dihidrolipoil transacetilasa (E2)	<u>BAC72076</u>	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>
	bkdF	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	<u>BAC72088</u>	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>
	bkdG	subunidad b de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	<u>BAC72089</u>	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>
	bkdH	dihidrolipoil transacetilasa (E2)	<u>BAC72090</u>	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Streptomyces avermitilis</i>

FIG. 1I

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	bkdAA	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	NP_390285	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Bacillus subtilis</i>
	bkdAB	subunidad b de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	NP_390284	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Bacillus subtilis</i>
	bkdB	dihidrolipoil transacetilasa (E2)	NP_390283	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Bacillus subtilis</i>
	bkdA1	subunidad a de a-cetoácido deshidrogenasa de cadena ramificada (E1a)	AAA65614	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Pseudomonas putida</i>
	bkdA2	subunidad b de a-cetoácido deshidrogenasa de cadena ramificada (E1b)	AAA65615	CE 1.2.4.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Pseudomonas putida</i>
	bkdC	dihidrolipoil transacetilasa (E2)	AAA65617	CE 2.3.1.168	preparación de precursores de acil-CoA de cadena ramificada	<i>Pseudomonas putida</i>
	lpd	dihidrolipoamida deshidrogenasa (E3)	P09063	1.8.1.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Pseudomonas putida</i>
	lpd	dihidrolipoamida deshidrogenasa (E3)	NP_414658	1.8.1.4	preparación de precursores de acil-CoA de cadena ramificada	<i>Escherichia coli</i>
	lve	aminoácido aminotransferasa de cadena ramificada	YP_026247	2.6.1.42	preparación de a-cetoácidos ramificados	<i>Escherichia coli</i>
	lve	aminoácido aminotransferasa de cadena ramificada	AAF34406	2.6.1.42	preparación de a-cetoácidos ramificados	<i>Lactococcus lactis</i>

FIG. 1J

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	livE	aminoácido aminotransferasa de cadena ramificada	NP_745648	2.6.1.42	preparación de α-cetoácidos ramificados	<i>Pseudomonas putida</i>
	livE	aminoácido aminotransferasa de cadena ramificada	NP_629657	2.6.1.42	preparación de α-cetoácidos ramificados	<i>Streptomyces coelicolor</i>
	ccr	crotonil-CoA reductasa	NP_630556	1.6.5.5, 1.1.1.1	Conversión de crotonil-CoA en butiril-CoA	<i>Streptomyces coelicolor</i>
	ccr	crotonil-CoA reductasa	AAD53915	1.6.5.5, 1.1.1.1	Conversión de crotonil-CoA en butiril-CoA	<i>Streptomyces cinnamonensis</i>
	IcmA, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad A	NP_629554	5.4.99.2	conversión de butiril-CoA en isobutiril-CoA	<i>Streptomyces coelicolor</i>
	IcmA, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad A	AAC08713	5.4.99.2	conversión de butiril-CoA en isobutiril-CoA	<i>Streptomyces cinnamonensis</i>
	IcmB, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad B	NP_630904	5.4.99.2	conversión de butiril-CoA en isobutiril-CoA	<i>Streptomyces coelicolor</i>
	IcmB, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad B	CAB59633	5.4.99.2	conversión de butiril-CoA en isobutiril-CoA	<i>Streptomyces cinnamonensis</i>

FIG. 1K

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	Genes de FabH, ACP y fabF con especificidad para acil-CoA de cadena ramificada					
	IlvE	aminoácido aminotransferasa de cadena ramificada	CAC12788	CE 2.6.1.42	aminoácido de cadena ramificada amino transferasa	<i>Staphylococcus carnosus</i>
	FabH1	beta-cetoacil-ACP sintasa III	NP_626634	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Streptomyces coelicolor</i>
	FabH3	beta-cetoacil-ACP sintasa III	NP_823466	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Streptomyces avermitilis</i>
	FabH_A	beta-cetoacil-ACP sintasa III	NP_389015	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>
	FabH_B	beta-cetoacil-ACP sintasa III	NP_388898	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>

FIG. 1L

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	SmaIDRAFT_08 18	beta-cetoacil-ACP sintasa III	ZP 01643059	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Stenotrophomonas maltophilia</i>
	FabH	beta-cetoacil-ACP sintasa III	YP 123672	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Legionella pneumophila</i>
	FabH	beta-cetoacil-ACP sintasa III	NP 415609	2.3.1.180	inicio de la biosíntesis de ácidos grasos de cadena ramificada	<i>Escherichia coli</i>
Para producir ácidos grasos cíclicos						
	AnsJ	deshidratasa (supuesta)	no disponible	no disponible	biosíntesis de ciclohexil carbonil-CoA	<i>Streptomyces collinus</i>
	AnsK	CoA ligasa (supuesta)	no disponible	no disponible	biosíntesis de ciclohexil carbonil-CoA	<i>Streptomyces collinus</i>
	AnsL	deshidrogenasa (supuesta)	no disponible	no disponible	biosíntesis de ciclohexil carbonil-CoA	<i>Streptomyces collinus</i>
	ChcA	enoi-CoA reductasa	U72144	CE 1.3.1.34	biosíntesis de ciclohexil carbonil-CoA	<i>Streptomyces collinus</i>
	AnsM	oxidorreductasa (supuesta)	no disponible	no disponible	biosíntesis de ciclohexil carbonil-CoA	<i>Streptomyces collinus</i>

FIG. 1M

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	PlmJ	deshidratasa (supuesta)	<u>AAQ84158</u>	no disponible	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp.HK803</i>
	PlmK	CoA ligasa (supuesta)	<u>AAQ84158</u>	no disponible	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp.HK803</i>
	PlmI	deshidrogenasa (supuesta)	<u>AAQ84159</u>	no disponible	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp.HK803</i>
	ChcA	enoiI-CoA reductasa	<u>AAQ84160</u>	CE 1.3.1.34	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp.HK803</i>
	PlmM	oxidorreductasa (supuesta)	<u>AAQ84161</u>	no disponible	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp.HK803</i>
	ChcB	enoiI-CoA isomerasa	AF268489	no disponible	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
	ChcB/CaiD	enoiI-CoA isomerasa	NP_629292	4.2.1.-	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces coelicolor</i>
	ChcB/CaiD	enoiI-CoA isomerasa	NP_824296	4.2.1.-	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces avermitilis</i>

FIG. 1N

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
<u>2C. Control del nivel de saturación</u>						
	Sfa	Supresor de FabA	AAN79592, AAC44390	NINGUNO	Aumento de ácidos grasos monoinsaturados	<i>E. coli</i>
	véase también FabA y FabB en sec. 1				producción de ácidos grasos insaturados	
	GnsA	supresores de la mutación nula secG	ABD18647.1	NINGUNO	aumento ésteres de ácidos grasos insaturados	<i>E. coli</i>
	GnsB	supresores de la mutación nula secG	AAC74076.1	NINGUNO	aumento ésteres de ácidos grasos insaturados	<i>E. coli</i>
	véanse también los elementos de la sección 2A - con :0 son insaturados (sin dobles enlaces) y con :1 son saturados (1 doble enlace)					
	fabB	3-oxoacil-[proteína-portadora-de-acilo] sintasa I	BAA16180	CE: 2.3.1.41	modulación de la producción de ácidos grasos insaturados	<i>Escherichia coli</i>

FIG. 10

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	fabK	trans-2-enoil-ACP reductasa II	AAF98273	1.3.1.9	modulación de la producción de ácidos grasos insaturados	<i>Streptococcus pneumoniae</i>
	fabL	enoiI-(proteína portadora de acilo) reductasa	AAU39821	1.3.1.9	modulación de la producción de ácidos grasos insaturados	<i>Bacillus licheniformis</i> DSM13
	fabM	trans-2, cis-3-decenoiI-ACP isomerasa	DAA05501	4.2.1.17	modulación de la producción de ácidos grasos insaturados	<i>Streptococcus mutans</i>
	des	D5 acil grasa desaturasa	034653	1.14.19	modulación de la producción de ácidos grasos insaturados	<i>Bacillus subtilis</i>
3. Producción de producto final						
3A. Producción de cera						
	AT3G51970	alcohol de cadena larga O-grasa-aciltransferasa	NP_190765	2.3.1.26	producción de cera	<i>Arabidopsis thaliana</i>
		tiosterasa (véase la sección de control de la longitud de la cadena)			aumento de la producción de ácido graso	
		acil-CoA reductasa formadora de alcohol graso		1.1.1.*	conversión de acil-CoA en alcohol graso	
	acr1	acil-CoA reductasa (ACR1)	YP_047869	1.2.1.42	conversión de acil-CoA en alcohol graso	<i>Acinetobacter</i> sp. ADP1
	yqhD	alcohol deshidrogenasa	AP_003562	1.1.-.-	aumento	<i>E. coli</i> W3110
	ELO1	Ácido graso elongasa	BAD98251	2.3.1.-	producción de ácidos grasos de longitud de cadena muy larga	<i>Pichia angusta</i>

FIG. 1P

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	plsC	aciltransferasa	AAA16514	2.3.1.51		<i>Saccharomyces cerevisiae</i>
	DAGAT/DGAT	diacilglicerol aciltransferasa	AAF19262	2.3.1.20	producción de cera	<i>Arabidopsis thaliana</i>
	hWS	acil-CoA alcohol de cera aciltransferasa	AAX48018	2.3.1.20	producción de cera	<i>Homo sapiens</i>
	aft1	éster de cera sintasa/acil-CoA:diacilglicerol aciltransferasa bifuncional	AAO17391	2.3.1.20	producción de cera	<i>Acinetobacter sp. ADP1</i>
	WS377	éster de cera sintasa	ABO21021	2.3.1.20	producción de cera	<i>Marinobacter hydrocarbonoclasticus</i>
	mWS	éster de cera sintasa (<i>simmondsia</i>)	AAD38041	2.3.1.-	producción de cera	<i>Simmondsia chinensis</i>
<u>3B. Producción de alcoholes grasos</u>						
		diversas tioesterasias (véase la Sec. 2A)			hidrólisis de acil-CoA graso/ACP	
	acr1	acil-CoA reductasa	YP_047869	1.2.1.42	reducción de acil-CoA graso en aldehídos grasos	<i>Acinetobacter sp. ADP1</i>
	yqhD	alcohol deshidrogenasa	AP_003562	1.1.-.-	reducción de aldehídos grasos en alcoholes grasos	<i>Escherichia coli W3110</i>
	alrA	alcohol deshidrogenasa	CAG70252	1.1.-.-	reducción de aldehídos grasos en alcoholes grasos	<i>Acinetobacter sp. ADP1</i>

FIG.1Q

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	BmFAR	FAR (acil-CoA reductasa formadora de alcoholes grasos)	BAC79425	1.1.1.*	reducción de acil-CoA graso en alcohol graso	<i>Bombyx mori</i>
	GTNG_1865	aldehído deshidrogenasa de cadena larga	YP_001125970	1.2.1.3	reducción de aldehídos grasos en alcoholes grasos	<i>Geobacillus thermodenitrificans</i> NG80-2
	AAR	Acil-ACP reductasa	YP_400611	1.2.1.42	reducción de acil-ACP graso/CoA en aldehídos grasos	<i>Synechococcus elongatus</i>
	carB	Ácido graso reductasa	YP_889972	6.2.1.3, 1.2.1.42	reducción de ácidos grasos en aldehído graso	<i>Mycobacterium smegmatis</i>
	FadD	acil-CoA sintetasa	NP_416319	6.2.1.3	activación de ácidos grasos a acil-CoA grasos	<i>E. coli</i> K12
Para preparar Butanol						
	atoB	acetil-CoA acetiltransferasa	YP_049388	2.3.1.9	producto fresco	<i>Erwinia carotovora</i>
	hbd	Beta-hidroxibutiril-CoA deshidrogenasa	BAD51424	1.1.1.157	producto fresco	<i>Butyrivibrio fibrisolvens</i>
	CPE0095	crotonasa	BAB79801	4.2.1.55	producto fresco	<i>Clostridium perfringens</i>
	bcd	butiril-CoA deshidrogenasa	AAM14583	1.3.99.2	producto fresco	<i>Clostridium beijerinckii</i>
	ALDH	aldehído deshidrogenasa de acilación de la coenzima A	AAT66436	1.2.1.3	producto fresco	<i>Clostridium beijerinckii</i>
	AdhE	aldehído-alcohol deshidrogenasa	AAN80172	1.1.1.1 1.2.1.10	producto fresco	<i>Escherichia coli</i> CFT073
3C. Producción de éster acetílico de alcohol graso						

FIG. 1R

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
	tioesterasa	véase la sección de control de la longitud de la cadena			producto fresco	
	acr1	acil-CoA reductasa	YP_047869	1.2.1.42	producto fresco	<i>Acinetobacter sp. ADP1</i>
	yqhD	alcohol deshidrogenasa	APJD03562	1.1.-.-	producto fresco	<i>E. Coli K12</i>
	AAT	alcohol O-acetiltransferasa	AAG13130	2.3.1.84	producto fresco	<i>Fragaria x ananassa</i>
<u>4. Exportación</u>						
	AtMRP5	asociado a resistencia a múltiples fármacos de <i>Arabidopsis thaliana</i>	NP_171908	NINGUNO	productos de exportación	<i>Arabidopsis thaliana</i>
	AmiS2	Transportador de ABC AmiS2	JC5491	NINGUNO	productos de exportación	<i>Rhodococcus sp.</i>
	AtPGPI	P GLUCOPROTEÍNA1 DE <i>ARABIDOPSIS THALIANA</i>	NP_181228	NINGUNO	productos de exportación	<i>Arabidopsis thaliana</i>
	AcrA	supuesta proteína de transporte de salida de múltiples fármacos acrA	CAF23274	NINGUNO	productos de exportación	<i>Candidatus Protochlamydia amoebophila UWE25</i>
	AcrB	probable proteína transportadora de salida de múltiples fármacos, acrB	CAF23275	NINGUNO	productos de exportación	<i>Candidatus Protochlamydia amoebophila UWE25</i>
	TolC	Proteína de membrana externa [biogénesis de la envoltura celular,	ABD59001	NINGUNO	productos de exportación	<i>Francisella tularensis subsp. novicida</i>
	AcrE	la proteína transmembrana afecta a la formación de septos y a la permeabilidad de la membrana celular	YP_312213	NINGUNO	productos de exportación	<i>Shigella sonnei Ss046</i>
	AcrF	proteína F de resistencia a la acriflavina	P24181	NINGUNO	productos de exportación	<i>Escherichia coli</i>
	tlI1619	transportador de salida de múltiples fármacos	NP_682409.1	NINGUNO	productos de exportación	<i>Thermosynechococcus elongatus BP-1</i>
	tlI0139	transportador de salida de múltiples fármacos	NP_680930.1	NINGUNO	productos de exportación	<i>Thermosynechococcus elongatus BP-1</i>

FIG. 1S

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO DE CE	USO	ORGANISMO
5. Fermentación						
	Genes de punto de control de la replicación				aumento de la eficiencia de la producción	
	umuD	ADN polimerasa V, subunidad	YP_310132	3.4.21.-	aumento de la eficiencia de la producción	<i>Shigella sonnei</i> Ss046
	umuC	ADN polimerasa V, subunidad	ABC42261	2.7.7.7	aumento de la eficiencia de la producción	<i>Escherichia coli</i>
	NADH: NADPH transhidrogenasa (subunidades alfa y beta) (pntA, pntB)		P07001, P0AB70	1.6.1.2	aumento de la eficiencia de la producción	<i>Shigella flexneri</i>

FIG. 1T

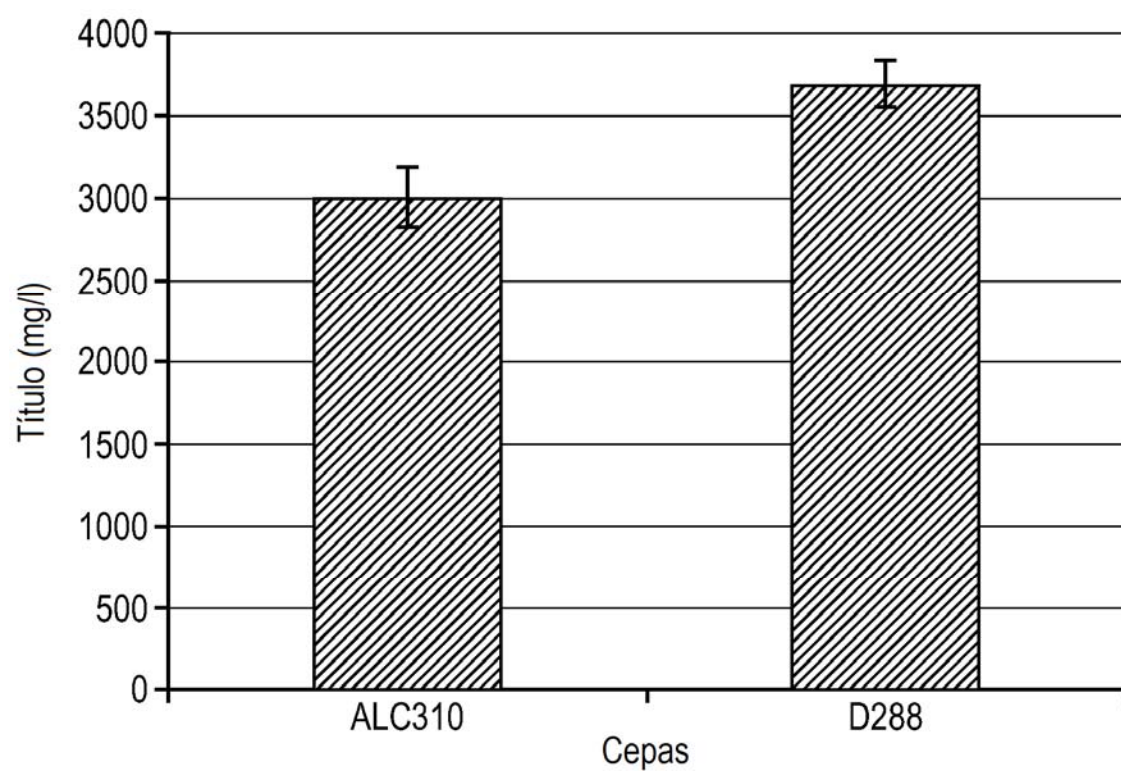


FIG. 2

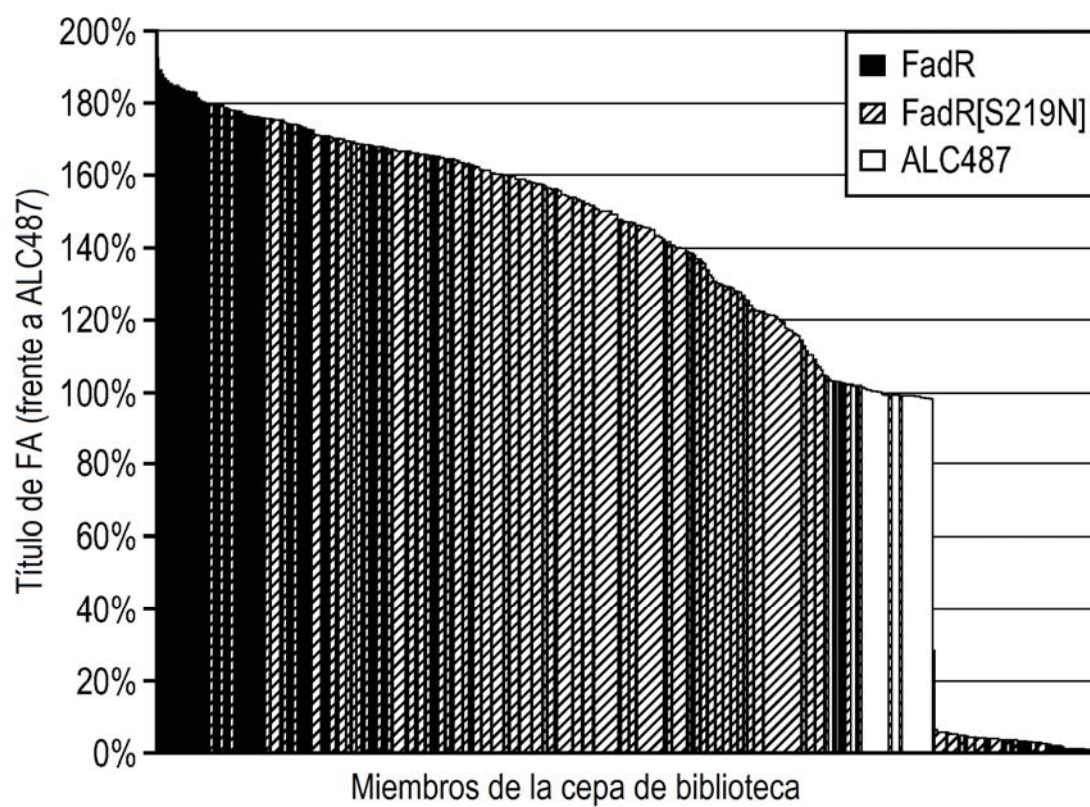


FIG. 4

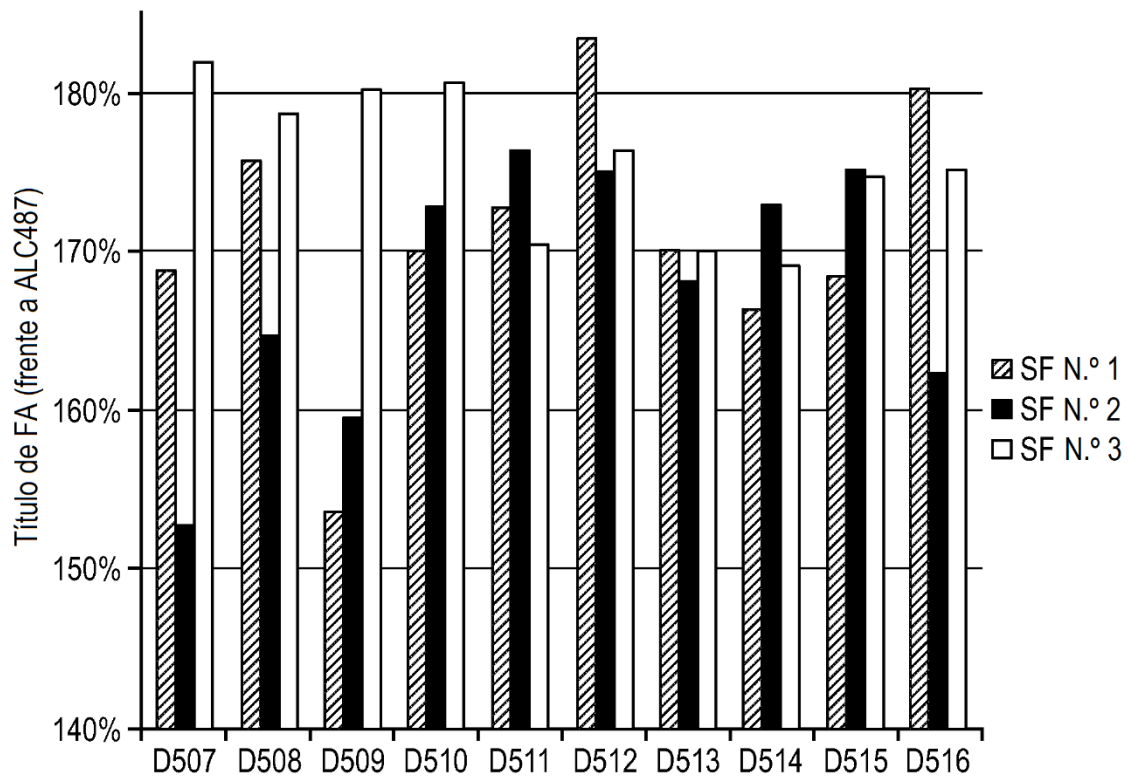


FIG. 5

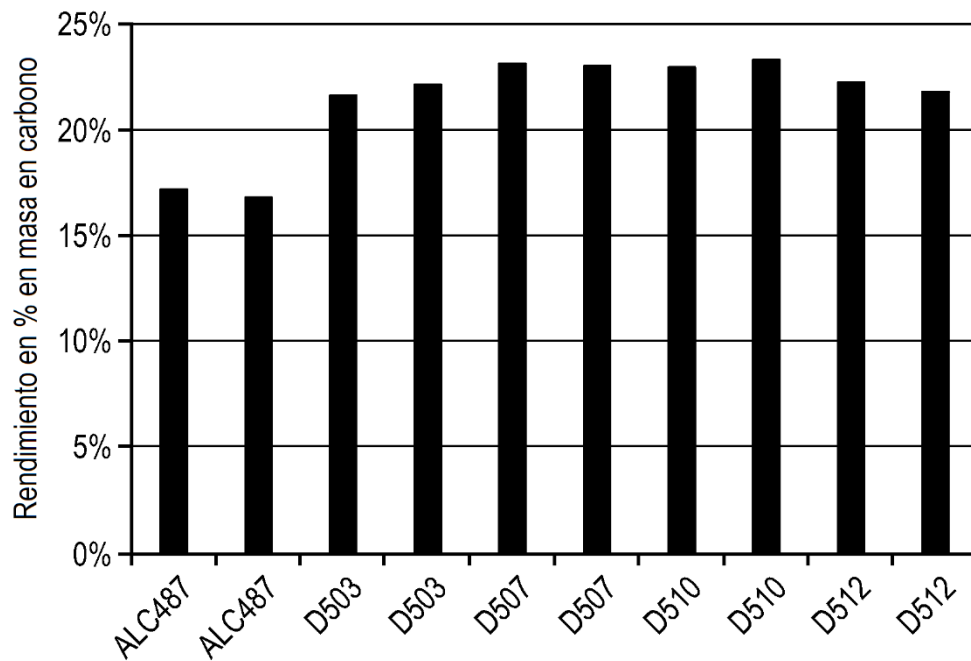


FIG. 6

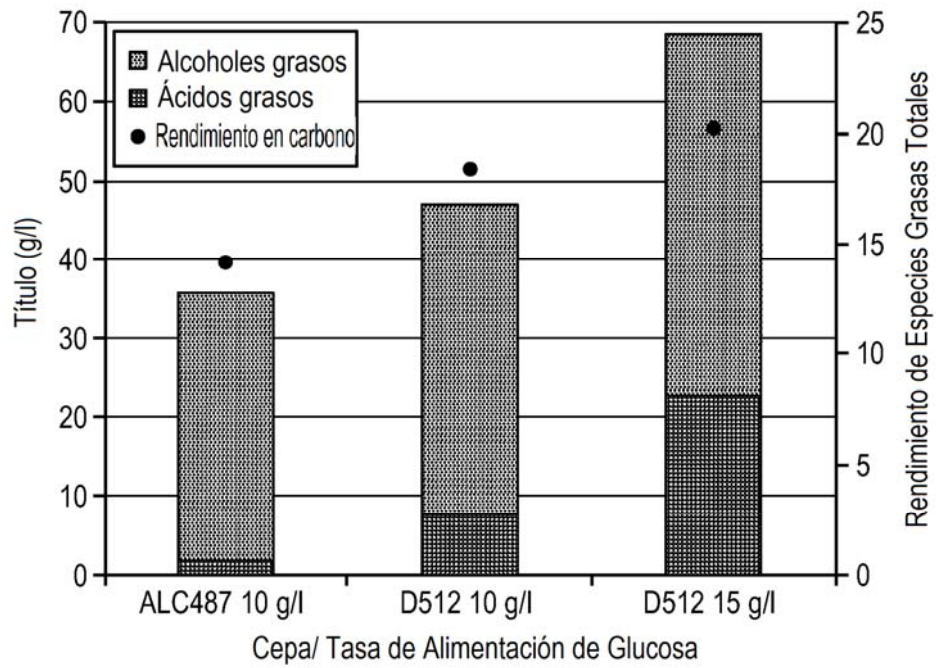


FIG. 7

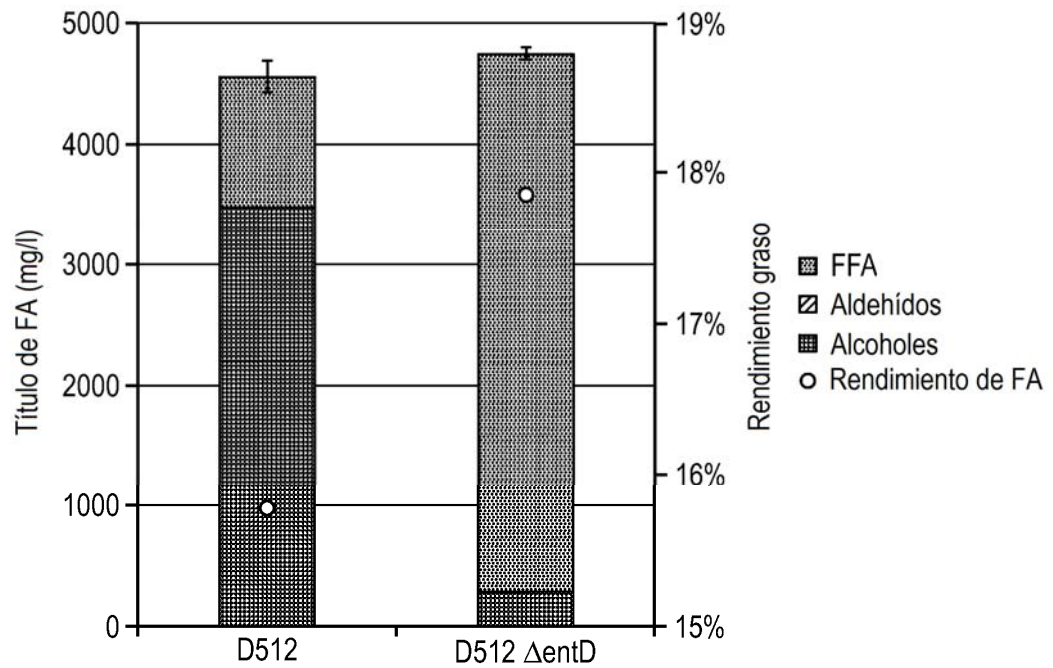


FIG. 8

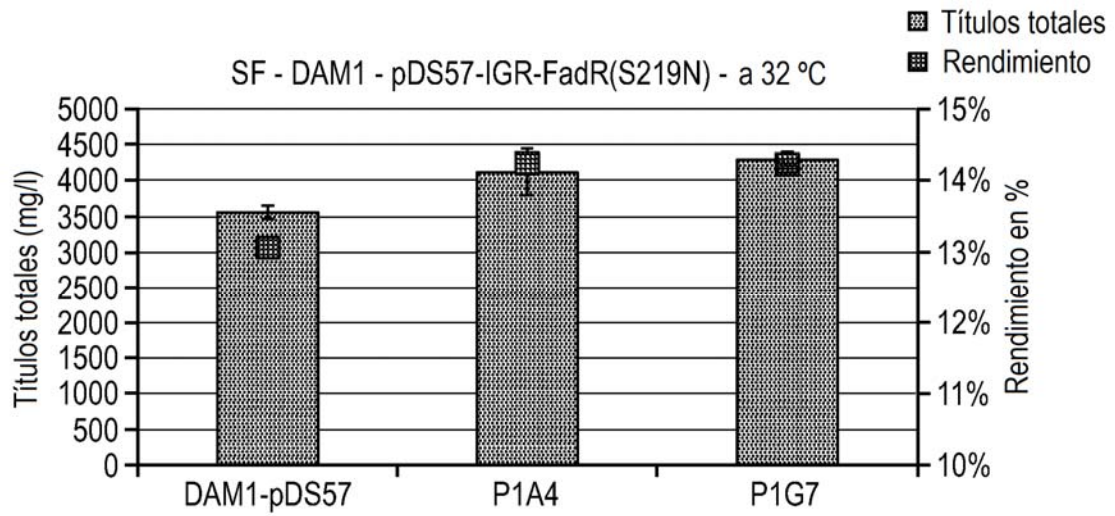


FIG. 9

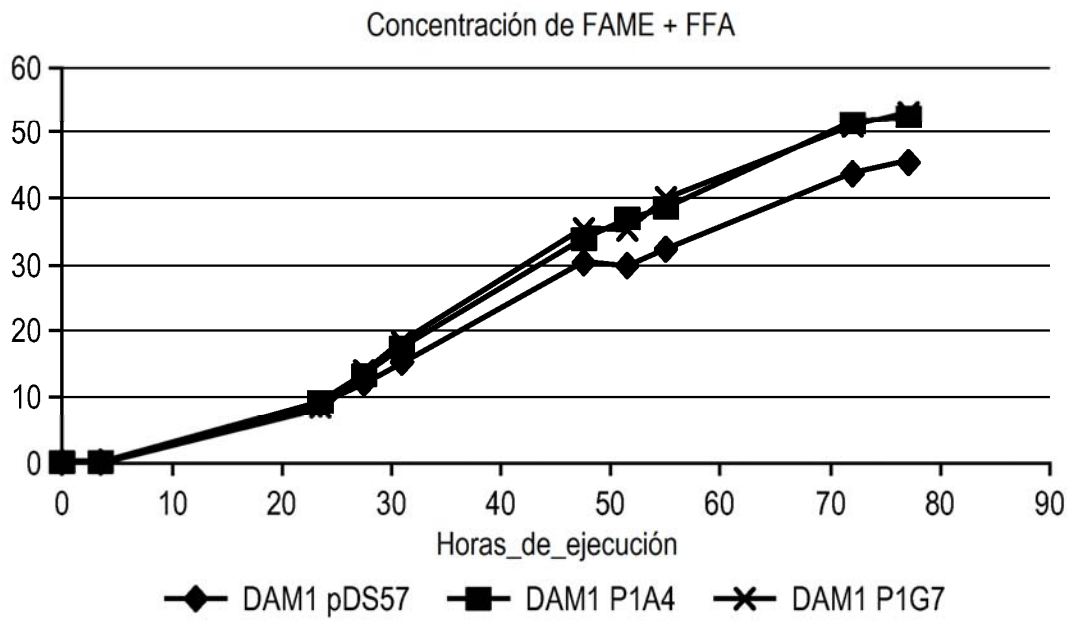


FIG. 10

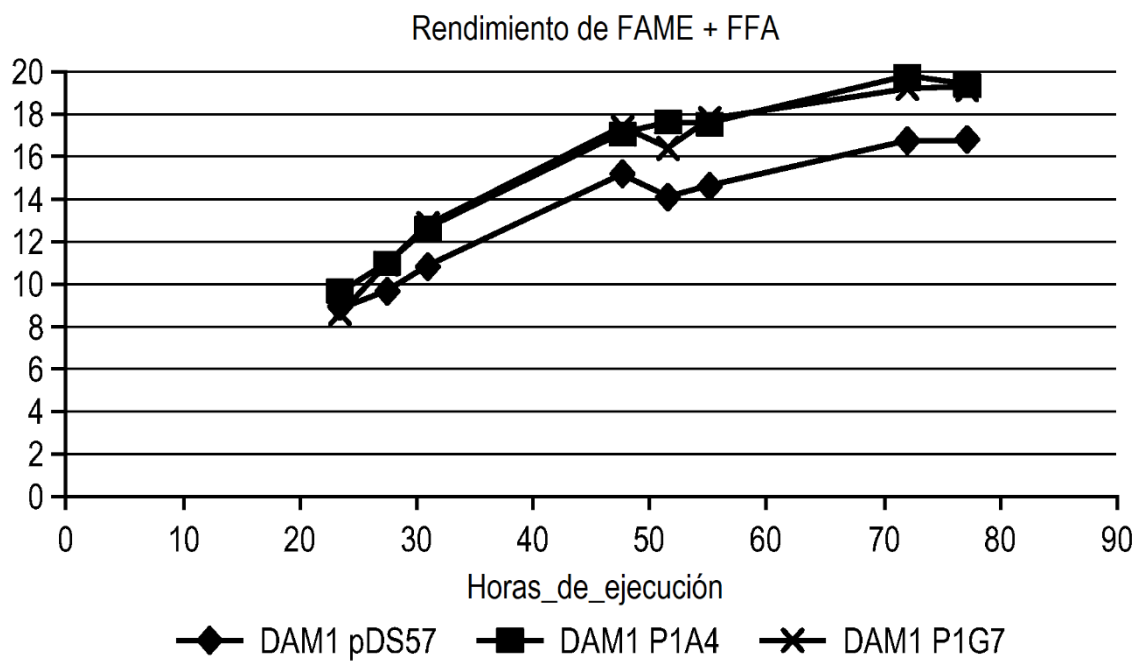


FIG. 11