

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 413 482**

(51) Int. Cl.:

**C12N 15/54** (2006.01)  
**C12N 9/10** (2006.01)  
**C12P 7/64** (2006.01)  
**C12N 1/21** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **10.11.2008 E 10007946 (6)**

(97) Fecha y número de publicación de la concesión europea: **27.03.2013 EP 2327769**

---

(54) Título: **Organismos hiperfotosintéticos**

(30) Prioridad:

**10.11.2007 US 987046 P**  
**28.02.2008 US 32169 P**  
**22.08.2008 US 90933 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**16.07.2013**

(73) Titular/es:

**JOULE UNLIMITED TECHNOLOGIES, INC.**  
(100.0%)  
18 Crosby Drive  
Bedford, MA 01730, US

(72) Inventor/es:

**DEVROE, ERIC;**  
**KOSURI, SRIRAM;**  
**BERRY, DAVID;**  
**AFEYAN, NOUBAR;**  
**SKRALY, FRANK;**  
**ROBERTSON, DAN;**  
**GREEN, BRIAN y**  
**RIDLEY, CHRISTIAN**

(74) Agente/Representante:

**UNGRÍA LÓPEZ, Javier**

**ES 2 413 482 T3**

---

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

**DESCRIPCIÓN**

Organismos hiperfotosintéticos

**5 Remisión a solicitudes relacionadas**

Esta solicitud reivindica prioridad de las solicitudes provisionales de Estados Unidos 60/987.046, presentada el 10 de noviembre de 2007; 61/032.169 presentada el 28 de febrero de 2008; y 61/090.933, presentada el 22 de agosto de 2008.

10

**Campo**

La presente invención se refiere al campo de biología sintética, y más particularmente a organismos fotoautotróficos industrializados diseñados para convertir de forma eficaz el dióxido de carbono y luz en biomasa y productos de interés basados en carbono.

15

**Antecedentes**

20 La fotosíntesis es un proceso por el cual las entidades biológicas utilizan la luz solar y el CO<sub>2</sub> para producir azúcares para la energía. Los organismos fotoautotróficos existentes (es decir, plantas, algas, y bacterias fotosintéticas) son poco adecuadas para el bioprocесamiento industrial. En particular, la mayoría de los organismos tienen lento tiempos de duplicación (10-72 horas) en comparación con los organismos heterotróficos industrializados tales como *Escherichia coli* (20 minutos). Además, los organismos fotoautotróficos a menudo son susceptibles a variaciones moderadas en estrés ambiental común incluyendo pH, temperatura y tolerancia salina. Dichas susceptibilidades

25 hacen de las aplicaciones industriales de los fotoautótrofos ineficaces. Además, crecientemente factores ambientales tóxicos (por ejemplo, contaminantes tóxicos incluyendo metales pesados, subproductos industriales basados en nitrógeno y azufre) pueden además limitar las aplicaciones de los fotoautótrofos a usos industriales particulares.

30 Los productos deseables que pueden producirse potencialmente en microorganismos (por ejemplo, etanol y otros alcoholes superiores de cadena ramificada producidos en *E. coli* modificada por ingeniería (Atsumi, et al. Nature (2008) vol. 451: 86-90)) se han hallado difíciles de procesar en fotoautótrofos a causa de vías metabólicas incompatibles o ineficaces de producción o la ausencia completa de biocatalizadores necesarios basados en células.

35 Por tanto, existe la necesidad en la técnica de organismos fotoautotróficos mejorados que sean más adecuados para bioprocесamiento industrial.

Nomura et al. (1995) Plant Physiology: 107; 703-708 describe la transformación de *Synechococcus sp.* PCC7942 con un plásmido lanzadera que codifica un fragmento de 9 kb que codifica el grupo génico bet de *E. coli*.

40

La entrada en la base de datos UniProt con número de acceso P25665 y Gonzalez et al (1992) Biochemistry: 31:6045-6056 describen 5-metiltetrahidropteroiltriglutamato - homocisteína metiltransferasa (Met E.).

45

**Sumario**  
En este documento se describen métodos y composiciones para diseñar por ingeniería vías que utilicen de forma estable organismos fotosintéticos por, por ejemplo, mejora genética de sus eficacias de captura de la luz y fijación del carbono y por optimización de sus propiedades de crecimiento para la propagación en fotobioreactores. También se describen las células "hiperfotosintéticas" modificadas por ingeniería resultantes y organismos que posibilitan la conversión eficaz de dióxido de carbono y luz en productos de interés basados en carbono.

50 La presente invención proporciona una célula cianobacteriana modificada por ingeniería que comprende un ácido nucleico modificado por ingeniería que codifica una proteína 5-metiltetrahidropteroiltriglutamato - homocisteína metiltransferasa independiente de vitamina B<sub>12</sub> (metE) que comprende una secuencia que es al menos un 95% idéntica a la secuencia mostrada en la Figura 4. Esta secuencia es la secuencia de la proteína metE de *Escherichia coli* K12 (número de acceso NP\_418273.1).

55 La presente invención también proporciona el uso de una proteína 5-metiltetrahidropteroiltriglutamato - homocisteína metiltransferasa independiente de vitamina B<sub>12</sub> que comprende una secuencia que es al menos un 95% idéntica a la secuencia mostrada en la Figura 4 para producir una célula cianobacteriana independiente de vitamina B<sub>12</sub>.

60 La presente invención también proporciona un método para producir un producto de interés basado en carbono, que comprende cultivar una célula cianobacteriana modificada por ingeniería como se ha descrito anteriormente en presencia de CO<sub>2</sub> y luz en condiciones adecuadas para producir el producto de interés basado en carbono.

65

**Figuras**

La **Figura 1** proporciona la Tabla 1 que identifica los genes que pueden expresarse o regularse positivamente.

5 La **Figura 2** proporciona la Tabla 2 que identifica los genes que pueden regularse negativamente o eliminarse.

La **Figura 3** representa la viabilidad de *Synechococcus* sp. PCC 7002 de tipo silvestre y modificada por ingeniería genética en placas de agarosa de medio A<sup>+</sup> sin B12.

10 La **Figura 4** muestra la secuencia de la proteína metE de *Escherichia coli* K12 (número de acceso NP\_418273.1).

**Abreviaturas y términos**

15 Las siguientes explicaciones de términos y métodos se proporcionan para describir mejor la presente descripción y guiar a los especialistas en la técnica en la práctica de la presente descripción. Como se usa en este documento, "que comprende" significa "que incluye" y las formas singulares "un" o "una" o "el/la" incluyen referencias plurales salvo que el contexto indique claramente lo contrario. Por ejemplo, referencia a "que comprende una célula" incluye una o una pluralidad de dichas células, y referencia a "que comprende una tioesterasa" incluye referencia a uno o más péptidos tioesterasa y equivalentes de los mismos conocidos para los especialistas en la técnica, y similares. El 20 término "o" se refiere a un elemento individual de elementos alternativos indicados o una combinación de dos o más elementos, salvo que el contexto indique claramente lo contrario.

25 Salvo que se explique de otro modo, todos los términos técnicos y científicos usados en este documento tienen el mismo significado que el habitualmente comprendido para los especialistas en la técnica a la que pertenece esta descripción. Aunque métodos y materiales similares o equivalentes a los descritos en este documento pueden usarse en la práctica o ensayo de la presente descripción, a continuación se describen los métodos y materiales adecuados. Los materiales, métodos, y ejemplos son solamente ilustrativos y no pretenden ser limitantes. Otras características de la descripción son evidentes a partir de la siguiente descripción detallada y las reivindicaciones.

30 **Números de acceso:** Los números de acceso durante toda esta descripción corresponden a los hallados en las siguientes bases de datos: NCBI (National Center for Biotechnology Information), TIGR (The Institute for Genomic Research), y KEGG (Kyoto Encyclopedia of Genes and Genomes). Los números de acceso son los proporcionados en las bases de datos a 10 de noviembre de 2007.

35 **Números de clasificación de enzimas (EC):** Los números EC proporcionados durante toda esta descripción se obtienen de la base de datos de ligandos de KEGG, mantenida por la Kyoto Encyclopedia of Genes and Genomes, patrocinada en parte por la Universidad de Tokio. Los números EC son los proporcionados en la base de datos a 10 de noviembre de 2007.

40 **ADN:** Ácido desoxirribonucleico. El ADN es un polímero de cadena larga que incluye el material genético de la mayoría de los organismos vivos (algunos virus tienen genes que incluyen ácido ribonucleico, ARN). Las unidades repetidas en polímeros de ADN son cuatro nucleótidos diferentes, cada uno de los cuales incluye una de las cuatro bases adenina, guanina, citosina y timina unida a un azúcar desoxirribosa al cual está adherido un grupo fosfato.

45 **Codón:** Tripletes de nucleótidos, mencionados como codones, en moléculas de ADN codifican aminoácidos en un péptido. El término codón también se usa para las secuencias correspondientes (y complementarias) de tres nucleótidos en el ARNm en el cual se transcribe el ADN.

50 **Endógeno:** Como se usa en este documento con referencia a una molécula de ácido nucleico y una célula particular o microorganismo se refiere a una secuencia de ácido nucleico o péptido que está en la célula y no se introdujo en la célula usando técnicas de ingeniería recombinante. Por ejemplo, un gen que estaba presente en la célula cuando la célula se aisló originalmente de la naturaleza se considera que es endógeno. Un gen aún se considera endógeno si las secuencias de control, tales como un promotor o secuencias potenciadoras que activan la transcripción o traducción se han alterado a través de técnicas recombinantes.

55 **Exógeno:** Como se usa en este documento con referencia a una molécula de ácido nucleico y una célula particular o microorganismo se refiere a una secuencia de ácido nucleico o péptido que no estaba presente en la célula cuando la célula se aisló originalmente de la naturaleza. Por ejemplo, un ácido nucleico que se originó en un microorganismo diferente y se modificó por ingeniería en una célula alternativa usando técnicas de ADN recombinante u otros métodos para suministrar dicho ácido nucleico se considera que es exógeno.

60 **Expresión:** El proceso por el cual la información codificada de un gen se convierte en las moléculas que soportan las estructuras y funciones de una célula, tal como una proteína, ARN transferente, o ARN ribosómico. Los genes expresados incluyen aquellos que se transcriben en ARNm y después se traducen en proteína y aquellos que se transcriben en ARN pero no se traducen en proteína (por ejemplo, ARN transferentes y ribosómicos).

- Sobre-expresión:** La sobre-expresión se refiere a cualquier estado en el cual se causa que un gen se transcriba a una velocidad elevada en comparación con la velocidad de transcripción endógena para ese gen. En algunos ejemplos, la sobre-expresión incluye adicionalmente una velocidad elevada de traducción del gen en comparación con la velocidad de traducción endógena para ese gen. Los métodos para ensayar la sobre-expresión son bien conocidos en la técnica, por ejemplo los niveles de ARN transcripto pueden evaluarse usando reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) y los niveles de proteína pueden evaluarse usando análisis de electroforesis en gel de poliacrilamida con dodecil sulfato sódico (SDS-PAGE). Además, se considera que un gen se sobre-expresa cuando muestra actividad elevada en comparación con su actividad endógena, que puede ocurrir, por ejemplo, a través de la reducción en la concentración o actividad de su inhibidor, o mediante la expresión de una versión mutante con actividad elevada. En realizaciones preferidas, cuando la célula hospedadora codifica un gen endógeno con una actividad bioquímica deseada, es útil para sobre-expresar un gen exógeno, que permite un control regulador más explícito en el bioprocесamiento y un medio para mitigar potencialmente los efectos de la regulación del metabolismo central, que se centra alrededor de los genes nativos explícitamente.
- Regulación negativa:** La regulación negativa se refiere a cualquier estado en el cual se causa que un gen se transcriba a una velocidad reducida en comparación con la velocidad de transcripción génica endógena para ese gen. En ciertas realizaciones, la expresión génica está regulada negativamente mediante la expresión de ácidos nucleicos, tales como oligonucleótidos antisentido, ARN bicatenario, ARN interferente pequeño, ARN de horquilla pequeño, microARN, ribozimas, y similares. En algunos ejemplos, la regulación negativa incluye adicionalmente un nivel reducido de traducción del gen en comparación con la velocidad de traducción endógena para ese gen. Además, un gen se considera que está regulado negativamente cuando muestra actividad disminuida en comparación con su actividad endógena, que puede suceder, por ejemplo, a través de un aumento en la concentración o actividad de su inhibidor, o mediante la expresión de una versión mutante con actividad reducida. Los métodos para ensayar la regulación negativa son bien conocidos para los especialistas en la técnica, por ejemplo, los niveles de ARN transcripto pueden evaluarse usando RT-PCR y los niveles de proteínas pueden evaluarse usando análisis SDS-PAGE.
- Eliminación:** Un gen cuyo nivel de expresión o actividad se ha reducido a cero. En algunos ejemplos, un gen se elimina mediante delección o remplazo de algo de o toda su secuencia codificante. En otros ejemplos, un gen se elimina mediante la introducción o retirada de uno o más nucleótidos en su fase de lectura abierta, que provoca la traducción de un producto proteico sin sentido o no funcional de otro modo.
- Autótrofo:** Los autótrofos (u organismos autotróficos) son organismos que producen compuestos orgánicos complejos a partir de moléculas inorgánicas simples y una fuente de energía externa, tal como luz (fotoautótrofos) o reacciones químicas de compuestos inorgánicos.
- Heterótrofo:** Los heterótrofos (u organismos heterotróficos) son organismos que, a diferencia de los autótrofos, no pueden obtener la energía directamente de la luz o de compuestos químicos inorgánicos, y por tanto deben alimentarse de substratos de carbono orgánico. Obtienen energía química descomponiendo las moléculas orgánicas que consumen. Los heterótrofos incluyen animales, hongos, y numerosos tipos de bacterias.
- Hiperfotosintético:** Una célula u organismo que expresa proteínas fotosintéticas que a través de técnicas de ADN recombinante se han modificado por ingeniería específicamente para expresar ácidos nucleicos endógenos y/o exógenos que provocan una o más mejoras funcionales relacionadas con el bioprocесamiento industrial y la conversión de dióxido de carbono y luz en productos de carbono reducido o masa celular. Una célula hiperfotosintética también abarca una célula u organismo modificado por ingeniería como se ha descrito anteriormente que se ha evolucionado o mutagenizado para conseguir una o más mejoras funcionales.
- Hidrocarburo:** Generalmente se refiere a un compuesto químico que consta de los elementos carbono (C), opcionalmente oxígeno (O), e hidrógeno (H).
- Vía biosintética:** También mencionada como "vía metabólica", se refiere a una serie de reacciones bioquímicas anabólicas o catabólicas para convertir (transmutar) una especie química en otra. Por ejemplo, una vía biosintética de hidrocarburo se refiere a una serie de reacciones bioquímicas que convierten aportaciones y/o metabolitos en intermedios tipo producto de hidrocarburo y después en hidrocarburos o productos de hidrocarburo. Las vías anabólicas implican construir una molécula más grande a partir de moléculas más pequeñas, un proceso que requiere energía. Las vías catabólicas implican descomponer moléculas más grandes, a menudo liberando energía.
- Celulosa:** La celulosa  $[(C_6H_{10}O_5)_n]$  es un carbohidrato polisacárido polimérico de cadena larga, de beta-glucosa. Forma el componente estructural principal de las plantas y no es digerible por seres humanos. La celulosa es un material común en paredes celulares vegetales y se apreció por primera vez como tal en 1838. Existe de forma natural en forma casi pura solamente en fibra de algodón; en combinación con lignina y cualquier hemicelulosa, se halla en todo el material vegetal.
- Tensioactivos:** Los tensioactivos son sustancias capaces de reducir la tensión superficial de un líquido en el cual se disuelven. Están típicamente compuestos de una cabeza soluble en agua y una cadena o cola de hidrocarburo. El

grupo soluble en agua es hidrófilo y puede ser iónico o no iónico, y la cadena de hidrocarburo es hidrófoba.

**Biocombustible:** Un biocombustible es cualquier combustible que se obtiene de una fuente biológica.

- 5   **Ácido nucleico modificado por ingeniería:** Un "ácido nucleico modificado por ingeniería" es una molécula de ácido nucleico que incluye al menos una diferencia a partir de la molécula de ácido nucleico de origen natural. El ácido nucleico modificado por ingeniería incluye todas las secuencias heterólogas modificadas y no modificadas exógenas (es decir, secuencias obtenidas de un organismo o célula diferente al que alberga el ácido nucleico modificado por ingeniería) así como genes endógenos, operones, secuencias codificantes, o secuencias no codificantes, que se han modificado, mutado o que incluyen delecciones o inserciones en comparación con una secuencia de origen natural. Los ácidos nucleicos modificados por ingeniería también incluyen todas las secuencias, independientemente del origen, que están ligadas a un promotor inducible o a otra secuencia de control con la cual no están asociados de forma natural. Los ácidos nucleicos modificados por ingeniería incluyen adicionalmente todas las secuencias que pueden usarse para regular negativamente o eliminar la expresión de un gen endógeno. Estos incluyen moléculas antisentido, moléculas de ARN y, construcciones para producir recombinación homóloga, construcciones cre-lox, y similares.

- 20   **Ácido nucleico de captura de luz:** Un "ácido nucleico de captura de luz" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una o más proteínas que convierten la energía de la luz (es decir, fotones) en energía química tal como un gradiente de protones, energía reductora, o una molécula que contiene al menos un enlace fosfato de alta energía tal como ATP o GTP. Los ácidos nucleicos de captura de luz ejemplares incluyen aquellos que codifican bombas de protones activadas por luz tales como rodopsina, xantorrodopsina, proteorrodopsina y bacteriorrodopsina, así como ácidos nucleicos que codifican proteínas que modulan directamente (complejos recolectores de luz, LHC I y LHC II) o indirectamente (antena de recolección de luz truncada, *tla1*) las eficacias de captura recolectora de la luz. Un ácido nucleico de captura de luz incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión reduce la captura de luz.

- 30   **Ácido nucleico de la vía de fijación de dióxido de carbono:** Un "ácido nucleico de la vía de fijación de dióxido de carbono" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína que posibilita la fijación autotrófica del carbono. Un ácido nucleico de la vía de fijación de dióxido de carbono también incluye ácidos nucleicos que dirigen el flujo de carbono a intermedios celulares clave requeridos para el crecimiento eficaz o la formación de productos basados en carbono, tales como acetil-CoA. Ácidos nucleicos de la vía de fijación de dióxido de carbono ejemplares incluyen aquellos que codifican la propionil-CoA carboxilasa, piruvato sintasa, formato deshidrogenasa, y ribulosa-1,5-bisfosfato carboxilasa/oxygenasa (RuBisCO). Un ácido nucleico de la vía de fijación de dióxido de carbono incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión reduce la fijación del dióxido de carbono.

- 40   **Ácido nucleico de la vía NADH:** Un "ácido nucleico de la vía NADH" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína para mantener un suministro apropiadamente equilibrado de NAD reducido para realizar la fijación del carbono. Ácidos nucleicos de la vía NADH ejemplares incluyen aquellos que codifican una isocitrato deshidrogenasa dependiente de NAD<sup>+</sup> y malato deshidrogenasa. Un ácido nucleico de la vía NADH incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión afecta al mantenimiento de un suministro apropiadamente equilibrado de NAD reducido para realizar la fijación del carbono y otros procesos biológicos necesarios.

- 50   **Ácido nucleico de la vía NADPH:** Un "ácido nucleico de la vía NADPH" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína para mantener un suministro apropiadamente equilibrado de NADPH reducido para realizar la fijación del carbono. Ácidos nucleicos de la vía NADPH ejemplares incluyen aquellos que codifican la fosfogluconolactonasa y la piridina nucleótido transhidrogenasa soluble. Un ácido nucleico de la vía NADPH incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión afecta al mantenimiento de un suministro apropiadamente equilibrado de NADP reducido para realizar la fijación del carbono y otros procesos biológicos necesarios.

- 55   **Ácido nucleico de tolerancia térmica:** Un "ácido nucleico de tolerancia térmica" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición provoca un aumento en la tolerancia térmica. Ácidos nucleicos de tolerancia térmica ejemplares incluyen aquellos que codifican ClpC/Hsp100, groESL1, HspA, y PsbU. Un ácido nucleico de tolerancia térmica incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión altera la tolerancia térmica.

- 65   **Ácido nucleico de tolerancia al pH:** Un "ácido nucleico de tolerancia al pH" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición posibilita el crecimiento a un pH elevado o reducido. Ácidos nucleicos de tolerancia al pH ejemplares incluyen aquellos que codifican glutamato descarboxilasa y superóxido dismutasa. Un ácido nucleico de tolerancia al pH incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes

endógenos cuya expresión altera la tolerancia al pH.

- Tolerancia a gases de combustión:** Un "ácido nucleico de tolerancia a gases de combustión" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición posibilita el crecimiento en presencia de componentes de gases de combustión incluyendo dióxido de carbono, SO<sub>x</sub>, NO<sub>x</sub>, y N<sub>2</sub>. Ácidos nucleicos de tolerancia a gases de combustión ejemplares incluyen aquellos que codifican la superóxido dismutasa, catalasa, cisteína sintasa, y NirK. Un ácido nucleico de tolerancia a gases de combustión incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión altera la tolerancia a gases de combustión.

**Ácido nucleico de independencia de nutrientes:** Un "ácido nucleico de independencia de nutrientes" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición posibilita la propagación en ausencia de, o en concentraciones reducidas de, un nutriente exógeno. Ácidos nucleicos de independencia de nutrientes ejemplares incluyen aquellos que codifican MetE y NrdB. Un gen de independencia de nutrientes incluye adicionalmente un ácido nucleico usado para reducir la expresión de o eliminar uno o más genes endógenos cuya expresión altera la propagación en ausencia de, o en concentraciones reducidas de, un nutriente exógeno.

**Ácido nucleico de tolerancia salina:** Un "ácido nucleico de tolerancia salina" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición posibilita la propagación en condiciones de elevada salinidad, tal como agua de mar. Ácidos nucleicos de tolerancia salina ejemplares incluyen aquellos que codifican el antiportador Na<sup>+</sup>/H<sup>+</sup>, y conservación básica de mama.

**Ácido nucleico del flujo de acetil-CoA:** Un "ácido nucleico del flujo de acetil-CoA" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición provoca un aumento en la acetil-CoA producida sobre una unidad de tiempo. Ácidos nucleicos ejemplares que pueden sobre-expresarse incluyen pantotenato quinasa y piruvato deshidrogenasa. Los ácidos nucleicos que pueden regularse negativamente, inhibirse, o eliminarse incluyen acil coenzima A deshidrogenasa, glicerol 3-fosfato deshidrogenasa biosintética, y lactato deshidrogenasa.

**Ácido nucleico de clorofila d:** Un "ácido nucleico de clorofila d" se refiere a un ácido nucleico que solo o en combinación con otro ácido nucleico codifica una proteína cuya sobre-expresión, regulación negativa, o inhibición provoca la biosíntesis de clorofila d y su incorporación en centros de reacción, el fotosistema (PS)I y PSII. Un ácido nucleico de clorofila d posibilita la propagación en, por ejemplo, condiciones en las que la radiación fotosintéticamente disponible se usa probablemente de forma completa por organismos que absorben luz usando clorofila a y/o clorofila b de modo que el entorno tiene baja intensidad de luz visible pero elevada intensidad cercana a los infrarrojos donde ningún otro organismo fotosintético absorbe fuertemente de modo que un organismo que usa la clorofila d para la captura de la luz puede prosperar. Como alternativa, un ácido nucleico de clorofila d posibilita la propagación en condiciones en las que la longitud de onda de la luz artificial usada por la iluminación se selecciona para permitir la propagación de un organismo que usa clorofila d para la captura de la luz.

**Identidad de secuencia:** Los términos "idéntico" o porcentaje de "identidad", en el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de restos aminoacídicos o nucleótidos que son iguales (es decir, aproximadamente el 60% de identidad, preferiblemente el 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o mayor identidad sobre una región especificada, cuando se comparan y alinean para una correspondencia máxima sobre una ventana de comparación o región designada) medida usando algoritmos de comparación de secuencia BLAST o BLAST 2.0 con parámetros por defecto descritos a continuación, o por alineación manual e inspección visual (véase, por ejemplo, el sitio web NCBI o similares). Dichas secuencias entonces se dice que son "sustancialmente idénticas". Esta definición también se refiere a, o puede aplicarse a, el complementario de una secuencia de ensayo. La definición también incluye secuencias que tienen delecciones y/o adiciones, así como aquellas que tienen sustituciones. Como se describe a continuación, los algoritmos preferidos pueden tener en cuenta huecos y similares. Preferiblemente, existe identidad sobre una región que es al menos de aproximadamente 25 aminoácidos o nucleótidos de longitud, o más preferiblemente sobre una región que es de 50-100 aminoácidos o nucleótidos de longitud.

Para la comparación de secuencias, típicamente una secuencia actúa como secuencia de referencia, con la cual se compara las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se introducen en un ordenador, se designan las coordenadas de subsecuencia, si es necesario, y se designan los parámetros del programa del algoritmo de secuencia. Preferiblemente, pueden usarse los parámetros por defecto del programa, o pueden designarse parámetros alternativos. El algoritmo de comparación de secuencia entonces calcula el porcentaje de identidad de secuencia para las secuencias de ensayo con relación a la secuencia de referencia, en base a los parámetros del programa.

**Ventana de comparación:** Una "ventana de comparación", como se usa en este documento, incluye la referencia a un segmento de una cualquiera de varias posiciones contiguas seleccionadas entre el grupo compuesto por 20 a 600, habitualmente de aproximadamente 50 a aproximadamente 200, más habitualmente de aproximadamente 100 a aproximadamente 150 en que una secuencia puede compararse con una secuencia de referencia de la misma cantidad de posiciones contiguas después de que las dos secuencias se hayan alineado de forma óptima. Los métodos de alineación de secuencias para la comparación son bien conocidos en la técnica. La alineación óptima de secuencias para comparación puede realizarse, por ejemplo, por el algoritmo de homología local de Smith Y Waterman, Adv. Appl. Math. 2:482 (1981), por el algoritmo de alineación de homología de Needleman y Wunsch, J. Mol. Biol. 48:443 (1970), por el método de búsqueda de similitud de Pearson y Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), por implementaciones informatizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA del paquete de software Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o por alineación manual e inspección visual (véase, por ejemplo, Current Protocols in Molecular Biology (Ausubel y col., eds. suplemento 1995)).

Un ejemplo preferido de algoritmo que es adecuado para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul y col., Nuc. Acids Res. 25:3389-3402 (1977) y Altschul y col., J. Mol. Biol. 215: 403-410 (1990), respectivamente. BLAST y BLAST 2.0 se usan, con los parámetros descritos en este documento, para determinar el porcentaje de identidad de secuencia para los ácidos nucleicos y proteínas de la invención. El software para realizar análisis BLAST está disponible al público a través del sitio web del National Center for Biotechnology Information. Este algoritmo implica identificar primero pares de secuencias con alta valoración (HSP) identificando palabras cortas de longitud W en la secuencia en cuestión, que coinciden o satisfacen algún valor umbral T de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T se menciona como el umbral del valor de palabra vecina (Altschul y col., supra). Estos aciertos de palabra vecina inicial actúan como semillas para iniciar búsquedas para hallar HSP más largos que los contengan. Los aciertos de palabra se extienden en ambas direcciones a lo largo de cada secuencia lo más lejos para que pueda aumentarse el valor de alineación cumulativo. Los valores cumulativos se calculan usando, para secuencias de nucleótidos, los parámetros M (valor de bonificación para un par de restos apareados; siempre >0) y N (valor de penalización para restos desapareados; siempre <0). Para secuencias de aminoácidos, se usa una matriz de valoración para calcular el valor cumulativo. La extensión de los aciertos de palabra en cada dirección se detienen cuando: el valor de alineación cumulativo queda fuera por la cantidad X desde su valor máximo conseguido; el valor cumulativo llega a cero o por debajo, debido a la acumulación de una o más alineaciones de restos de valoración negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de palabra de 3, una expectativa (E) de 10, y la matriz de valoración BLOSUM62 (véase Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)) alineaciones (B) de 50, expectativas (E) de 10, M=5, N=-4, y una comparación de ambas hebras.

#### 40 Descripción detallada de la invención

##### Organismos

La presente invención posibilita la conversión de una cianobacteria fotoautotrófica en un organismo hiperfotosintético.

Las algas y cianobacterias incluyen, aunque sin limitación, los siguientes géneros: *Acanthoceras*, *Acanthococcus*, *Acaryocloris*, *Achnanthes*, *Achnanthidium*, *Actinastrum*, *Actinocloris*, *Actinococcus*, *Actinotaenium*, *Amphichrysis*, *Amphidinium*, *Amphikrikos*, *Amphipleura*, *Amphiprora*, *Amphithrix*, *Amphora*, *Anabaena*, *Anabaenopsis*, *Aneumastus*, *Ankistrodesmus*, *Ankyra*, *Anomoeoneis*, *Apatococcus*, *Aphanizomenon*, *Aphanocapsa*, *Aphanochaete*, *Aphanothece*, *Apicystis*, *Apistonema*, *Arthrodesmus*, *Artherospira*, *Ascochloris*, *Asterionella*, *Asterococcus*, *Audouinella*, *Aulacoseira*, *Bacillaria*, *Balbiania*, *Bambusina*, *Bangia*, *Basichlamys*, *Batrachospermum*, *Binuclearia*, *Bitrichia*, *Blidingia*, *Botriopsis*, *Botrydium*, *Botryococcus*, *Botryosphaerella*, *Brachiomonas*, *Brachysira*, *Brachytrichia*, *Brebissonia*, *Bulbochaete*, *Burnilleria*, *Bumilleriopsis*, *Caloneis*, *Calothrix*, *Campylodiscus*, *Capsosiphon*, *Carteria*, *Catena*, *Cavinula*, *Centritractus*, *Centronella*, *Ceratium*, *Chaetoceros*, *Chaetocloris*, *Chaetonella*, *Chaetonema*, *Chaetopeltis*, *Chaetophora*, *Chaetosphaeridium*, *Chamaesiphon*, *Chara*, *Characioclasis*, *Characiopsis*, *Characium*, *Charales*, *Chilomonas*, *Chlainomonas*, *Chlamydolepharis*, *Chlamydocapsa*, *Chlamydomonas*, *Chlamydomonopsis*, *Chlamydomyxia*, *Chlamydomyxa*, *Chlamydophysis*, *Chlorangiopsis*, *Chlorella*, *Chlorobotrys*, *Chlorobrachis*, *Chlorochytrium*, *Chlorococcum*, *Chlorogloea*, *Chlorogonium*, *Chlorolobion*, *Chloromonas*, *Chlorophysema*, *Chlorophytia*, *Chlorosaccus*, *Chlorosarcina*, *Choricystis*, *Chromophyton*, *Chromulina*, *Chroococcidiopsis*, *Chroococcus*, *Chroodactylon*, *Chroomonas*, *Chroothece*, *Chrysamoeba*, *Chrysapsis*, *Chrysidiastrum*, *Chrysocapsa*, *Chrysocapsella*, *Chrysochaete*, *Chrysochromulina*, *Chrysococcus*, *Chrysocrinus*, *Chrysolepidomonas*, *Chrysolykos*, *Chrysonebula*, *Chrysophyta*, *Chrysopyxis*, *Chrysosaccus*, *Chrysphaerella*, *Chrysostephanothaera*, *Cladophora*, *Clastidium*, *Closteriopsis*, *Closterium*, *Coccozymxa*, *Coccineis*, *Coelastrella*, *Coelastrum*, *Coelosphaerium*, *Coenocloris*, *Coenococcus*, *Coenocystis*, *Colacium*, *Coleochaete*, *Collodictyon*, *Compsogonopsis*, *Compsopogon*, *Conjugatophyta*, *Conochaete*, *Coronastrum*, *Cosmarium*, *Cosmioneis*,

- 5 *Cosmocladium, Crateriportula, Craticula, Crinalium, Crucigenia, Crucigeniella, Cryptoaulax, Cryptomonas, Cryptophyta, Ctenophora, Cyanodictyon, Cyanonephron, Cyanophora, Cyanophyta, Cyanothomonas, Ciclonexis, Ciclostephanos, Ciclotella, Cylindrocapsa, Cylindrocystis, Cylindrospermum, Cylindrotheca, Cymatopleura, Cymbella, Cymbellonitzschia, Cystodinium Dactylococcopsis, Debarya, Denticula, Dermatochrysis, Dermocarpa, Dermocarpella, Desmatractum, Desmidium, Desmococcus, Desmonema, Desmosiphon, Diacanthos, Diacronema, Diadesmis, Diatoma, Diatomella, Dicella, Dichothrix, Dichotomococcus, Dicranochaete, Dictyochloris, Dictyococcus, Dictyosphaerium, Didymocystis, Didymogenes, Didymosphenia, Dilabifilum, Dimorfococcus, Dinobryon, Dinococcus, Diplochloris, Diploneis, Diplostauron, Distriponella, Docidium, Draparnaldia, Dunaliella, Dysmorfococcus, Ecballocystis, Elakatothrix, Ellerbeckia, Encyonema, Enteromorpha, Entocladia, Entomoneis, Entophysalis, Epichrysis, Epipyxis, Epitheria, Eremosphaera, Euastropsis, Euastrum, Eucapsis, Eucocconeis, Eudorina, Euglena, Euglenophyta, Eunotia, Eustigmatophyta, Eutreptia, Fallacia, Fischerella, Fragilaria, Fragilariforma, Franceia, Frustulia, Curcilla, Geminella, Genicularia, Glauccystis, Glaucophyta, Glenodiniopsis, Glenodinium, Gloeocapsa, Gloeochaete, Gloeochrysis, Gloeococcus, Gloeocystis, Gloeodendron, Gloemonas, Gloeoplax, Gloeothece, Gloeotila, Gloeotrichia, Gloiodictyon, Golenkinia, Golenkiniopsis, Gomontia, Gomphocymbella, Gomphonema, Gomphosphaeria, Gonatozygon, Gongrosia, Gongrosira, Goniocloris, Gonium, Gonyostomum, Granulocloris, Granulocystopsis, Groenbladia, Gymnodinium, Gymnozyga, Gyrosigma, Haematococcus, Hafniomonas, Hallasia, Hammatoidea, Hannaea, Hantzschia, Hapalosiphon, Haplotaenium, Haptophyta, Haslea, Hemidinium, Hemitonnia, Heribaudiella, Heteromastix, Heterothrix, Hibberdia, Hildenbrandia, Hillea, Holopedium, Homoeothrix, Hormanthonema, Hormotila, Hyalobrachion, Hyalocardium, Hyalodiscus, Hyalogonium, Hyalotheca, Hydriatum, Hydrococcus, Hydrocoleum, Hydrocoryne, Hydrodictyon, Hydrosera, Hydrurus, Hyella, Hymenomonas, Isthmocoron, Johannesbaptistia, Juranyiella, Karayevia, Kathablepharis, Katodinium, Kephryion, Keratococcus, Kirchneriella, Klebsormidium, Kolbesia, Koliella, Komarekia, Korshikoviella, Kraskella Lagerheimia Lagynion Lamprothamnium, Lemanea, Lepocinclis, Leptosira, Lobococcus, Lobocystis, Lobomonas, Luticola, Lyngbya, Malleocloris, Mallomonas, Mantiella, Marssonella, Martyana, Mastigocoleus, Gastogloia, Melosira, Merismopedia, Mesostigma, Mesotaenium, Micractinium, Micrasterias, Microchaete, Microcoleus, Microcysts, Microglena, Micromonas, Microspora, Microthamnion, Mischococcus, Monochrysis, Monodus, Monomastix, Monoraphidium, Monostroma, Mougeotia, Mougeotopsis, Myocloris, Myromecia, Myxosarcina, Naegeliella, Nannocloris, Nautococcus, Navicula, Neglectella, Neidium, Nephroclamy, Nephrocytium, Nephrodiella, Nephroelmis, Netrium, Nitella, Nitellopsis, Nitzschia, Nodularia, Nostoc, Ochromonas, Oedogonium, Oligochaetophora, Onychonema, Oocardium, Oocystis, Opephora, Ophiocytium, Orthoseira, Oscillatoria, Oxyneis, Pachycladella, Palmella, Palmodictyon, Pnadorina, Pannus, Paralia, Pascherina, Paulschulzia, Pediastrum, Pedinella, Pedinomonas, Pedinopera, Pelagodictyon, Penium, Peranema, Peridiniosis, Peridinium, Peronia, Petroneis, Phacotus, Phacus, Phaeaster, Phaeodermatium, Phaeophyta, Phaeosphaera, Phaeothamnion, Phormidium, Phycopeltis, Phyllariocloris, Phyllocardium, Phyllocladus, Pinnularia, Pitophora, Placoneis, Planctonema, Planktosphaeria, Planothidium, Plectonema, Pleodorina, Pleuroastrum, Pleurocapsa, Pleurocladia, Pleurodiscus, Pleurosiga, Pleurosira, Pleurotaenium, Pocillomonas, Podohedra, Polyblepharides, Polychaetophora, Polyedriella, Polyedriopsis, Polygonioclasis, Polyepidomonas, Polyaenia, Polytoma, Polytomella, Porphyridium, Posteriocromonas, Prasinocloris, Prasinocladus, Prasinophyta, Prasiola, Prochlorophyta, Prochlorothrix, Protoderma, Protosiphon, Provasoliella, Prymnesium, Psammocystis, Psammocystidium, Pseudanabaena, Pseudocyclotium, Psuedocarteria, Psuedochate, Psuedocharacium, Psuedococcozymxa, Psuedodictyosphaerium, Psuedokephyrion, Psuedonobyrsa, Psuedoquadrigula, Psuedosphaerocysts, Psuedostaurastrum, Psuedostaurosira, Psuedotetrastrum, Pteromonas, Punctastrata, Piramichlamys, Piramimonas, Pirrophyta, Quadricloris, Quadrigula, Radiococcus, Radiofilum, Raphidiopsis, Raphidocleis, Raphidonema, Raphidophyta, Peimeria, Rhabdoderma, Rhabdomonas, Rhizoclonium, Rhodomomas, Rhodophyta, Rhoicosphenia, Rhopalodia, Rivularia, Rosenvingiella, Rossithidium, Roya, Scenedesmus, Scherffelia, Schizochlamydella, Schizochlamys, Schizomeris, Schizophyllum, Schroederia, Scolioneis, Scotiella, Scotiellopsis, Scourfieldia, Scytonema, Selenastrum, Selenocloris, Sellaphora, Semiorbis, Siderocelis, Diderocystopsis, Dimonsenia, Siphononema, Sirocladium, Sirogonium, Skeletonema, Sorastrum, Spermatozopsis, Sphaerellopsis, Sphaerellopsis, Sphaerodinium, Sphaeroplaea, Sphaerozozma, Spiniferomonas, Spirogyra, Spirotaenia, Spirulina, Spondylomorum, Spondylosium, Sporotetras, Spumella, Staurastrum, Staurodesmus, Stauroneis, Staurosira, Staurosirella, Stenopterobia, Stephanocostis, Stephanodiscus, Stephanoporus, Stephanosphaera, Stichococcus, Stichogloea, Stigeoclonium, Stigonema, Stipitococcus, Stokesiella, Strombomonas, Stylochrysalis, Stylocladum, Stylocladus, Styloxyis, Stylosphaeridium, Surirella, Sykidion, Symploca, Synechococcus, Synechocystis, Synedra, Synochromonas, Synura, Tabellaria, Tabularia, Teilingia, Temnogametum, Tetmemorus, Tetraclorella, Tetracyclus, Tetradesmus, Tetraedriella, Tetraedron, Tetraselmis, Tetraspora, Tetrastrum, Thalassiosira, Thamniochaete, Thorakocloris, Thorea, Tolypella, Tolypothrix, Trachelomonas, Trachydiscus, Trebouxia, Trentepohlia, Treubaria, Tribonema, Trichodesmium, Trichodiscus, Trochiscia, Tryblionella, Ulothrix, Uroglena, Uronema, Urosolenia, Urospora, Uva, Vacuolaria, Vaucheria, Volvox, Volvulina, Westella, Woloszynskia, Xanthidium, Xanthophyta, Xenococcus, Zygnea, Zygernopsis, y Zygonium.*
- 60 La conversión hiperfotosintética requiere modificación genética extensiva (Tablas 1 y 2); por tanto, el organismo fotoautotrófico precursor se transforma con ADN exógeno.

65 Los organismos preferidos para conversión hiperfotosintética incluyen *Synechococcus* sp PCC 7002 y, *Synechococcus* sp. PCC 7942; *Synechocystis* sp. PCC 6803, y *Thermosynechococcus elongatus* BP-1 (cianobacteria).

Propagación

- Los métodos para el cultivo de organismos fotosintéticos en medio líquido y en placas que contienen agarosa son bien conocidos para los especialistas en la técnica (véanse, por ejemplo, los sitios web asociados con la ATCC, y 5 con el Instituto Pasteur). Por ejemplo, las células de *Synechococcus* sp. PCC 7002 (disponibles en la Pasteur Culture Collection of Cyanobacteria) se cultivan en medio BG-11 ( $\text{NaNO}_3$  17,65 mM,  $\text{K}_2\text{HPO}_4$  0,18 mM,  $\text{MgSO}_4$  0,3 mM,  $\text{CaCl}_2$  0,25 mM, ácido cítrico 0,03 mM, citrato de amonio férrico 0,03 mM, EDTA 0,003 mM,  $\text{Na}_2\text{CO}_3$  0,19 mM, 10 2,86 mg/l de  $\text{H}_3\text{BO}_3$ , 1,81 mg/l de  $\text{MnCl}_2$ , 0,222 mg/l de  $\text{ZnSO}_4$ , 0,390 mg/l de  $\text{Na}_2\text{MoO}_4$ , 0,079 mg/l de  $\text{CuSO}_4$ , y 0,049 mg/l de  $\text{Co}(\text{NO}_3)_2$ , pH 7,4) suplementado con 16 µg/l de biotina,  $\text{MgSO}_4$  20 mM, KCl 8 mM, y NaCl 300 mM (véase, por ejemplo, el sitio web asociado con el Instituto Pasteur, y Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L. "Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl. Acad. Sci USA (2004). 101 (52): 18228-33). Típicamente, los cultivos se mantienen a 28°C y se burbujean de forma continua con  $\text{CO}_2$  al 5% a una intensidad de luz de 120 µmol de fotones/m<sup>2</sup>/s.
- 15 *Thermosynechococcus elongatus* BP-1 (disponible en el Kazusa DNA Research Institute, Japón) se propaga en medio BG11 suplementado con TES-KOH 20 mM (pH 8,2) como se ha descrito [Iwai M, Katoh H, Katayama M, Ikeuchi M. "Improved genetic transformation of the thermophilic cyanobacterium, Thermosynechococcus elongatus BP-1," Plant Cell Physiol (2004). 45(2):171-175)]. Típicamente, los cultivos se mantienen a 50°C y se burbujean de forma continua con  $\text{CO}_2$  al 5% a una intensidad de luz de 38 µmol de fotones/m<sup>2</sup>/s.
- 20 Lo anterior define condiciones de propagación típicas. Según sea apropiado, las incubaciones se realizan usando medios alternativos o composiciones gaseosas, temperaturas alternativas (5-75°C), y/o flujos de luz (0-5500 µmol de fotones/m<sup>2</sup>/s).
- 25 La luz se suministra a través de una diversidad de mecanismos, incluyendo iluminación natural (luz solar), incandescencia convencional, fluorescencia, o bombillas halógenas, o mediante propagación en cámaras de crecimiento iluminadas especialmente diseñadas (por ejemplo, cámara de crecimiento iluminada modelo LI15 (Sheldon Manufacturing, Inc. Cornelius, OR). Para experimentos que requieren longitudes de onda y/o intensidades 30 específicas, la luz se distribuye mediante diodos emisores de luz (LED), en los cuales pueden controlarse cuidadosamente los espectros de longitud de onda y la intensidad (Philips).
- 35 El dióxido de carbono se suministra mediante inclusión de suplementos de medio sólido (es decir, bicarbonato sódico) o en forma de un gas mediante su distribución en la incubadora de crecimiento o medio. La mayoría de los experimentos se realizan usando dióxido de carbono gaseoso concentrado, a concentraciones entre el 1 y el 30%, que se burbujea directamente en el medio de cultivo a velocidades suficientes para proporcionar la mezcla para los 40 organismos. Cuando se utiliza dióxido de carbono gaseoso concentrado, el gas se origina en forma pura a partir de cilindros disponibles en el mercado, o preferentemente a partir de fuentes concentradas que incluyen gases residuales o gases de combustión de plantas carboníferas, refinerías, instalaciones de producción de cemento, instalaciones de gas natural, destilerías, y similares.

Plásmidos

- Los plásmidos relevantes para ingeniería genética típicamente incluyen al menos dos elementos funcionales 1) un origen de replicación que posibilita la propagación de la secuencia de ADN en el organismo hospedador, y 2) un marcador de selección (por ejemplo, un marcador de resistencia a antibióticos que confiere resistencia a ampicilina, kanamicina, zoecina, cloranfenicol, tetraciclina, espectinomicina, y similares). Los plásmidos a menudo se mencionan como "vectores de clonación" cuando su propósito principal es posibilitar la propagación de un inserto de ADN heterólogo deseado. Los plásmidos también pueden incluir secuencias reguladoras de acción en cis para dirigir la transcripción y la traducción de insertos de ADN heterólogo (por ejemplo, promotores, terminadores de la transcripción, sitios de unión al ribosoma); dichos plásmidos se mencionan frecuentemente como "vectores de expresión". Cuando los plásmidos contienen elementos funcionales que permiten la propagación en más de una especie, dichos plásmidos se mencionan como "vectores lanzadera". Los vectores lanzadera son bien conocidos para los especialistas en la técnica. Por ejemplo, pSE4 es un vector lanzadera que permite la propagación en *E. coli* y *Synechococcus* [Maeda S, Kawaguchi Y, Ohy T, y Omata T. J. Bacteriol. (1998). 180:4080-4088]. Los vectores lanzadera son particularmente útiles en la presente invención para permitir la fácil manipulación de genes y secuencias reguladoras en *E. coli* antes de la transformación en el organismo fotoautotrófico de interés.

Técnicas de transformación

- 60 Los métodos convencionales para la transformación de procariotas son bien conocidos para los especialistas en la técnica [Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volumen 152 Academic Press, Inc., San Diego, Calif.; Sambrook y col. (1989) Molecular Cloning--A Laboratory Manual (2<sup>a</sup> ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, N.Y.; y Current Protocols in Molecular Biology, F. M. Ausubel y col., eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley y Sons, Inc., (hasta e incluyendo el suplemento 1997)].

Muchos organismos procariotas son competentes de forma natural; otros pueden volverse competentes por tratamientos químicos. Ejemplos no limitantes de técnicas de transformación incluyen incubación directa en presencia de ADN exógeno, transformación por choque térmico, transformación por electroporación, transformación por bombardeo con partículas biológicas, transformación mediante adición de lípidos o agentes fusogénicos (es decir, polietilenglicol), conjugación con un microorganismo heterólogo, o transducción mediante partículas víricas.

5 Las células de *Synechococcus* sp. PCC 7002 se transforman de acuerdo con el protocolo optimizado descrito previamente [Essich ES, Stevens Jr E, Porter RD "Chromosomal Transformation in the Cyanobacterium Agmenellum quadruplicatum". J Bacteriol (1990). 172 (4): 1916-1922]. Las células se cultivan en medio A (18 g/l de NaCl, 5 g/l de 10 MgSO<sub>4</sub> · 7 H<sub>2</sub>O, 30 mg/l de Na<sub>2</sub>EDTA, 600 mg/l de KCl, 370 mg/l de CaCl<sub>2</sub> · 2 H<sub>2</sub>O, 1 g/l de NaNO<sub>3</sub>, 50 mg/l de KH<sub>2</sub>PO<sub>4</sub>, 1 g/l de Trizma base pH 8,2, 4 µg/l de vitamina B 12, 3,89 mg/l de FeCl<sub>3</sub> · 6 H<sub>2</sub>O, 34,3 mg/l de H<sub>3</sub>BO<sub>3</sub>, 4,3 mg/l de MnCl<sub>2</sub> · 4 H<sub>2</sub>O, 315 µg/l de ZnCl<sub>2</sub>, 30 µg/l de MoO<sub>3</sub>, 3 µg/l de CuSO<sub>4</sub> · 5 H<sub>2</sub>O, 12,2 µg/l de CoCl<sub>2</sub> · 6 H<sub>2</sub>O) 15 [Stevens SE, Patterson COP, y Myers J. "The production of hydrogen peroxide by green algae: a survey." J. Phycology (1973). 9:427-430] más 5 g/l de NaNO<sub>3</sub>, a aproximadamente 10<sup>8</sup> células/ml. Se mezclan nueve volúmenes de células con un volumen de 1-10 µg/ml de ADN en NaCl 0,15 M/citrato Na<sub>3</sub> 0,015 M y se incuban a 27-30°C durante 3 horas antes de la adición de un volumen de DNase 1 a una concentración final de 10 µg/ml. Las 20 células se siembran en placas en 2,5 ml de agar recubierto con medio A al 0,6% que se templó a 45°C y se incuban. Las células se estimulan con antibiótico por refuerzo de 2,0 ml de agar con medio A al 0,6% que contiene una concentración apropiada de antibiótico con una pipeta Pasteur estéril. Los transformantes se pican 3-4 días 25 después. Las selecciones se realizan típicamente usando 200 µg/ml de kanamicina, 8 µg/ml de cloranfenicol, 10 µg/ml de espechinomicina en medio sólido, mientras que se emplean 150 µg/ml de kanamicina, 7 µg/ml de cloranfenicol, y 5 µg/ml de espechinomicina en medio líquido.

25 Las células *Thermosynechococcus elongatus* BP-1 se transforman de acuerdo con el protocolo optimizado previamente descrito [Iwai M, Katoh H, Katayama M, y Ikeuchi M. "Improved genetic transformation of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1, Plant Cell Physiol (2004). 45(2):171-175]. Se incuban cultivos en fase semi-exponencial con ADN exógeno (típicamente 3-20 µg) y se someten a electroporación en una fuerza de campo de 10 kV/cm usando un BioRad Gene Pulsar Xcell (Bio-Rad Laboratories, Hercules, CA). Después de la electroporación, las células se recuperan en 1 ml de medio BG11 durante 24 horas a 45°C. Las 30 células se siembran directamente en placas BG11 que contienen el antibiótico apropiado u opcionalmente se premezclan con medio BG11 que contiene agar fundido al 0,35% (p/v) (Difco, Estados Unidos).

35 La transformación típicamente implica la incubación de las células receptoras con ADN plasmídico purificado aislado de *E. coli*. En contraste, la conjugación bacteriana proporciona un medio alternativo para transferir directamente ADN de una especie bacteriana a otra. Por ejemplo, se han descrito técnicas que posibilitan la conjugación entre *E. coli* y cianobacterias incluyendo *Synechocystis* PCC 6803 y PCC 6714 y *Synechococcus* PCC 7942 y PCC 6301, así como *Synechococcus elongatus* termófilo [Marraccini P, Bulteau S, Cassier-Chauvat C, Mermet-Bouvier P, y Chauvat F. "A conjugative plasmid vector for promoter analysis in several cyanobacteria of the genera *Synechococcus* and *Synechocystis*." Plant Molecular Biology (1993). 23(4):905-909; Muhlenhoff U y Chauvat; "Gene 40 transfer and manipulation in the thermophilic cyanobacterium *Synechococcus elongatus*." Molecular and General Genetics MGG (1996). 252(1-2):93-100].

45 Las Tablas 1 y 2 definen genes preferidos para transmitir las propiedades hiperfotosintéticas a un organismo fotoautotrófico existente.

50 La Tabla 1 enumera genes que se sobre-expresan para potenciar las tasas de fijación del carbono, la tolerancia térmica, la tolerancia al pH, la tolerancia a gases de combustión, la tolerancia salina, las eficacias de recolección de la luz, la generación de energía reductora, y la independencia de nutrientes, junto con información sobre vías asociadas, números de la Comisión de Enzimas (EC), nombres de genes ejemplares, organismo fuente, número de acceso a GenBank, y homólogos de fuentes alternativas. Cuando el organismo precursor codifica un gen con la actividad enzimática indicada, no obstante es útil sobre-expresar estos componentes para mejorar la fijación del CO<sub>2</sub>. En algunas circunstancias, la secuencia enzimática nativa puede sobre-expresarse. En otras circunstancias, es útil sobre-expresar un gen exógeno, que permita un control regulador más explícito en el bioproceso y un medio para mitigar potencialmente los efectos de la regulación del metabolismo central, que se centra alrededor de los genes 55 nativos explícitamente.

60 Las secuencias de nucleótidos para los genes indicados (o secuencias de ADN que codifican los polipéptidos idénticos u homólogos, pero que abarcan sustituciones de nucleótidos para 1) alterar los niveles de expresión en base a la tabla de uso de codones del organismo hospedador; 2) añadir o eliminar estructuras secundarias; 3) añadir o eliminar secuencias de reconocimiento de endonucleasas de restricción; y/o 4) facilitar la síntesis y ensamblaje de genes) se ensamblan por Codon Devices Inc (Cambridge, MA). Se usan, según se ha indicado, proveedores alternativos incluyendo DNA2.0 (Menlo Park, CA), Blue Heron Biotechnology (Bothell, WA), y Geneart (Regensburg, Alemania). Las secuencias insostenibles por fuentes comerciales pueden prepararse usando reacción en cadena de la polimerasa (PCR) a partir de muestras de ADN o ADNc, o bibliotecas de ADNc/BAC. Los insertos se propagan inicialmente y se secuencian en un vector de clonación, tal como pUC19. De forma importante, la síntesis primaria y

la verificación de secuencia del gen de interés en pUC19 proporciona flexibilidad para transferir cada unidad en diversas combinaciones para que vectores de destino alternativos dirijan la transcripción y traducción de las enzimas deseadas. Se incluyen sitios de clonación específicos y/o únicos en los extremos 5' y 3' de las fases de lectura abierta (ORF) para facilitar las transferencias moleculares.

- 5 Las vías metabólicas necesarias se codifican inicialmente en cassetes de expresión dirigidos por promotores constitutivos que están siempre "activados". Muchos de estos promotores son conocidos, por ejemplo, el operón de la proteína ribosómica *spc* ( $P_{spc}$ ), el promotor del gen de la beta-lactamasa de pBR322 ( $P_{bla}$ ), el promotor del bacteriófago lambda  $P_L$ , los promotores de control de la replicación del plásmido pBR322 ( $P_{RNAl}$  o  $P_{RNAlI}$ ), o los promotores  $P_1$  o  $P_2$  del operón de ARN ribosómico *rrnB* [Liang ST, Bipatnath M, Xu YC, Chen SL, Dennis P, Ehrenber M, Bremer H. Activities of Constitutive Promoters in Escherichia coli. J. Mol Biol (1999). Vol. 292, Número 1, pág. 19-37], los promotores del virus *Chlorella* descritos en la patente de Estados Unidos 5.846.744 ["Chlorella virus promoters"], el promotor del virus del mosaico de la coliflor 35S [Zheng X, Deng W, Luo K, Duan H, Chen Y, McAvoy R, Song S, Pei Y, Li Y. "The cauliflower mosaic virus (CaMV) 35S promoter sequence alters the level and patterns of activity of adjacent tissue-and organ-specific gene promoters." Plant Cell Rep (2007). 26(8): 11195-1203], el promotor constitutivo *petH* de *Anabaena* [Valladares A, Muro-Pastor AM, Fillat MF, Herrero A, Flores E. "Constitutive and nitrogen-regulated promoters of the petH gene encoding ferredoxin:NADP+ reductase in the heterocyst-forming cyanobacterium Anabaena sp." FEBS Lett (1999). 449(2-3): 159-64], el promotor *RbcS2* de *Chlamydomonas* [Leon R, Couso I, Fernandez E. "Metabolic engineering of ketocarotenoids biosynthesis in the unicellular microalgae Chlamydomonas reinhardtii." J Biotechnol (2007). 130(2): 143-152], y la secuencia promotora central de *psbA* de *Microcystis aeruginosa* K-81 [Shibato J, Asayama M, Shirai M. "Specific recognition of the cyanobacterial *psbA* promoter by RNA polymerases containing principal sigma factors." Biochim. Biophys Acta (1998). 1442(2-3): 296-303].
- 10 25 Según sea necesario, tras diseñar y ensayar las vías, se "afina" la potencia de los promotores constitutivos para aumentar o disminuir los niveles de transcripción para optimizar una red, por ejemplo, modificando los elementos conservados -35 y -10 o el espaciado entre estos elementos [Alper H, Fischer C, Nevoigt E, Stephanopoulos G. "Tuning genetic control through promoter engineering." PNAS (2005). 102(36):12678-12783; Jensen PR y Hammer K. "The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters." Appl Environ Microbiol (1998). 64(1):82-87; Mijakovic I, Petranovic D, Jensen PR. Tunable promoters in system biology. Curr Opin Biotechnol (2005). 16: 329-335; De Mey M, Maertens J, Lequeux GJ, Soetaert WK, Vandamme EJ. "Construction and model-based analysis of a promoter library from E. coli: an indispensable tool for metabolic engineering." BMC Biotechnology (2007) 7:34].
- 15 30 35 40 45 50 55 60 65 70 75 80 85 90 95
- Cuando se demuestra una expresión constitutiva no óptima (es decir, tiene efectos perjudiciales, está desincronizada con la red, etc.) se usan promotores inducibles. Los promotores Inducibles están "inactivados" (no transcritos) antes de la adición de un agente inductor, frecuentemente una molécula pequeña o metabolito. Ejemplos de sistemas promotores inducibles adecuados incluyen el promotor inducible de arabinosa  $P_{bad}$  [Khlebnikov A, Datsenko KA, Skaug T, Wanner BL, Keasling JD. "Homogeneous expression of the P(BAD) promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter." Microbiology (2001). 147 (Pt 12): 3241-7], el promotor inducible de rhamnosa  $rhaP_{BAD}$  [Haldimann A, Daniels L, Wanner B. J Bacteriol (1998). "Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the Escherichia coli phosphate regulon." 180: 1277-1286], el promotor inducible de propionato pPRO [Lee SK y Keasling JD. "A propionate-inducible expression system for enteric bacteria." Appl Environ Microbiol (2005). 71 (11):6856-62)], el promotor *lac* inducible por IPTG [Gronenborn. Mol Gen Genet (1976). "Overproduction of phage lambda repressor under control of the lac promoter of Escherichia coli." 148:243-250], el promotor sintético *tac* [De Boer HA, Comstock LJ, Vasser M. "The tac promoter: a functional hybrid derived from the trp and lac promoters." PNAS (1983). 80:21-25], el promotor sintético *trc* [Brosius J, Erfle M, Storella J. "Spacing of the -10 y -35 regions in the tac promoter. Effect on its in vivo activity." J Biol Chem (1985). 260:3539-3541], o el sistema de la ARN polimerasa T7 [Studier FW y Moffatt BA. "Use of bacteriophage T7 RNA polimerase to direct selective high-level expression of cloned genes." J Mol Biol (1986). 189:113-130], el sistema promotor/operador *tetA* inducible por tetraciclina o anhidrotetraciclina [Skerra A. "Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli" Gene (1994). 151:131-135], los promotores *Cpx1* y *Cyc6* inducibles por níquel de *Chlamydomonas reinhardtii* [Quinn JM, Kropat J, Merchant S. "Copper response element and Crr1-dependent nickel-responsive promoter for induced, reversible gene expression in Chlamydomonas reinhardtii." Eukaryotic Cell (2003). 2 (5):995-1002], el promotor *nirA* inducible por nitrito de *Synechococcus* [Qi Q, Hao M, Ng W, Slater SC, Baszis SR, Weiss JD, y Valentin HE. "Application of the Synechococcus *nirA* promoter to establish an inducible expression system for engineering the Synechocystis tocopherol pathway." Appl. Environ. Microb (2005) 71(10): 5678-5684], el promotor de la arilsulfatasa sensible a azufre de *Volvox* [Hallman A y Sumper M. "Reporter genes and highly regulated promoters as tools for transformation experiments in Volvox carteri." Proc Natl Acad Sci (1994). 91 (24):11562-11566]. Se emplean éstos y otros promotores inducibles de origen natural u obtenidos sintéticamente (véase, por ejemplo, la patente de Estados Unidos Nº 7.235.385; Methods for enhancing expression of recombinant proteins).
- Se seleccionan orígenes de replicación alternativos para proporcionar capas adicionales de control de la expresión. La cantidad de copias por célula contribuye al "efecto de dosificación génica." Por ejemplo, se usan los orígenes de elevadas copias pMB1 o *colE1* para generar 300-1000 copias de cada plásmido por célula, lo que contribuye a un

elevado nivel de expresión génica. En contraste, los plásmidos que codifican orígenes de bajas copias, tales como pSC101 o p15A, se aprovechan para restringir la cantidad de copias a aproximadamente 1-20 copias por célula. Las técnicas y secuencias para modular adicionalmente la cantidad de copias de plásmidos son conocidas (véase, por ejemplo, la patente de Estados Unidos Nº 5.565.333, Plasmid replication origin increasing the copy number of

5 plasmid containing said origin; la patente de Estados Unidos Nº 6.806.066, Expression vectors with modified ColE1 origin of replication for control of plasmid copy number). Ciertos organismos, incluyendo *Synechococcus* sp. PCC 7002, codifican plásmidos endógenos con cantidades variables de copias [Yano S, Kawata Y, Kojima H. "Salinity-dependent copy number changes of endogenous plasmids in *Synechococcus* sp. PCC 7002]; por tanto, la dosificación génica puede modificarse dirigiendo los cassetes de expresión a distintos plásmidos endógenos.

10 Los niveles de expresión también se optimizan por modulación de la eficacia de traducción. En *E. coli*, una secuencia de Shine-Dalgarno (SD) [Shine J y Dalgarno L. Nature (1975) "Determination of cistron specificity in bacterial ribosomes." 254(5495):34-8] es un secuencia consenso que dirige el ribosoma al ARNm y facilita el inicio 15 de la traducción alineando el ribosoma con el codón de inicio. La modulación de la secuencia SD se usa para aumentar o disminuir la eficacia de traducción según sea apropiado [de Boer HA, Comstock LJ, Hui A, Wong E, Vasser M. Gene Amplif Anal (1983). "Portable Shine-Dalgarno regions; nucleotides between the Shine-Dalgarno sequence and the start codon effect the translation efficiency." 3:103-16; Mattanovich D, Weik R, Thim S, Kramer W, Bayer K, Katinger H. Ann NY Acad Sci (1996). "Optimization of recombinant gene expression in Escherichia coli." 782:182-90]. Es de señalar que puede observarse un elevado nivel de traducción en ciertos contextos en ausencia 20 de una secuencia SD [Mutsuda M y Sugiura M. "Translation initiation of cyanobacterial rbcS mRNAs requires the 38-kDa ribosomal protein S1 but not the Shine-Dalgarno sequence." J Biol Chem (2006). 281(50):38314-38321; Xu J, Mironova R, Ivanov IG, Abouhaider MG. J Basic Microbiol (1999). "A polilinker-derived sequence, PL, highly increased translation efficiency in Escherichia coli." 39 (1):51-60]. La estructura secundaria del ARNm se modifica 25 por ingeniería dentro o fuera de los genes de interés para modular los niveles de expresión [Klinkert B, Elles I, Nickelsen J. "Translation of chloroplast psbD mRNA in Chlamydomonas is controlled by a secondary structure blocking the AUG start codon." Nucleic Acids Res (2006). 34(1):384-94; Cebe R y Geiser M. Protein Expr Purif (2006). "Rapid and easy thermodynamic optimization of 5'-end of mRNA dramatically increases the level of wild type protein expression in Escherichia coli." 45(2):374-80; Zhang W, Xiao W, Wei H, Zhang J, Tian Z. Biochem Biophys Res Commun (2006). "mRNA secondary structure at start AUG codon is a key limiting factor for human protein 30 expression in Escherichia coli." 349(1):69-78; Voges D, Watzele M, Nemetz C, Wizemann S, Buchberger B. Biochem Biophys Res Commun (2004). "Analyzing and enhancing mRNA translational efficiency in an Escherichia coli in vitro expression system." 318(2):601-14]. El uso de codones también se manipula para aumentar o disminuir los niveles 35 de traducción [Deng T. FEBS Lett (1997). "Bacterial expression and purification of biologically active mouse c-Fos proteins by selective codon optimization." 409(2): 269-72; Hale RS y Thompson G. Protein Expr Purif (1998). "Codon optimization of the gene encoding a domain from human type 1 neurofibromin protein results in a threefold improvement in expression level in Escherichia coli." 12(2): 185-8].

En algunas realizaciones, cada gen de interés se expresa en un plásmido único. En realizaciones preferidas, las vías biosintéticas deseadas se codifican en vectores plasmídicos multi-cistrónicos. Los vectores de expresión útiles se 40 diseñan internamente y se sintetizan por proveedores externos de síntesis génica.

#### Optimizaciones

45 Las siguientes vías biosintéticas y módulos se ensayan primero y se optimizan usando plásmidos episómicos descritos anteriormente. Optimizaciones no limitantes incluyen intercambio y afinado del promotor, manipulación del sitio de unión al ribosoma, alteración del orden génico (por ejemplo, gen *ABC* frente a *BAC*, *CBA*, *CAB*, *BCA*), co-expresión de chaperonas moleculares, mutagénesis aleatoria o dirigida de secuencias génicas para aumentar o disminuir la actividad, el plegamiento, o la regulación alostérica, expresión de secuencias génicas a partir de especies alternativas, manipulación de codones, adición o eliminación de secuencias de dirección intracelular tales 50 como secuencias señal, y similares.

Cada gen o módulo se optimiza individualmente, o como alternativa, en paralelo. El promotor funcional y las secuencias génicas se integran posteriormente en el cromosoma del hospedador fotoautotrófico para posibilitar la propagación estable en ausencia de presión selectiva (es decir, inclusión de antibióticos) usando técnicas convencionales conocidas para los especialistas en la técnica.

60 La Tabla 2 enumera los genes que pueden regularse negativamente o eliminarse para potenciar las tasas de fijación del carbono, tolerancia térmica, tolerancia al pH, tolerancia de gases de combustión, tolerancia salina, eficacias de recolección de la luz, generación de energía reductora, e independencia de nutrientes, junto con información sobre las vías asociadas, los números de la Comisión de Enzimas (EC), los nombres de genes ejemplares, el organismo fuente, y los números de acceso a GenBank (no de la invención).

#### Selecciones y ensayos

65 La presión selectiva proporciona un medio valioso para ensayar y optimizar los organismos hiperfotosintéticos. La capacidad de sobrevivir y replicarse en medios que carecen de vitamina B<sub>12</sub> como suplemento del medio confirma la

aplicación satisfactoria del módulo de vitamina B<sub>12</sub>.

Si se desea, puede introducirse una variación genética adicional antes de la presión selectiva por tratamiento con mutágenos, tales como luz ultra-violeta, alquilantes [por ejemplo, metanosulfonato de etilo (EMS), metano sulfonato de metilo (MMS), dietilsulfato (DES), y nitrosoguanidina (NTG, NG, MMG)], intercaladores de ADN (por ejemplo, bromuro de etidio), ácido nitroso, análogos de bases, bromouracilo, transposones, y similares.

Como alternativa o además de la presión selectiva, la actividad de la vía puede controlarse siguiendo el crecimiento en condiciones permisivas (es decir, no selectivas) midiendo el rendimiento de producto específico mediante diversos estudios de marcaje metabólico (incluyendo radiactividad), análisis bioquímicos (Michaelis-Menten), espectrometría de masas por cromatografía de gases (GC/MS), espectrometría de masas, espectrometría de masas por desorción e ionización por láser asistida por matriz-tiempo de vuelo (MALDI-TOF), electroforesis capilar (CE), y cromatografía líquida a alta presión (HPLC).

## 15 Métodos de fermentación

La producción y aislamiento de productos a partir de organismos hiperfotosintéticos puede potenciarse empleando técnicas de fermentación específicas. Un elemento esencial para maximizar la producción reduciendo al mismo tiempo los costes, es aumentando el porcentaje de la fuente de carbono que se convierte en dichos productos. Los 20 átomos de carbono, durante ciclos de vida celulares normales, van a funciones celulares incluyendo producción de lípidos, sacáridos, proteínas, y ácido nucleicos. Reduciendo la cantidad de carbono necesaria para actividades no relacionadas con productos, se puede aumentar la eficacia de la producción saliente. Esto se consigue cultivando primero los microorganismos hasta una densidad deseada. Una densidad preferida sería la conseguida en el pico de la fase log de crecimiento. En dicho punto, pueden aprovecharse genes de control de la replicación para detener el crecimiento de las células. Específicamente, pueden usarse mecanismos que detectan los requisitos (revisado en Camilli, A. y Bassler, B.L. Science 311:1113; Venturi, V. FEMS Microbiol Rev 30: 274; y Reading, N.C. y Sperandio, V. FEMS Microbiol Lett 254:1) para activar genes tales como p53, p21, u otros genes de control. Los genes que 25 pueden activarse para detener la replicación y crecimiento celular en *E. coli* incluyen genes *umuDC*, cuya sobre-expresión detiene el progreso de la fase exponencial en crecimiento estacionario (Murli, S., Opperman, T., Smith, B.T., y Walker, G.C. 2000 Journal of Bacteriology 182: 1127.). UmuC es una ADN polimerasa que puede realizar la 30 síntesis translesión sobre lesiones no codificantes - la base mechanística de la mayoría de las mutagénesis por UV y químicas. Los productos del gen *umuDC* son necesarios para el proceso de síntesis translesión y también sirven como control del daño al ADN. Los productos del gen UmuDC incluyen UmuC, UmuD, umuD', UmuD'2C, UmuD'2 y 35 UmuD<sub>2</sub>. Simultáneamente, se activan los genes de síntesis de producto, minimizando de este modo la necesidad de usar vías críticas de replicación y mantenimiento mientras se está produciendo el producto.

Como alternativa, el crecimiento celular y la producción de producto pueden conseguirse simultáneamente. En este método, las células se cultivan en biorreactores con un suministro continuo de aportaciones y retirada continua de producto. Las fermentaciones discontinuas, semicontinuas, y continuas son habituales y bien conocidas en la técnica 40 y pueden hallarse ejemplos en Thomas D. Brock en Biotechnology: A Textbook of Industrial Microbiology, Segunda Edición (1989) Sinauer Associates, Inc., Sunderland, Mass., o Deshpande, Mukund V., Appl. Biochem. Biotechnol (1992), 36:227.

En todos los métodos de producción, las aportaciones incluyen dióxido de carbono, agua, y luz. El dióxido de carbono puede ser de la atmósfera o de fuentes concentradas incluyendo gas residual o gases de combustión de plantas carboníferas, refinerías, instalaciones de producción de cemento, instalaciones de gas natural, destilerías, y similares. El agua puede ser no salina, de bajo contenido salino, marina, o de alto contenido salino. La luz puede ser solar o de fuentes artificiales incluyendo luces incandescentes, LED, fibra óptica, y luces fluorescentes.

Los organismos recolectores de luz están limitados en su productividad a los momentos en que la radiación solar es suficiente para actividad sus fotosistemas. En un bioprocessos preferido con organismo recolector de luz, se facilita que las células crezcan y produzcan producto con luz como impulsor energético. Cuando hay ausencia de luz suficiente, puede inducirse a las células a minimizar su tasa metabólica central. Para este fin, los promotores 50 inducibles específicos para la producción de productos pueden estimularse fuertemente para dirigir a la célula a procesar sus almacenes energéticos en el producto de elección. Con suficiente fuerza de inducción, la célula minimizará sus esfuerzos de crecimiento, y usa sus reservas de recolección de luz específicamente para la producción de producto. No obstante, se espera que la productividad neta sea mínima durante períodos en que se carece de luz suficiente ya que se capturan ninguno a pocos fotones.

En una realización preferida, la célula se modifica por ingeniería de tal modo que el producto final se libere de la célula. En realizaciones en que el producto final se libera de la célula, puede emplearse un proceso continuo. En este enfoque, puede ensamblarse un reactor con organismos productores de productos deseables en múltiples formas. En una realización, el reactor se hace funcionar en bloque de forma continua, con una parte del medio retirado y mantenido en un entorno menos agitado de modo que el producto acuoso se autoseparará con el producto 60 retirado y el resto vuelve a la cámara de fermentación. En realizaciones en las que el producto no se separa en una fase acuosa, el medio se retira y se emplean técnicas de separación apropiadas (por ejemplo, cromatografía,

destilación, etc.).

En una realización alternativa, el producto no se secreta por las células. En esta realización, se emplea un enfoque de fermentación semicontinua. En dichos casos, las células se cultivan en exposición continuada a las aportaciones

5 (luz, agua, y dióxido de carbono) como se ha especificado anteriormente hasta que la cámara de reacción se satura con células y producto. Una parte significativa de la totalidad del cultivo se retira, se lisan las células, y se aislan los productos por técnicas apropiadas de separación (por ejemplo, cromatografía, destilación, filtración, centrifugación, etc.).

10 En una realización preferida, la cámara de fermentación encerrará una fermentación que está experimentando una fermentación reductora continua. En este caso, se crea un entorno reductor estable. El equilibrio de electrones se mantiene por la liberación de dióxido de carbono (en forma gaseosa). Aumentar el equilibrio NAD/H y NADP/H, como se ha descrito anteriormente, también puede ser de ayuda para estabilizar el equilibrio de electrones.

15 **Detección y análisis de productos génicos y celulares**

Cualquiera de los métodos analíticos convencionales, tales como espectrometría de masas por cromatografía de gases, y espectrometría de masas por cromatografía líquida, HPLC, electroforesis capilar, espectrometría de masas por desorción e ionización por láser asistida por matriz-tiempo de vuelo, etc., puede usarse para analizar los niveles

20 y la identidad del producto producido por los organismos modificados de la presente invención.

La capacidad para detectar la formación de una nueva vía bioquímica funcional en la célula hiperfotosintética es importante para la práctica de los presentes métodos. En general, los ensayos se realizan para detectar reacciones de transformación bioquímica heterólogas de la célula hospedadora que producen, por ejemplo, moléculas orgánicas

25 pequeñas y similares como parte de una vía de síntesis *de novo*, o por modificación química de moléculas de forma ectópica proporcionadas en el entorno de la célula hospedadora. La generación de dichas moléculas por la célula hospedadora puede detectarse en "extractos de ensayo", que pueden ser medios condicionados, lisados celulares, membranas celulares, o productos de fraccionamiento semi-purificados o purificados de los mismos. Lo último puede prepararse, como se ha descrito anteriormente, por técnicas clásicas de fraccionamiento/purificación, incluyendo

30 separación de fases, separación cromatográfica, o fraccionamiento de en disolvente (por ejemplo, metanol etanol, acetona, acetato de etilo, tetrahidrofurano (THF), acetonitrilo, benceno, éter, sales de bicarbonato, diclorometano, cloroformo, éter de petróleo, hexano, ciclohexano, éter dietílico y similares). Cuando el ensayo se establece con una célula respondedora para ensayar el efecto de una actividad producida por la célula hospedadora sobre una célula completa en lugar de un fragmento celular, la célula hospedadora y la célula de ensayo pueden co-cultivarse juntas (opcionalmente separadas por un inserto de cultivo, por ejemplo Collaborative Biomedical Products, Bedford, MA, nº de Catálogo 40446).

35 En ciertas realizaciones, el ensayo se establece para detectar directamente, por técnicas químicas o fotométricas, una especie molecular que se produce (o destruye) por una vía biosintética de la célula hospedadora recombinante.

40 Dicha producción o degradación de la especie molecular debe depender, al menos en parte, de la expresión del ADN genómico heterólogo. En otras realizaciones, la etapa de detección del presente método implica la caracterización de medios fraccionados/lisados celulares (el extracto de ensayo), o la aplicación del extracto de ensayo a un sistema de detección bioquímico o biológico. En otras realizaciones, el ensayo detectar indirectamente la formación de productos de una vía heteróloga observando un cambio fenotípico en la célula hospedadora, por ejemplo, de un modo autocrino, que depende del establecimiento de una vía biosintética heteróloga en la célula hospedadora.

50 En ciertas realizaciones, se buscan análogos relacionados con una clase conocida de compuestos, como por ejemplo análogos de alcaloides, aminoglucósidos, ansamacrólidos, beta-lactamas (incluyendo penicilinas y cefalosporinas), carbapenems, terpinoides, hormonas prostanoïdes, azúcares, ácidos grasos, lincosaminidas, macrólidos, nitrofuranos, nucleósidos, oligosacáridos, oxazolidinonas, péptidos y polipéptidos, fenazinas, polienos, poliésteres, quinolonas, tetraciclinas, estreptograminas, sulfonamidas, esteroides, vitaminas y xantinas. En dichas realizaciones, si existe un ensayo disponible para identificar directamente y/o aislar el producto natural, y se espera que los análogos se comporten de forma similar en esas condiciones, la etapa de detección del presente método

55 puede ser tan sencilla como la detección directa de análogos de interés en el medio de cultivo celular o la preparación de la célula. Por ejemplo, puede realizarse separación cromatográfica u otra separación bioquímica de un extracto de ensayo, y detectarse la presencia o ausencia de un análogo, por ejemplo, espectrofotométricamente, en la fracción en que existirían los compuestos conocidos en condiciones similares. En ciertas realizaciones, dichos compuestos pueden tener una fluorescencia característica o una fosforescencia que puede detectarse sin ninguna necesidad de fraccionar el medio y/o la célula recombinante.

60 En realizaciones relacionadas, puede ensayarse el medio de cultivo o lisado completo o fraccionado procedente de una célula hospedadora recombinante poniendo en contacto la muestra de ensayo con una célula heteróloga ("célula de ensayo") o componentes de la misma. Por ejemplo, una célula de ensayo, que puede ser procariota o eucariota, se pone en contacto con medio condicionado (completo o fraccionado) procedente de una célula hospedadora recombinante, y se ensaya la capacidad del medio condicionado de inducir una respuesta biológica o

bioquímica procedente de la célula de ensayo. Por ejemplo, el ensayo puede detectar un cambio fenotípico en la célula de ensayo, como por ejemplo un cambio en: la velocidad de transcripción o traducción o el patrón de corte y ayuste de un gen; la capacidad de una proteína; la fosforilación, prenilación, metilación, glucosilación u otra modificación post-traduccional de una proteína, ácido nucleico o lípido; la producción de segundos mensajeros, tales

5 como AMPc, inositol fosfatos y similares. Dichos efectos pueden medirse directamente, por ejemplo, aislando y estudiando un componente particular de la célula, o indirectamente tal como por expresión de un gen indicador, la detección de marcadores fenotípicos, y actividad citotóxica o citostática en la célula de ensayo.

10 Cuando se explora la bioactividad de compuestos de ensayo producidos por las células hospedadoras recombinantes, puede medirse directamente la generación de segundo mensajero intracelular. Se ha identificado una diversidad de efectores intracelulares. Por ejemplo, para exploraciones pretendidas para aislar compuestos, o los genes que codifican los compuestos, que son inhibidores o potenciadores de eventos regulados por receptor o canal de iones, puede detectarse el nivel de producción de segundo mensajero a partir de las proteínas de señalización corriente abajo, tales como adenil ciclase, fosfodiesterasas, fosfoinositidasas, fosfoinositol quinasas, y 15 fosfolipasas, como también los niveles intracelulares de una diversidad de iones.

En otras realizaciones más, la señal detectable puede producirse por el uso de enzimas o sondas cromogénicas/fluorescentes cuyas actividades son dependientes de la concentración de un segundo mensajero, por ejemplo, tal como calcio, productos de la hidrólisis de inositol fosfato, AMPc, etc.

20 Muchos genes indicadores y elementos reguladores de la transcripción son conocidos para los especialistas en la técnica y otros pueden identificarse o sintetizarse por métodos conocidos para los especialistas en la técnica. Ejemplos de genes indicadores incluyen, aunque sin limitación, CAT (cloranfenicol acetil transferasa) (Alton y Vapnek (1979), Nature 282: 864-869) luciferasa, y otros sistemas de detección enzimáticos, tales como betagalactosidasa; luciferasa de luciérnaga (de Wet et al. (1987), Mol. Cell. Biol. 7:725-737); luciferasa bacteriana (Engebrecht y Silverman (1984), PNAS 1: 4154-4158; Baldwin et al. (1984), Biochemistry 23: 3663-3667); fosfatasa alcalina (Toh et al. (1989) Eur. J. Biochem. 182:231-238, Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), fosfatasa alcalina secretada placentaria humana (Cullen y Malim (1992) Methods in Enzymol. 216:362-368); β-lactamasa o GST.

30 Los elementos de control de la transcripción para su uso en las construcciones génicas indicadoras, o para modificar el locus genómico de un gene indicador incluyen, aunque sin limitación, promotores, potenciadores, y sitios de unión represores y activadores. Los elementos reguladores de la transcripción adecuados pueden obtenerse de las regiones reguladoras de la transcripción de genes cuya expresión se induce rápidamente, generalmente en minutos, del contacto entre la proteína de la superficie celular y la proteína efectora que modula la actividad de la proteína de superficie celular. Ejemplos de dichos genes incluyen, aunque sin limitación, los genes tempranos inmediatos (véase, Sheng et al. (1990) Neuron 4: 477-485), tal como c-fos. Los genes tempranos inmediatos son genes que se inducen rápidamente tras la unión de un ligando a una proteína de superficie celular. Los elementos de control de la transcripción que se prefieren para su uso en las construcciones génicas incluyen elementos de control de la 35 transcripción de genes tempranos inmediatos, elementos obtenidos de otros genes que muestran alguna o todas las características de los genes tempranos inmediatos, o elementos sintéticos que se construyen de modo que genes en unión funcional con los mismos muestren dichas características. Las características de genes preferidos a partir de los cuales se obtiene los elementos de control de la transcripción incluyen, aunque sin limitación, expresión baja o indetectable en células quiescentes, inducción rápida a nivel transcripcional en minutos de estimulación extracelular, 40 inducción que es transitoria e independiente de nueva síntesis proteica, la posterior desactivación de la transcripción requiere nueva síntesis proteica, y los ARNm transcriptos a partir de estos genes tienen una corta semi-vida. No es necesario que estén presentes todas estas propiedades.

45 En otras realizaciones más, la etapa de detección se proporciona en forma de un sistema sin células, por ejemplo, una preparación de lisado celular o de proteína o ácido nucleico purificada o semi-purificada. Las muestras obtenidas de las células hospedadoras recombinantes pueden ensayarse para actividades tales como inhibición o potenciación de dichos complejos por pares (el "complejo diana") que implican interacciones proteína-proteína, interacciones proteína-ácido nucleico, interacciones proteína-ligando, interacciones ácido nucleico-ácido nucleico, y similares. El ensayo puede detectar la ganancia o pérdida de los complejos diana, por ejemplo, por actividades 50 endógenas o heterólogas asociadas con una o ambas moléculas del complejo.

55 A menudo se prefieren ensayos que se realizan en sistemas sin células, tales como los que se pueden obtener con proteínas purificadas o semi-purificadas, como exploraciones "primarias" porque pueden generarse para permitir un rápido desarrollo y una detección relativamente fácil de una alteración en una diana molecular cuando se pone en contacto con una muestra de ensayo. Además, los efectos de toxicidad celular y/o biodisponibilidad de la muestra de ensayo pueden ignorarse generalmente en el sistema *in vitro*, centrándose principalmente el ensayo, en cambio, en el efecto de la muestra sobre al diana molecular ya que puede manifestarse en una alteración de la afinidad de unión con otras moléculas o cambios en las propiedades enzimáticas (si fuera aplicable) de la diana molecular. La detección y cuantificación de los complejos por pares proporciona un medio para determinar la eficacia de las muestras de ensayo en la inhibición (o potenciación) de la formación de complejos. La eficacia del compuesto puede 60 valorarse generando curvas de respuesta a dosis a partir de datos obtenidos usando diversas concentraciones de la

muestra de ensayo. Además, también puede realizarse un ensayo de control para proporcionar una medida inicial para la comparación. Por ejemplo, en el ensayo de control puede añadirse medio condicionado de las células hospedadoras no transformadas.

- 5 La cantidad de complejo diana puede detectarse por una diversidad de técnicas. Por ejemplo, la modulación en la formación de complejos puede cuantificarse usando, por ejemplo, proteínas marcadas de forma detectable o similares (por ejemplo, radiomarcadas, marcadas de forma fluorescente, o marcadas enzimáticamente), por inmunoensayo, o por detección cromatográfica.
- 10 En otras realizaciones más, puede usarse una enzima purificada o semi-purificada para ensayar las muestras de ensayo. La capacidad de una muestra de ensayo de inhibir o potenciar la actividad de la enzima puede detectarse convenientemente siguiendo la velocidad de conversión de un sustrato para la enzima.
- 15 En otras realizaciones más, la etapa de detección puede diseñarse para detectar un cambio fenotípico en la célula hospedadora que se induce por productos de la expresión de las secuencias genómicas heterólogas. Muchos de los formatos de ensayo basados en células mencionados anteriormente también pueden usarse en la célula hospedadora, por ejemplo, de un modo tipo autocrino.
- 20 Además de proporcionar una base para aislar moléculas biológicamente activas producidas por las células hospedadoras recombinantes, la etapa de detección también puede usarse para identificar clones genómicos que incluyen genes que codifican vías biosintéticas de interés. Además, por métodos de subclonación iterativa y/o combinatoria que dependen de dichas etapas de detección, pueden clonarse los genes individuales que confieren la vía detectada a partir del fragmento genómico más grande.
- 25 Los presentes métodos de exploración pueden realizarse en un formato diferencial, por ejemplo, comparando la eficacia de una muestra de ensayo en un ensayo de detección obtenido con componentes humanos con aquellos obtenidos de, por ejemplo, componentes fúngicos o bacterianos. Por tanto, la selectividad como bactericida o fungicida puede ser un criterio en el protocolo de selección.
- 30 La cepa hospedadora no tiene que producir elevados niveles de los nuevos compuestos para que el método sea satisfactorio. La expresión de los genes puede no ser óptima, pueden no estar presentes factores reguladores globales, o las combinaciones de metabolitos pueden no soportar la producción máxima del producto. La capacidad de detectar el metabolito a menudo no requerirá los niveles máximos de producción, particularmente cuando el bioensayo es sensible a pequeñas cantidades de productos naturales. Por tanto, la producción submáxima inicial de compuestos no tiene que ser una limitación para el éxito del presente método.

Finalmente, como se ha indicado anteriormente, la muestra de ensayo puede obtenerse de, por ejemplo, medios condicionados o lisados celulares. Con respecto a los últimos, se anticipa que en ciertos casos puede haber compuestos expresados de forma heteróloga que pueden no exportarse apropiadamente desde la célula hospedadora.

- 40 Existe una diversidad de técnicas disponible en la técnica para lisar células. Un enfoque preferido es el uso de un agente de lisis específico de célula hospedadora. Por ejemplo puede usarse un fago (por ejemplo, P1, λ, φ80) para lisar selectivamente *E. coli*. Asimismo, pueden usarse cianófagos para lisar selectivamente cianobacterias, tales como *Synechococcus* y *Prochlorococcus*. La adición dicho fago a cultivos de crecimiento de células hospedadoras puede maximizar el acceso a los productos heterólogos de nuevas vías biosintéticas en la célula. Además, dichos agentes no interfieren con el crecimiento de un organismo de ensayo, por ejemplo, una célula humana, que puede co-cultivarse con la biblioteca de células hospedadoras.

#### Optimización metabólica

- 50 Como parte del proceso de optimización, se proporcionan etapas para eliminar reacciones secundarias indeseables, si las hay, que puedan consumir el carbono y la energía pero que no produzcan productos útiles (tales como hidrocarburos, ésteres de cera, tensioactivos y otros productos de hidrocarburo). Estas etapas pueden ser útiles porque pueden ayudar a mejorar los rendimientos de los productos deseados.
- 55 Puede usarse una combinación de diferentes enfoques. Dichos enfoques incluyen, por ejemplo, metabolómica (que puede usarse para identificar productos indeseables e intermedios metabólicos que se acumulan en el interior de la célula), modelado metabólico y marcaje isotópico (para determinar el flujo a través de reacciones metabólicas que contribuyen a la producción de hidrocarburos), y técnicas genéticas convencionales (para eliminar o inutilizar sustancialmente reacciones metabólicas indeseadas). Por ejemplo, el modelado metabólico proporciona un medio para cuantificar los flujos a través de las vías metabólicas de la célula y determinar el efecto de la eliminación de etapas metabólicas clave. Además, la metabolómica y el modelado metabólico posibilitan una mejor comprensión del efecto de la eliminación de etapas metabólicas clave sobre la producción de productos deseados.
- 60 Para predecir cómo una manipulación particular del metabolismo afecta al metabolismo celular y la síntesis del producto deseado, se desarrolló un marco teórico para describir los flujos molares a través de todas las vías metabólicas conocidas de la célula. Varios aspectos importantes de este marco teórico incluyen: (i) una base de

datos relativamente completa de vías conocidas, (ii) la incorporación de la dependencia en la tasa de crecimiento de la composición celular y los requisitos energéticos, (iii) las mediciones experimentales de la composición de aminoácidos de las proteínas y la composición de ácidos grasos de las membranas a diferentes tasas de crecimiento y tasas de dilución y (iv) las mediciones experimentales de las reacciones secundarias que se sabe que suceden como resultado de la manipulación del metabolismo. Estos nuevos desarrollos permiten una predicción significativamente más precisa de los flujos en las vías metabólicas clave y la regulación de la actividad enzimática. (Keasling, J. D. et al., "New tools for metabolic engineering of *Escherichia coli*", en *Metabolic Engineering*, Publisher Marcel Dekker, Nueva York, Nym 1999; Keasling, J.D, "Gene-expression tools for the metabolic engineering of bacteria", *Trends in Biotechnology*, 17, 452-460, 1999; Martin, V. J. J., et al., "Redesigning cells for production of complex organic molecules", *ASM News* 68, 336-343 2002; Henry, C. S., et al., "Genome-Scale Thermodynamic Analysis of *Escherichia coli* Metabolism", *Biophys. J.*, 90, 1453-1461, 2006.)

Dichos tipos de modelos se han aplicado, por ejemplo, para analizar los flujos metabólicos en organismos responsables de la eliminación potenciada de fósforo biológico en reactores de tratamiento de aguas residuales y en hongos filamentosos productores de policétidos. Véase, por ejemplo, Pramanik, et al., "A stoichiometric model of *Escherichia coli* metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements." *Biotechnol. Bioeng.* 56, 398-421, 1997; Pramanik, et al., "Effect of carbon source and growth rate on biomass composition and metabolic flux predictions of a stoichiometric model." *Biotechnol. Bioeng.* 60, 230-238, 1998; Pramanik et al., "A flux-based stoichiometric model of enhanced biological fosforus removal metabolism." *Wat. Sci. Tech.* 37, 609-613, 1998; Pramanik et al., "Development and validation of a flux-based stoichiometric model for enhanced biological fosforus removal metabolism." *Water Res.* 33, 462-476, 1998.

### Productos

25 Los organismos hiperfotosintéticos de la presente invención pueden modificarse por ingeniería para producir categorías de productos, incluyendo aunque sin limitación, azúcares biológicos, productos de hidrocarburo, formas sólidas, y agentes farmacéuticos.

30 Los azúcares biológicos incluyen aunque sin limitación, glucosa, almidón, celulosa, hemicelulosa, glucógeno, xilosa, dextrosa, fructosa, lactosa, fructosa, galactosa, ácido urónico, maltosa, y policétidos. En realizaciones preferidas, el azúcar biológico puede ser glucógeno, almidón, o celulosa.

35 La celulosa es la forma más abundante de biomasa terrestre viva (Crawford, R. L. 1981, *Lignin biodegradation and transformation*, John Wiley y Sons, Nueva York.). La celulosa, especialmente las pelusas de algodón, se usa en la fabricación de nitrocelulosa. La celulosa es también el constituyente principal del papel. Los monómeros de celulosa (beta-glucosa) se unen juntos a través de enlaces 1,4 glucosídicos. La celulosa es una cadena recta (no existen espirales). En microfibrillas, los múltiples grupos hidróxido se unen por hidrógeno entre sí, manteniendo las cadenas firmemente juntas y contribuyendo a su elevada resistencia a tracción. Dado un material de celulosa, la parte que no se disuelve en una solución al 17,5% de hidróxido sódico a 20°C es alfa celulosa, que es celulosa verdadera; la parte que se disuelve y después precipita tras la acidificación es beta celulosa, y la proporción que se disuelve pero no precipita es gamma celulosa. La hemicelulosa es una clase de polisacárido de pared celular vegetal que puede ser cualquier de varios heteropolímeros. Éstos incluyen xilano, xiloglucano, arabinoxilano, arabinogalactano, glucuronoxilano, glucomanano, y galactomanano. Esta clase de polisacáridos se halla en casi todas las paredes celulares junto con la celulosa. La hemicelulosa es de menor peso que la celulosa, y no puede extraerse por agua caliente o agentes quelantes, pero puede extraerse por álcalis acuoso. Las cadenas poliméricas se unen a pectina y celulosa, formando una red de fibras entrecruzadas.

40 Hay esencialmente tres tipos de productos de hidrocarburo: (1) productos de hidrocarburo aromático, que tiene al menos un anillo aromático; (2) productos de hidrocarburo saturado, que carecen de enlaces dobles, triples o aromáticos; y (3) productos de hidrocarburo insaturado, que tienen uno o más dobles o triples enlaces entre los átomos de carbono. Un "producto de hidrocarburo" puede definirse adicionalmente como un compuesto químico que consta de C, H, y opcionalmente O, con un esqueleto de carbono y átomos de hidrógeno y oxígeno, unidos al mismo. El oxígeno puede unirse de forma sencilla o doble al esqueleto y puede unirse por hidrógeno. En el caso de éteres y ésteres, el oxígeno puede incorporarse en el esqueleto, y unirse por dos enlaces sencillo, a las cadenas de carbono. Un único átomo de carbono puede unirse a uno o más átomos de oxígeno. Los productos de hidrocarburo también pueden incluir los anteriores compuestos unidos a agentes biológicos incluyendo proteínas, coenzima A y acetil coenzima A. Los productos de hidrocarburo incluyen, aunque sin limitación, hidrocarburos, alcoholes, aldehídos, ácido carboxílico, éteres, ésteres, carotenoides, y cetonas.

45 50 55 60 65 Los productos de hidrocarburo también incluyen alcanos, alquenos, alquinos, dienos, isoprenos, alcoholes, aldehídos, ácidos carboxílicos, tensioactivos, ésteres de cera, agentes químicos poliméricos [polifthalato carbonato (PPC), poliéster carbonato (PEC), polietileno, polipropileno, poliestireno, polihidroxialcanoatos (PHA), poli-beta-hidroxibutirato (PHB), polilactida (PLA), y policaprolactona (PCL)], agentes químicos monoméricos [propilenglicol, etilenglicol, y 1,3-propanodiol, etileno, ácido acético, ácido butírico, ácido 3-hidroxipropanoico (3-HPA), ácido acrílico, y ácido malónico], y combinaciones de los mismos. En algunas realizaciones preferidas, los productos de hidrocarburo son alcanos, alcoholes, tensioactivos, ésteres de cera y combinaciones de los mismos. Otros productos

de hidrocarburo incluyen ácidos grasos, hidrocarburos unidos a acetil-CoA, carbohidratos unidos a acetil-CoA, e intermedios policétidos.

- 5 Los organismos recombinantes pueden modificarse por ingeniería para producir productos de hidrocarburo e intermedios sobre un amplio intervalo de tamaños. Alcanos específicos que pueden producirse incluyen, por ejemplo, etano, propano, butano, pentano, hexano, heptano, octano, nonano, decano, undecano, dodecano, tridecano, tetradecano, pentadecano, hexadecano, heptadecano, y octadecano. En realizaciones preferidas, los productos de hidrocarburo son octano, decano, dodecano, tetradecano, y hexadecano. Precursores de hidrocarburo tales como alcoholes que pueden producirse incluyen, por ejemplo, etanol, propanol, butanol, pentanol, hexanol, heptanol, octanol, nonanol, decanol, undecanol, dodecanol, tridecanol, tetradecanol, pentadecanol, hexadecanol, heptadecanol, y octadecanol. En realizaciones más preferidas, el alcohol se selecciona entre etanol, propanol, butanol, pentanol, hexanol, heptanol, octanol, nonanol, y decanol.
- 10 Se usan tensioactivos en una diversidad de productos, incluyendo detergentes y limpiadores, y también se usan como auxiliares para textiles, cuero y papel, en procesos químicos, en cosméticos y agentes farmacéuticos, en la industria alimentaria y en agricultura. Además, pueden usarse para ayudar en la extracción y aislamiento de petróleo crudo que resulta fuerte para los entornos de acceso o como emulsiones acuosas. Hay cuatro tipos de tensioactivos caracterizados por usos variados. Los tensioactivos aniónicos tienen actividad tipo detergente y generalmente se usan para aplicaciones de limpieza. Los tensioactivos catiónicos contienen hidrocarburos de cadena larga y a menudo se usan para tratar proteínas y polímeros sintéticos o son componentes de suavizantes textiles y acondicionadores capilares. Los tensioactivos anfotéricos también contienen hidrocarburos de cadena larga y se usan típicamente en champúes. Los tensioactivos no iónicos se usan generalmente en productos de limpieza.
- 15 20 25 30 35
- 20 Los hidrocarburos pueden producirse adicionalmente como biocombustibles. Un biocombustible es cualquier combustible que se obtenga de una fuente biológica - recientemente organismos vivos o sus sub-productos metabólicos, tales como estiércol de vacas. Un biocombustible puede definirse adicionalmente como combustible obtenido de un producto metabólico de un organismo vivo. Los biocombustibles preferidos incluyen, aunque sin limitación, biodiesel, biocrudo, etanol, "petróleo renovable", butanol, y propano.
- 30 Las formas sólidas de carbono incluyen, por ejemplo, carbón, grafito, grafeno, cemento, nanotubos de carbono, negro de carbono, diamantes, y perlas. Los sólidos de carbono puro tales como carbono y diamante son las formas sólidas preferidas.
- 35 Pueden producirse agentes farmacéuticos incluyendo, por ejemplo, taxol basado en isoprenoide y artemisinina, u oseltamivir.

### **Construcciones plasmídicas**

#### **Construcción del plásmido base pJB5**

- 40 El plásmidos base pJB5 se diseñó como un vector de expresión vacío para recombinación en *Synechococcus* sp. PCC 7002. Se diseñaron dos regiones de homología, la región de homología corriente arriba (UHR) y la región de homología corriente abajo (DHR) para flanquear el gen o genes clonados de interés. Estas regiones de homología de 500 pb corresponden a las posiciones 3301-3800 y 3801-4300 en el plásmido natural pAQ1 (acceso a Genbank NC\_005025) para UHR y DHR respectivamente. El promotor aadA, la secuencia génica, y el terminador se diseñaron para conferir resistencia a espectinomicina y estreptomicina a la construcción integrada. Para la expresión, se diseño pJB5 con el promotor del casete de resistencia a kanamicina *aph2* y el sitio de unión al ribosoma (RBS). Corriente abajo de este promotor y RBS, se diseñó e insertó el sitio de reconocimiento para endonucleasa de restricción para *NdeI* y *EcoRI*, así como los sitios para *XbaI*, *BamHI*, *SpeI* y *Pad*. Después del sitio *EcoRI*, se incluyó el terminador natural del gen de la alcohol deshidrogenasa de *Zymomonas mobilis* (*adhII*). Sitios de restricción *XbaI* convenientes flanquean la UHR y la DHR permitiendo la escisión del ADN pretendido para recombinación del resto del vector. Se construyeron pJB5, pJB6 y pJB7 por DNA2.0 (Menlo Park, CA).

### **Ejemplos**

#### **Ejemplo 1: Independencia de nutrientes**

- 55 60
- Además de CO<sub>2</sub> y luz, los organismos fotoautotróficos típicamente requieren fuentes inorgánicas de nutrientes y vitaminas. Los nutrientes requeridos se suplementan generalmente al medio de cultivo durante propagación a escala experimental de dichos organismos. Sin embargo, dichos nutrientes son prohibitivamente caros en el contexto de bioprocesamiento a escala industrial.

- 65
- El nitrógeno es un constituyente clave de una diversidad de macromoléculas celulares, incluyendo aminoácidos y nucleótidos. La modificación por ingeniería de fotoautótrofos para utilizar de forma eficaz fuentes económicas proporciona significativas ventajas económicas y prácticas.

Los fotoautótrofos pueden modificarse por ingeniería para fijar N<sub>2</sub>, que se halla presente en concentraciones de casi el 80% (v/v) en el aire y los gases de combustión. En este escenario, se sobre-expresan los genes necesarios para la fijación el nitrógeno, además de los genes necesarios para la síntesis de los cofactores necesarios y la proteína accesoria [Herrero A, Muro-Pastor AM, Flores E. J Bacteriol (2001) "Nitrogen Control in Cyanobacteria." 183(2): 411-425]. Dieciocho de dichos genes ejemplares se hallan en *Nostoc* sp PCC 7120: proteína de biosíntesis del cofactor FeMo (NifB), locus NP\_485557; [4Fe-4S]ferredoxina (FdxN), locus BAB77882; L-cisteína desulfurasa (NifS), EC 2.8.1.7, locus NP\_485499; proteína accesoria del grupo Fe (NifU), locus NP\_485498; nitrogenasa - subunidad Fe (NifH), EC 1.18.6.1, locus NP\_485497; nitrogenasa-subunidad alfa (NifD), EC 1.18.6.1, locus NP\_485484; nitrogenasa - subunidad beta (NifK), EC 1.18.6.1, locus NP\_485483; proteína de biosíntesis del cofactor FeMo (NifE), locus NP\_485481; proteína de biosíntesis del cofactor FeMo (NifN), locus NP\_485480; proteína accesoria del grupo FeS (NifX), locus NP\_485479; proteína accesoria del cofactor FeMo (NifW), locus NP\_485476; proteína accesoria del cofactor FeMo (HesA), locus NP\_485475; proteína accesoria del cofactor FeS (HesB), locus NP\_485474; ferredoxina específica de fijación de nitrógeno (FdxH), locus NP\_485473; piruvato-flavodoxina oxidoreductasa (NifJ), EC 1.2.7.1, locus NP\_486843; homocitrato sintasa (NifV), EC 2.3.3.14, locus NP\_485450; proteína accesoria del cofactor FeMo (NifZ), locus NP\_485451; y proteína accesoria del cofactor FeMo (NifT), locus NP\_485452). La sobre-expresión de los anteriores genes facilita que los organismos fotoautotróficos crezcan con N<sub>2</sub> gaseoso como única fuente de nitrógeno.

Además o como alternativa, los fotoautótrofos se modifican por ingeniería para asimilar nitrito, que es un componente traza de los gases de combustión. Las células deben modificarse por ingeniería para expresar un transportador de nitrito/nitrato y transmitirse con la capacidad de convertir nitrito en amoniaco. Secuencias génicas ejemplares que codifican transportadores activos de nitrato/nitrato se hallan en *Synechococcus* sp. PCC 6301 [Herrero A, Muro-Pastor AM, Flores E. J Bacteriol (2001) "Nitrogen Control in Cyanobacteria." 183(2): 411-425] y se sobre-expresan en células para importar nitratos y nitrito, por ejemplo: proteína de unión al sustrato del sistema de transporte de nitrato/nitrito tipo ABC (NrtA), locus YP\_171021; proteína permeasa del sistema de transporte de nitrato/nitrito tipo ABC (NrtB), locus YP\_171022; proteína de unión a ATP del sistema de transporte de nitrato/nitrito tipo ABC (NrtC), locus YP\_171023; proteína de unión a ATP del sistema de transporte de nitrato/nitrito tipo ABC (NrtD), locus YP\_171024). Como alternativa, se sobre-expresa individualmente el polipéptido del gen del transportador de nitrato/nitrito (NrtP) de *Synechococcus* sp. PCC 7002 [Sakamoto T, Inoue-Sakamoto K, Bryant DA. J Bacteriol (1999). "A Novel Nitrate/Nitrite Permease in the Marine Cyanobacterium *Synechococcus* sp. Strain PCC 7002." 181(23):7363-7372], permeasa de nitrito/nitrato (NrtP), locus AAD45941.

A concentraciones elevadas, el nitrito es tóxico para la mayoría de las células. Para aliviar la toxicidad por nitrito, las células fotoautotróficas se modifican por ingeniería para sobre-expresar genes dentro del operón de tolerancia a nitrito hallado en *Nitrosomonas europaea* ATCC 19718 [Beaumont HJE, Lens SI, Westerhoff HV, van Spanning RJM. "Novel nirK Cluster Genes in *Nitrosomonas europaea* Are Required for NirK-Dependent Tolerance to Nitrite." J Bacteriol (2005). 187(19): 6849-6851], multicobre oxidasa tipo 1 (NirK), locus NP\_840998; citocromo c, clase IC (NcgA), locus NP\_841001; citocromo c, clase I (NcgB), locus NP\_841000; citocromo c, clase IC (NcgC), locus NP\_840999.

Además o como alternativa, las células fotoautotróficas se modifican por ingeniería para asimilar el amoniaco. En este escenario, las células se modifican por ingeniería para sobre-expresar una amonio permeasa. Una amonio permeasa ejemplar hallada en *Synechocystis* sp. PCC 6803 [Montesinos ML, Muro-Pastor AM, Herrero A, Flores E. "Ammonium/Methylammonium Permeases of a Cyanobacterium. Identification and analysis of three nitrogen-regulated amt genes in *Synechocystis* sp. PCC 6803." J Biol Chem (1998). 273(47):31463-31470] que se sobre-expresa es amonio/metilammonio permeasa de alta afinidad (Amt1), locus NP\_442561; amonio/metilammonio permeasa (Amt2), locus NP\_440272; amonio/metilammonio permeasa (Amt3), locus NP\_442793.

Además o como alternativa, las células fotoautotróficas se modifican por ingeniería para asimilar urea. En este escenario, las células deben modificarse por ingeniería para sobre-expresar un transportador de urea para posibilitar la captación eficaz de urea en la célula. Se sobre-expresa un transportador de urea ejemplar hallado en *Nostoc* sp. PCC 7120 [Valladares A, Montesinos ML, Herrero A, Flores E. "An ABC-type, high-affinity urea permease identified in cyanobacteria." Molecular Microbiology (2002). 43(3): 703-715] que comprende el transportador tipo ABC de cinco genes para la captación de urea de alta afinidad, por ejemplo, urea permeasa de alta afinidad, tipo ABC, dominio periplásmico (UrtA), locus CAB70948.1; urea permeasa de alta afinidad, tipo ABC, dominio de membrana (UrtB), locus CAB70949.1; urea permeasa de alta afinidad, tipo ABC, dominio de membrana (UrtC), locus CAB70950.1; urea permeasa de alta afinidad, tipo ABC, dominio de unión a ATP (UrtD), locus CAB70951.1 y urea permeasa de alta afinidad, tipo ABC, dominio de unión a ATP (UrtE), locus CAB70952.1.

Además, se sobre-expresa la urea amidohidrolasa (EC 3.5.1.5) ("ureasa") y sus proteínas accesorias asociadas, que catalizan la conversión de urea en amoniaco y dióxido de carbono. Se sobre-expresa una ureasa ejemplar hallada en *Synechococcus* sp. WH 7805 [Collier J, Brahamsha B, Palenik B. "The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease to utilize urea as a nitrogen source: molecular-genetic and biochemical analysis of the enzyme" Microbiology (1999). 145(2): 447-459] que comprende los genes de ureasa ureABC: urea amidohidrolasa, subunidad gamma (UreA), locus AAC61500; urea amidohidrolasa, subunidad beta (UreB), locus AAC61501; urea amidohidrolasa, subunidad alfa (UreC), locus AAC61502; y los genes ureDE-FG que codifican las proteínas

accesorias proteína accesoria de ureasa (UreD), locus AAC61499; proteína accesoria de ureasa (UreE), locus AAC61498; proteína accesoria de ureasa (UreF), locus AAC61497; y proteína accesoria de ureasa (UreG), locus AAC61496.

- 5 La vitamina B 12 es un cofactor vitamínico que facilita la canalización de reacciones basadas en radicales. Muchos organismos, incluyendo al menos la mitad de todas las microalgas estudiadas, tales como *Synechococcus* sp. PCC 7002, requieren fuentes externas de vitamina B12 para el crecimiento, lo que es prohibitivamente caro en bioprocesamiento industrial a gran escala [Croft MT, Warren MJ, Smith AG. "Algae Need Their Vitamins", *Eukaryotic Cell* (2006) 5(8): 1175-1183]. La necesidad de vitamina B12 puede obviarse por modificación por ingeniería de células fotoautotróficas para que expresen la vía de biosíntesis de la vitamina B12. Se sobre-expresa una vía de biosíntesis ejemplar hallada en *Salmonella typhimurium* (no de la invención), incluyendo, aunque sin limitación, los siguientes 20 genes: uroporfirina-III C-metiltransferasa (CysG), EC 2.1.1.107, locus NP\_462380; sirohidroclorocobalto-quelatasa (CbiK), EC 4.99.1.3, locus NP\_460970; precorrina-2 C20-metiltransferasa (CbiL), EC 2.1.1.130, locus NP\_460969; precorrina-3B metilasa (CbiH), EC 2.1.1.131, locus NP\_460972; CbiG/precorrina metiltransferasa bifuncional (CbiG), locus NP\_460973; precorrina-4 C11-metiltransferasa (CbiF), EC 2.1.1.133, locus NP\_460974; proteína de biosíntesis de cobalamina (CbiD), locus NP\_460977; precorrina-6A reductasa NADPH-dependiente (CbiJ), EC 1.3.1.54, locus NP\_460971; precorrina-6B C5,15-metiltransferasa (CbiE), EC 2.1.1.132, locus NP\_460976; precorrina-6B C12 descarboxilasa (CbiT), EC 2.1.1.132, locus NP\_460975; precorrina-8X-metilmutasa (CbiC), EC 5.4.1.2, locus NP\_460978; ácido cobirínico A,C-diamida sintasa (CbiA), EC 6.3.1.-, locus NP\_460980; ácido cob(I)irínico a,c-diamida adenosiltransferasa (BtuR), EC 2.5.1.17, locus NP\_460677; ácido cobirínico sintasa (CbiP), EC 6.3.5.10, locus NP\_460964; ácido cobírico descarboxilasa (CobD), EC 4.1.1.81, locus NP\_459636; adenosilcobinamida-fosfato sintasa (CbiB), EC 6.3.1.10, locus NP\_460979; alfa ribazol-5'-P fosfatasa (CobC), EC 3.1.3.73, locus NP\_459635; cobalamina(5'-fosfato) sintasa (CobS), EC 2.7.8.26, locus NP\_460962; cobinamida fosfato guanilil transferasa (CobU), EC 2.7.7.62, locus NP\_460963; y nicotinato-nucleótido dimetilbencimidazol-P foforribosil transferasa (CobT), EC 2.4.2.21, locus NP\_460961).

Además, para permitir la captación de cobalto y su incorporación en la vitamina B12, se sobre-expresan los genes que codifican el transportador de cobalto. Se sobre-expresa la proteína transportadora de cobalto ejemplar hallada en *Salmonella typhimurium* sistema de transporte de Co<sup>2+</sup> tipo ABC, componente permeasa (CbiM), locus NP\_460968; sistema de transporte de cobalto tipo ABC, componente periplásmico (CbiN), locus NP\_460967; y sistema de transporte de cobalto tipo ABC, componente permeasa (CbiQ), locus NP\_461989.

En la invención, los organismos fotoautotróficos se modifican por ingeniería para sobre-expresar la enzima independiente de vitamina B12 para obviar la necesidad de este cofactor completamente. En la mayoría de los organismos fotoautotróficos, solamente la metionina sintasa (EC 2.1.1.13) y las ribonucleótido reductasas de clase II requieren vitamina B12. Por lo tanto se sobre-expresa una metionina sintasa ejemplar independiente de vitamina B12 (EC 2.1.1.14) de *Thermotoga maritima*: 5-metiltetrahidropteroltriglutamato-homocisteína metiltransferasa (MetE), locus NP\_229090 (no de la invención).

40 Además se contempla que la independencia de nutrientes (por ejemplo, vitamina B12) de células hospedadoras puede realizarse por expresión de diversas proteínas que codifican 5-metiltetrahidropteroltriglutamato-homocisteína metiltransferasa (metE) que comprenden secuencias que son al menos un 95%, 96%, 97%, 98%, 99% o 100% idénticas a metE.

#### 45 Construcción de cepas

El vector de expresión, pJB5, contiene dos regiones de ADN de 500 pb homólogas a secuencias en el plásmido natural pAQ1 de *Synechococcus* sp. PCC 7002. Entre las dos regiones de homología hay un promotor y un sitio de unión al ribosoma, sitios para la inserción de la secuencia de ADN, y un casete de resistencia, aadA1, que confiere resistencia a espectinomicina. Los genes *metE* de *E. coli* (NP\_418273) y *Thermosynechococcus elongatus* BP-1 (NP\_681881) se codificaron en pJB6 y pJB7 respectivamente. Todos los plásmidos se dirigieron con *NdeI* (New England Biolabs) y *EcoRI* (New England Biolabs), y el fragmento de ~1 kb de pJB6 y pJB7, y el fragmento grande de pJB5, se aislaron en gel y se purificaron usando técnicas convencionales (Qiagen). Los fragmentos de pJB6 y pJB7 se ligaron y transformaron en pJB5 usando técnicas convencionales para formar pJB5-6 y pJB5-7.

55 *Synechococcus* sp. PCC 7002 se transformó con pJB5-6 y pJB5-7 del siguiente modo. Se dirigieron pJB5-6 y pJB5-7 y se inactivaron con *SbfI* (New England Biolabs) durante 1 hora y se inactivaron por calor. El ADN se incubó con células frescas a OD730 de 1 durante 4 horas en una incubadora en la oscuridad a 37°C en A<sup>+</sup>. Las células después se dejaron diluir en 20 ml de medio A<sup>+</sup> fresco con luz moderada y se burbujearon con aire con CO<sub>2</sub> al 1% durante 24 horas. Las células después se diluyeron 1 en 20 en medio A<sup>+</sup> fresco que contenía 10 ug/ml de espectinomicina y se dejaron crecer durante 5 días en las mismas condiciones. Las células se diluyeron de nuevo 1 en 20 en 25 ml de medio A<sup>+</sup> que carecía de vitamina B12, y esto se repitió una segunda vez después de que las células crecieran a alta densidad. Después de la segunda extensión, las células se sembraron en placas A<sup>+</sup> que carecían de vitamina B12.

65 Las Fig. 3A-D muestran *Synechococcus* sp. PCC 7002 de tipo silvestre y células que expresan de forma transgénica *MetE* de *E. coli* y *MetE* de *Thermosynechococcus elongatus* BP-1 diluidas y cultivadas durante una noche en medio

- A<sup>+</sup> que carecía de vitamina B12. Las células se diluyeron, después se sembraron y se dejaron crecer 1 semana a 37°C en una incubadora iluminada en placas tanto con suficiente vitamina B12 como deficientes en la misma. La **Fig. 3A** ilustra *Synechococcus* de tipo silvestre en placa con suficiente B 12 mientras que la **Fig. 3B** muestra la placa B12 deficiente. La **Fig. 3C** representa una cepa transgénica de *Synechococcus* con *MetE* de *E. coli* en la placa vitamina B 12 deficiente y la **Fig. 3D** muestra *MetE* de *Thermosynechococcus elongatus* en la placa vitamina B12 deficiente. Los resultados muestran la capacidad de la metionina sintasa expresada de forma transgénica de rescatar los requisitos de vitamina B12 de las cianobacterias.

**REIVINDICACIONES**

- 5 1. Una célula cianobacteriana modificada por ingeniería que comprende un ácido nucleico modificado por ingeniería que codifica una proteína 5-metiltetrahidropteroiltriglutamato-homocisteína metiltransferasa independiente de vitamina B12 que comprende una secuencia que es al menos un 95% idéntica a la secuencia mostrada en la Figura 4.
- 10 2. La célula cianobacteriana modificada por ingeniería de la reivindicación 1, donde dicha proteína 5-metiltetrahidropteroiltriglutamato-homocisteína metiltransferasa independiente de vitamina B12 tiene una secuencia de aminoácidos idéntica a la mostrada en la Figura 4.
- 15 3. La célula cianobacteriana modificada por ingeniería de una cualquiera de las reivindicaciones 1 ó 2, donde dicha célula cianobacteriana es una especie de *Synechococcus*.
- 20 4. La célula cianobacteriana modificada por ingeniería de la reivindicación 3, donde dicha especie de *Synechococcus* es *Synechococcus* sp. PCC 7002.
- 25 5. Uso de una proteína 5-metiltetrahidropteroiltriglutamato-homocisteína metiltransferasa independiente de vitamina B12 que comprende una secuencia que es al menos un 95% idéntica a la secuencia mostrada en la Figura 4 para producir una célula cianobacteriana independiente de vitamina B12.
6. El uso de la reivindicación 5, donde dicha proteína 5-metiltetrahidropteroiltriglutamato-homocisteína metiltransferasa independiente de vitamina B12 tiene una secuencia de aminoácidos idéntica a la secuencia mostrada en la Figura 4.
- 25 7. El uso de la reivindicación 5 ó 6, donde dicha célula cianobacteriana es una especie de *Synechococcus*.
8. El uso de la reivindicación 7, donde dicha especie de *Synechococcus* es *Synechococcus* sp. PCC 7002.
- 30 9. El uso de una cualquiera de las reivindicaciones 5 a 8, donde una capacidad de sobrevivir y replicarse en medio que carece de vitamina B<sub>12</sub> como suplemento del medio confirma la aplicación satisfactoria de independencia de vitamina B<sub>12</sub>.
- 35 10. Un método para producir un producto de interés basado en carbono, que comprende cultivar una célula cianobacteriana modificada por ingeniería de una cualquiera de las reivindicaciones 1 a 4 en presencia de CO<sub>2</sub> y luz en condiciones adecuadas para producir el producto de interés basado en carbono.

TABLA 1 GENES A SOBRE-EXPRESAR

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.2	Acetil-CoA carboxilasa (subunidad alfa )	<i>Escherichia coli</i>	AAA70370	Homo sapiens [ACACA, NC0000017.9]
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.2	Acetil-CoA carboxilasa (subunidad beta)	<i>Escherichia coli</i>	AAA23807	<i>Arabidopsis thaliana</i> [AtCg00500]
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.2	Proteína vehículo de biotina-carboxilo (accB)	<i>Escherichia coli</i>	JW3223	Bacillus halodurans [BH1132], <i>Vibrio cholerae</i> [EAZ76879.1 o A5E_0311]
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.2	biotina-carboxilasa	<i>Escherichia coli</i>	AAA23748	<i>Photobacterium profundum</i> 3TCK [EAS42088.1 o 90325619]
Fijación del carbono	Ciclo de hidroxipropionato	3- 1.1.1.59	malonil-CoA reductasa	<i>Chloroflexus aurantiacus</i>	AY530019	solamente un I puede encontrarse mediante BLASTp
Fijación del carbono	Ciclo de hidroxipropionato	3-	3-hidroxipropionil-CoA sintasa	<i>Chloroflexus aurantiacus</i>	AF445079	sinetasa y ligasa dependiente de AMP, homólogo de BLASTp de 2702 bits, 81% positivo, ID con 4% de huecos [ABQ91563.1] de <i>Roseiflexus</i> sp RS-1
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.3	propionil-CoA carboxilasa (subunidad alfa )	Roseobacter denitrificans	RD1_2032	Gen PCCA mitocondrial de <i>Homo sapiens</i> [X14608]. Gen PCCA de <i>Mus musculus</i> [AY046947]
Fijación del carbono	Ciclo de hidroxipropionato	3- 6.4.1.3	propionil-CoA carboxilasa (subunidad beta)	Roseobacter denitrificans	RD1_2028	PCCB mitocondrial de <i>Rhodococcus erythropolis</i> [ABB80770.1], <i>Homo sapiens</i> [X73424J]

Figura 1

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del Ciclo carbono hidroxipropionato	de	3- 5.1.99.1.	metilmalonil-CoA epimerasa	Rhodobacter sphaerooides	CP000661	Homo sapiens MCEE [AF364547]
Fijación del Ciclo carbono hidroxipropionato	de	3- 5.1.99.2	metilmalonil-CoA mutasa	Escherichia coli	NC000913.2	Homo sapiens MUT [M65131]
Fijación del Ciclo carbono hidroxipropionato	de	3-	succinil-CoA:L-malato CoA transferasa aurantiacus (subunidad alfa )	Chloroflexus aurantiacus	DQ472736.1	Chloroflexus aggregans DSM 0485 [ZP_01516527.1 o EAIV09800.1]
Fijación del Ciclo carbono hidroxipropionato	de	3-	succinil-CoA:L-malato CoA transferasa aurantiacus (subunidad beta)	Chloroflexus aurantiacus	DQ472737.1	Chloroflexus aggregans DSM 9485 [ZP_01516526.1 o EAIV09799.1]
Fijación del Ciclo carbono hidroxipropionato	de	3- 1.3.1.6	fumarato reductasa - frdA - subunidad de flavoproteína	Escherichia coli	J01611; AAA23437.1	fumarato reductasa NP_458782.1 de Salmonella enterica subsp. Enterica serovar o Klebsiella pneumoniae ABR79907.1
Fijación del Ciclo carbono hidroxipropionato	de	3- 1.3.1.6	fumarato reductasa subunidad de hierro-azufre- frdB	Escherichia coli	J01611; EAY46226.1	succinato de hidrogenasa de Salmonella typhimurium LT2 [NP_463206.1]
Fijación del Ciclo carbono hidroxipropionato	de	3- 1.3.1.6	subunidad g15 [fumarato reductasa subunidad c]	Escherichia coli	J01611; NP_290787.1	Shigella flexneri 2a str. 301 [NP_710021.1], Klebsiella pneumoniae ABR79905.1
Fijación del Ciclo carbono hidroxipropionato	de	3- 1.3.1.6	subunidad g13 [fumarato reductasa subunidad D]	Escherichia coli	J01611, NP_757087.1	Salmonella enterica [YP_153210.1], Photorhabdus luminescens [NP_931317.1]
Fijación del Ciclo carbono hidroxipropionato	de	3- 4.2.1.2	fumarato hidratasa - clase I aeróbica (fumA)	Escherichia coli	CAA25204	fumarato hidratasa (fumB) AA233827 o clase II (fumC) CAA27698 de E. coli clase I anaeróbica

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono Ciclo de hidroxipropionato	3- 4.1.3.24	L-malil-CoA ligasa	Roseobacter denitrificans	NC_008209.1	Citratopomeroj BLASTp [YP_166806.1] y HTCC2255	Citratopomeroj DSS-3 [identificada a partir de proteobacterium ZP_01447127.1]
Fijación del carbono TCA reductor	2.3.3.8	ATP-citratotransfuerza, subunidad 1	Chlorobium tepidum	CT1089	Chlorobium ferrooxidans ZP_01385848.1	Chlorobium limicola [BAB21375.1], Chlorobium DSM 13031
Fijación del carbono TCA reductor	2.3.3.8	ATP-citratotransfuerza, subunidad 2	Chlorobium tepidum	CT1088	Chlorobium YP_911761.1, Chlorobium phaeobacteroides YP_911761.1, Chlorobium ferrooxidans ZP_01385849.1.	Chlorobium limicola [BAB21376.1], Chlorobium phaeobacteroides YP_911761.1, Chlorobium ferrooxidans ZP_01385849.1.
Fijación del carbono TCA reductor		citril-CoA sintasa (subunidad grande)	Hydrogenobacter thermophilus	BAD117844	Aquifex aeolicus [O67330], Leptospirillum sp. Grupo II UBA [A3ERU1]	Aquifex aeolicus [NP_214297.1], Leptospirillum sp. Grupo II UBA [A3ERU1]
Fijación del carbono TCA reductor		citril-CoA sintasa (subunidad pequeña)	Hydrogenobacter thermophilus	BAD117846	Aquifex Leptospirillum sp [EAY57418.1]	Aquifex aeolicus [NP_214297.1], Leptospirillum sp. Grupo II UBA [A3ERU1]
Fijación del carbono TCA reductor		citril-CoA ligasa	Hydrogenobacter thermophilus	BAD 17841	Aquifex Hydrogenobacter [ABU50085.1]	Aquifex aeolicus [NP_213101.1], Hydrogenobacter hydrogenophilus
Fijación del carbono TCA reductor	1.1.1.37	malato deshidrogenasa	Chlorobium tepidum	CAA56810	Prosthecochloris CAA56809.1, Pelodictyon luteolum DSM 273 [YP_375410.1]	vibriiformis vibrioformis
Fijación del carbono TCA reductor	4.2.1.2	fumarasa hidratasa (isozima aeróbica, fumA)	Escherichia coli	JW1604	isozima anaeróbica fumB de E. coli clase 1 (JW4083) y clase II fumC (JW1603)	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del TCA reductor carbono		1.3.99.1.	succinato deshidrogenasa (subunidad flavoproteína - SdhA)	Escherichia coli	NP_415251	Enterobacter sp. 638 [YP_001175956.1], <i>Serratia proteamaculans</i> [ZP_01538596.1]
Fijación del TCA reductor carbono		1.3.99.1.	subunidad de hierro-azufre SdhB	Escherichia coli	NP_415252	Salmonella enterica [YP_00151223.1], <i>Yersinia enterocolitica</i> [YP_001007133.1]
Fijación del TCA reductor carbono		1.3.99.1.	subunidad de anclaje a membrana SdhC	Escherichia coli	NP_415249	Enterobacter sp. 638 [ABP59903.1], <i>Yersinia frederiksenii</i> [ZP_00828037.1]
Fijación del TCA reductor carbono		1.3.99.1.	subunidad de anclaje a membrana SdhD	Escherichia coli	NP_415250	Enterobacter sp. 638 [YP_001175956.1], <i>Klebsiella pneumoniae</i> [YP_001334402.1]
Fijación del TCA reductor carbono		6.2.1.5	succinil-CoA sintetasa (sucD)	Escherichia coli	AAA23900	<i>Chlorobium tepidum</i> [AAM71515]
Fijación del TCA reductor carbono		6.2.1.5	succinil-CoA sintetasa	Escherichia coli	AAA23899	<i>Chlorobium tepidum</i> [AAM71626]
Fijación del TCA reductor carbono		1.2.7.3	alfa-cetoglutarato subunidad alfa -korA	Hydrogenobacter thermophilus	AB046568:46 -1869	<i>Chlorobium limicola</i> DSM 245; Números de acceso EAM42575, EAM42574, EAM42853, EAM42852.
Fijación del TCA reductor carbono		1.2.7.3	alfa-cetoglutarato subunidad beta -korB	Hydrogenobacter thermophilus	AB046568:18 83-2770	Hydrogenobacter thermophilus TK-6, para el grupo de genes DABGE 5
Fijación del TCA reductor carbono		1.1.1.42	Isocitrato deshidrogenasa -NADP dependent	Chlorobium limicola	EAM42635	<i>Synechococcus</i> sp WH 8102, icd, acceso CAE06681
Fijación del TCA reductor carbono		1.1.1.41	isocitrato deshidrogenasa -NAD depend. Subunidad 1	Saccharomyces cerevisiae	YNL037C	
Fijación del TCA reductor carbono		1.1.1.41	isocitrato deshidrogenasa -NAD depend. Subunidad 2	Saccharomyces cerevisiae	YOR136W	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	TCA reductor	4.2.1.3	aconitato hidratasa 1 (acnA)	Escherichia coli	b1276	
Fijación del carbono	TCA reductor	4.2.1.3	aconitato hidratasa 2 (acnB)	Escherichia coli	b0118	
Fijación del carbono	TCA reductor	1.2.7.1	Piruvato sintasa , subunidad A porA	Clostridium tetani	AA036986	
Fijación del carbono	TCA reductor	1.2.7.1	Piruvato sintasa , subunidad B porB	Clostridium tetani	AA036985	
Fijación del carbono	TCA reductor	1.2.7.1	Piruvato sintasa, subunidad C por C	Clostridium tetani	AA036988	
Fijación del carbono	TCA reductor	1.2.7.1	Piruvato sintasa , subunidad D por D	Clostridium tetani	AA036987	
Fijación del carbono	TCA reductor	2.7.9.2	Fosfoenolpiruvato sintasa - ppsA	Escherichia coli	AAA2431	Aquilex aeolicus VF5 ppsA (AAC07865).
Fijación del carbono	TCA reductor	4.1.1.31	PEP carboxilasa, ppo	Escherichia coli	CAA29332	
Fijación del carbono	Woods-Ljungdahl	1.2.1.4.3	Formato deshidrogenasa NADP-dependiente - subunidad A Mt-fdhA	Morella thermoacetica	AAB18330	
Fijación del carbono	Woods-Ljungdahl	1.2.1.4.3	Formato deshidrogenasa NADP-dependiente - subunidad B Mt-fdhB	Morella thermoacetica	AAB18329	
Fijación del carbono	Woods-Ljungdahl	6.3.4.3	formato tetrahidrofolato ligasa	Clostridium acidi- urici	M21507	AAB49329 de <i>Streptococcus mutans</i> (proteína = Q599925) o la proteína Q8XHL4 de <i>Clostridium perfringens</i> (codificada por "complemento" (BA00000 16.3:2825121 .2826791).

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	Woods-Ljungdahl	3.5.4.9 1.5.1.5	Meteniltetrahidrofolato ciclohidroasa	Escherichia coli	AAA23803	Moorella thermoacetica, AAO3B126 de Clostridium tetani, y BAB81529 de Clostridium perfringens. Todos son enzimas fold bifesionales.
Fijación del carbono	Woods-Ljungdahl	1.5.1.20	métíleno tetrahidrofolato reduktasa, metF	Escherichia coli	CAA24747	ABC19825 (fold) de AAC23094 de Haemophilus influenzae, o CAA30531 de <i>Salmonella typhimurium</i> .
Fijación del carbono	Woods-Ljungdahl	-	5-metiletetrahidrofolato proteína corriodo/hierro azufre metiltransferasa, acSE	Moorella thermoacetica	AAA53548	acsE de Carboxydothermus hydrogenoformas locus ABB15216.
Fijación del carbono	Woods-Ljungdahl	1.2.7.4/ 1.2.99.2	Monóxido de carbono deshidrogenasa/acetil-CoA sintasa - subunidad alfa	Moorella thermoacetica	AAA23229	Carboxydothermus hydrogenoformase [YP_360060]
Fijación del carbono	Woods-Ljungdahl	1.2.7.4/ 1.2.99.2	Monóxido de carbono deshidrogenasa/acetil-CoA sintasa - subunidad beta	Moorella thermoacetica	AAA23228	
Fijación del carbono	Derivación glioxilato	de 2.3.3.9	malato sintasa - aceB	Escherichia coli	JW3974	E. coli JW2943 locus malato sintasa G (glcB)
Fijación del carbono	Derivación glioxilato	de 4.1.3.1	Isocitrato liasa - aceA	Escherichia coli	JW3975	
Fijación del carbono	Derivación glioxilato	de 1.1.1.37	malato deshidrogenasa	Escherichia coli	JW3205	
Fijación del carbono	Gluconeogénesis	6.4.4.1	Piruvato carboxilasa	Saccharomyces cerevisiae	YGL062W	
Fijación del carbono	Gluconeogénesis	4.1.1.49	Fosfoenolpiruvato carboxiquinasa	Escherichia coli	JW3366	
Fijación del carbono	Gluconeogénesis	3.1.3.11	fructosa-1,6-bisfosfatasa	Escherichia coli	JW4191	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	Gluconeogénesis	3.1.3.68	glucosa-6-fosfatasa - dog1	Saccharomyces cerevisiae	YHR044C	locus YHR043C de <i>Saccharomyces cerevisiae</i> , dog2
Fijación del carbono	Piruvato síntesis	1.2.7.1	Piruvato ferredoxina:oxido-reductasa con actividad piruvato sintasa	Morella thermoacetum	Moth_0064	
Fijación del carbono	Ciclo de Calvin	EC 4.1.2.13	Fructosa-1,6-bisfotato aldolasa - clase I (Fda)	Synechocystis sp PCC 6803	NP_441723	Synechococcus sp. WH 7805 = ZP_01124026
Fijación del carbono	Ciclo de Calvin	EC 4.1.2.13	Fructosa-1,6-bisfotato aldolasa - clase II (FbaA)	Synechocystis sp PCC 6803	BAA10184	
Fijación del carbono	Ciclo de Calvin	EC 1.2.1.13	gliceralehído-3-fosfato deshidrogenasa (GAPDH), cbgG	Prochlorococcus marinus	NP_875968	
Fijación del carbono	Ciclo de Calvin	EC 2.7.1.19	Fosforibuloquinasa (PrkA)	Chlamydomonas reinhardtii	AAA333090	Prochlorococcus marinus = NP_894365
Fijación del carbono	Ciclo de Calvin		CP12	Thermosynechocoecus elongatus BP-1	BAC09372	Chlamydomonas reinhardtii = CAO03469; Synechococcus elongatus PCC 6301 = BAD79451
Fijación del carbono	Ciclo de Calvin	EC 2.2.1.1	Transketolasa (TKtA)	Synechocystis sp. PCC 6301	YPJ71693	
Fijación del carbono	Ciclo de Calvin	EC 3.1.3.11	Fructosa-1,6-bisfotata (Fbp)	Synechocystis sp. PCC 6803	NP_441738	
Fijación del carbono	Ciclo de Calvin	EC 5.1.3.1	Pentosa-5-fosfato-3-epimerasa (Rpe)	Synechocystis sp. mPCC 6301	YP_171630	
Fijación del carbono	Ciclo de Calvin	EC 4.1.2.13	Sedoheptulosa-1,7-bisfotato aldolasa (RpaA)	Thermosynechocoecus elongatus BP-1	NP_681166	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen exemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	Ciclo de Calvin	EC 3.1.3.37	Sedoheptulosa-1,7-bisfotatasa (SBPasa)	<i>Chlamydomonas reinhardtii</i>	CAA52439	
Fijación del carbono	Ciclo de Calvin	EC 5.3.1.6	Ribosa 5-fosfato Isomerasa (RpiA)	<i>Synechococcus elongatus PCC 6301</i>	YP_171649	
Fijación del carbono	Ciclo de Calvin	EC 2.7.2.3	Fosfoglicerato quinasa	<i>Synechococcus elongatus PCC 6301</i>	BAD78623	
Fijación del carbono	Ciclo de Calvin	EC 5.3.1.1	triosefotato isomerasa, tpiA	<i>Synechocystis sp PCC 6803</i>	Q59994	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxigenasa (RubisCo)- subunidad pequeña -cbbS	<i>Synechococcus spWH7803</i>	AAB48081	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxigenasa (RubisCo)- subunidad grande cblL	<i>Synechococcus spWH7803</i>	AAB8080	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxigenasa - subunidad pequeña (RbcS)	<i>Synechococcus elongatus PCC 6301</i>	YP_170839	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxigenasa - subunidad grande (RbcL)	<i>Synechococcus elongatus PCC 6301</i>	YP_170840	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxigenasa - subunidad pequeña (CbbS)	<i>Rhodobacter sphaeroides</i>	P27998	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxygenasa-subunidad grande (CbbL)	Rhodobacter sphaeroides	P27997	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxygenasa (CbbM)	Rhodobacter sphaeroides	P29278	
Fijación del carbono	Ciclo de Calvin	EC 4.1.1.39	Ribulosa-1,5-bisfotato carboxilasa/oxygenasa (RbcL)	Methanocaldococcus jannaschii	Q58632	
Fijación del carbono	Ciclo de Calvin		Rubisco activasa	Synechococcus sp. JA-3-3Ab	ABC98646	Chlamydomonas reinhardtii (EDP04194)
Fijación del carbono	Flujo de carbono-acetil-CoA	EC 2.7.1.33	pantotenato quinasa (panK)	Synechococcus sp. JA-3-3Ab	YP_473820	
Fijación del carbono	Flujo de carbono-acetil-CoA	EC 1.2.4.1	Piruvato, deshidrogenasa, subunidad alfa (pdhA)	Synechococcus PCC 6301	YP_172860	
Fijación del carbono	Flujo de carbono-acetil-CoA	EC 1.2.4.1	Piruvato, deshidrogenasa, subunidad alfa (pdhB)	Synechococcus PCC 6301	YP_172072	
Tolerancia a gases de combustión	Tolerancia a NOx		Multicobre oxidasa de tipo 1 (NirK)	Nitrosomonas europaea	NP_840998	
Tolerancia a gases de combustión	NOx@France		Multicobre oxidasa de tipo 1 (ncgA)	Nitrosomonas europaea	NP_841001	
Tolerancia a gases de combustión	Tolerancia a NOx		Citocromo c, clase I (ncgB)	Nitrosomonas europaea	NP_841000	
Tolerancia a gases de combustión	Tolerancia a NOx		Citocromo c, clase Ic (ncgC)	Nitrosomonas europaea	NP_840999	
Tolerancia a gases de combustión	SOx	EC 4.2.99.8 y 2.5.1.47	cisteína sintasa A (cysK)	Synechococcus PCC 7942	YP_398721	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia a gases de combustión	Tolerancia a SOx	EC 1.15.1.1	superóxido dismutasa (sodA)	Synechocystis PCC 6803	NP_441347	
Tolerancia a gases de combustión	Tolerancia a SOx	EC1.11.16	catalasa (katG)	Synechocystis PCC 6803	NP_441295	
Captura de luz	luz->PMF		Proteorrodopsina	Bacteria marina no cultivada HF10_19P19	ABL60988 gen HOT 0m1 (AF349978), el gen palE6 75m4 (AF349981), el gen SAR86 de eBAC31 (AF350002), y el gen A08 (AAG10475.	
Captura de luz	luz->PMF		bacteriorrodopsina	Halobacterium species NRC-1	NP_280292 gen de Halobacterium salinarum (V00474)	
Captura de luz	luz->PMF		deltarrodopsina 4	Haloterrigena sp arg-4	AB009620 Las alternativas incluyen la variante descrita en Kamo N et al, BBRC 2006, de Haloterrigena turkmenica, que difiere solamente en 2 posiciones en comparación con AB009620 (nunca se depositó número de acceso)	
Captura de luz	luz->PMF		xantorrodopsina	Salinibacter ruber DSM 13855	ABC44767	
Captura de luz	luz->PMF		opsina	Leptosphaeria maculans	AAG01180	
Captura de luz	Biosíntesis retinal	5.3.3.2	Isopentenil-difosfato delta-isomerasa	Bacteria marina no cultivada HF10_19P19	ABL60982 (JW2857) y Rhodococcus capsulatus (CAA77535.1)	Las alternativas incluyen E. coli
Captura de luz	Biosíntesis retinal	1.14.99.36	15,15'-beta-caroteno dioxygenasa	Bacteria marina no cultivada HF10U9P19	ABL60983	Homo sapiens (AAG153B0) y Mus musculus (AJ278064)

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso o	Alternativos
Captura de luz	Biosíntesis retinal		Licopeno ciclase	Bacteria marina no cultivada HF10_19P19	ABL60984	gen crtA de <i>Synechococcus</i> sp PCC 7002 (EF529626) y crtP de la misma especie (EF529627), y crtY de <i>Streptomyces coelicolor</i> (SCJ12,03, o NC_003888,3)
Captura de luz	Biosíntesis retinal		Fitoeno sintasa	Bacteria marina no cultivada HF10_19P19	ABL60985	<i>Streptomyces coelicolor</i> A3(2) [locusSC00187] o <i>Prochlorococcus marinus</i> crtB [Pro0166 o NC_005042,1]
Captura de luz	Biosíntesis retinal		Fitoeno deshidrogenasa	Bacteria marina no cultivada HF10_19P19	ABL60986	<i>Prochlorococcus marinus</i> [Pro0167] o <i>Thermosynechococcus elongatus</i> BP-1 [hasta 1561]
Captura de luz	Biosíntesis retinal		Garanilgeranil pirofosfato sintetasa	Bacteria marina no cultivada HF10_19P19	ABL60987	gen crtE de <i>Rhodobacter sphaeroides</i> 2.4.1 [RSP_0265] y <i>Arabidopsis thaliana</i> GGPS3 [AT3G14550]
Captura de luz	SalinixanthinA		beta-caroteno cetolasa	Salinibacter ruber DSM 13855	SRU_1502	los genes CrtO incluyen <i>Rhodococcus erythropolis</i> (AY705709), <i>Deinococcus radiodurans</i> R1 (pi1-1E755B1), y <i>Gloeoibacter violaceus</i> PCC 7421 [gvip239].
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F0 de la ATP sintasa unidad a membrana, subunidad a (Que consiste en atpB)	<i>Escherichia coli</i>	NP_418194	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F0 de la ATP sintasa unidad a membrana, subunidad c (AtpE)	<i>Escherichia coli</i>	NP_418193	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F0 de la ATP sintasa unidad a membrana, subunidad b (AtpF)	<i>Escherichia coli</i>	NP_418192	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F1 de la ATP sintasa unidad a membrana subunidad alfa (AtpA)	<i>Escherichia coli</i>	NP_418190	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F1 de la ATP sintasa unidad a membrana epsilon subunidad (AtpC)	<i>Escherichia coli</i>	NP_418187	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F1 de la ATP sintasa unidad a membrana subunidad beta (AtpD), I atpD (Locus NP_418188)	<i>Escherichia coli</i>	NP_418188	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F1 de la ATP sintasa unidad a membrana subunidad gamma (AtpG)	<i>Escherichia coli</i>	NP_418189	
Captura de luz	Síntesis de ATP	3.6.3.14	Sector F1 de la ATP sintasa unidad a membrana delta subunidad (AtpH)	<i>Escherichia coli</i>	NP_418191	
Independencia de nutriente	Asimilación de nitrato/nitrilo		Proteína de unión a sustrato del sistema de transporte de nitrato/nitrito tipo ABC (nrtA)	<i>Synechococcus elongatus</i> PCC 6301	YPJ71021	
Independencia de nutriente	Asimilación de nitrato/nitrilo		Proteína permeasa del sistema de transporte de nitrato/nitrito tipo ABC (nrtB)	<i>Synechococcus elongatus</i> PCC 6301	YP_171022	
Independencia de nutriente	Asimilación de nitrato/nitrilo		Proteína de unión a ATP del sistema de transporte de nitrato/nitrito tipo ABC (nrtC)	<i>Synechococcus elongatus</i> PCC 6301	YP_171023	

**Figura 1 (Continuación)**

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Independencia de nutriente	Asimilación de nitrato/nitrito		proteína de unión a ATP del sistema de transporte de nitrato/nitrito tipo ABC (nrdD)	Synechococcus elongatus PCC 6301	YP_171024	
Independencia de nutriente	Asimilación de nitrato/nitrito		Nitrito/nitrito permeasa (nrnP)	Synechococcus sp. PCC 7002	AAD45941	
Independencia de nutriente	Asimilación de nitrato/nitrito - Tolerancia a nitrito		Multicobre oxidasa de tipo 1 (NirK)	Nitrosomonas europaea	NP_840998	
Independencia de nutriente	Asimilación de nitrato/nitrito - Tolerancia a nitrito		Multicobre oxidasa de tipo 1 (ncgA)	Nitrosomonas europaea	NP_841001	
Independencia de nutriente	Asimilación de nitrato/nitrito - Tolerancia a nitrito		Citocromo c, clase I (ncgB)	Nitrosomonas europaea	NP_841000	
Independencia de nutriente	Asimilación de nitrato/nitrito - Tolerancia a nitrito		Citocromo c, clase IC (ncgC)	Nitrosomonas europaea	NP_840999	
Independencia de nutriente	Asimilación de nitrato/nitrito		nitrato reductasa (narA)	Synechococcus sp. PCC 7002	AAK49018	
Independencia de nutriente	Asimilación de nitrato/nitrito		nitrato reductasa (narB)	Synechococcus sp. PCC 7002	AAD45942	
Independencia de nutriente	Asimilación de nitrato/nitrito		amonio/metillamonio permeasa de alta afinidad (amt)	Synechocystis sp. PCC 6803	NP_442561	
Independencia de nutriente	Asimilación de amoníaco		amonio/metillamonio permeasa (amt2)	Synechocystis sp. PCC 6803	NP_440272	
Independencia de nutriente	Asimilación de amoníaco		amonio/metillamonio permeasa (amt3)	Synechocystis sp. PCC 6803	NP_442793	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/ acceso	Alternativos
Independencia de nutriente	Asimilación de urea	3.5.1.5	urea amidohidrolasa, subunidad gamma (ureA)	Synechococcus sp. WH 7805	AAC61500.1	
Independencia de nutriente	Asimilación de urea	3.5.1.5	urea amidohidrolasa, subunidad beta (ureB)	Synechococcus sp. WH 7805	AAC61501.1	
Independencia de nutriente	Asimilación de urea	3.5.1.5	urea amidohidrolasa, subunidad alfa (ureC)	Synechococcus sp.WH7805	AAC61502.1	
Independencia de nutriente	Asimilación de urea		proteína accesoria de ureasa (ureD)	Synechococcus sp. WH 7805	AAC61499.1	
Independencia de nutriente	Asimilación de urea		proteína accesoria de ureasa (ureE)	Synechococcus sp. WH 7805	AAC61498.1	
Independencia de nutriente	Asimilación de urea		proteína accesoria de ureasa (ureF)	Synechococcus sp. WH 7805	AAC61497.1	
Independencia de nutriente	Asimilación de urea		proteína accesoria de ureasa (ureG)	Synechococcus sp. WH 7805	AAC61496.1	
Independencia de nutriente	Asimilación de urea		Urea permeasa de alta afinidad, tipo ABC, dominio periplásmico(ureA)	Nostoc sp. PCC 7120	CAB70948.1	
Independencia de nutriente	Asimilación de urea		Urea permeasa de alta afinidad, tipo ABC, dominio de membrana (ureB)	Nostoc sp. PCC 7120	CAB70949.1	
Independencia de nutriente	Asimilación de urea		Urea permeasa de alta afinidad, tipo ABC, dominio de membrana (ureC)	Nostoc sp. PCC 7120	CAB70950.1	
Independencia de nutriente	Asimilación de urea		Urea permeasa de alta afinidad, tipo ABC, dominio de unión (ureD)	Nostoc sp. PCC 7120	CAB70951.1	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Independencia de nutriente	Asimilación de urea		Urea permeasa de alta afinidad, tipo ABC, dominio de unión a ATP (urnE)	Npstoc sp. PCC 7120	CAB70952.1	
Independencia de nutriente	Fijación de nitrógeno		Proteína de biosíntesis del cofactor FeMo (nifB)	Nostoc. Sp PCC 7120	NP_485557	
Independencia de nutriente	Fijación de nitrógeno		[4Fe-4S] ferredoxina (fclxN)	Nostoc. Sp PCC 7120	BA0000019.2	
Independencia de nutriente	Fijación de nitrógeno	2.8.1.7	L-Cisteína desulfurasa (nifS)	Nostoc. Sp PCC 7120	NP_485499	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del grupo Fe (nifU)	Nostoc. Sp PCC 7120	NP_485498	
Independencia de nutriente	Fijación de nitrógeno	1.18.6.1	Nitrogenasa - subunidad Fe (nifH)	Nostoc. Sp PCC 7120	NP_485497	
Independencia de nutriente	Fijación de nitrógeno	1.18.6.1	Nitrogenasa - subunidad alfa (nifD)	Nostoc. Sp PCC 7120	NP_485484	
Independencia de nutriente	Fijación de nitrógeno	1.18.6.1	Nitrogenasa - subunidad beta (nifK)	Nostoc. Sp PCC 7120	NP_485483	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifE)	Nostoc. Sp PCC 7120	NP_485481	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifN)	Nostoc. Sp PCC 7120	NP_485480	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifX)	Nostoc. Sp PCC 7120	NP_485479	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifW)	Nostoc. Sp PCC 7120	NP_485476	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (hesA)	Nostoc. 7120	Sp PCC NP_485475	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del ensamblaje del cofactor del grupo FeS (hesB)	Nostoc. 7120	Sp PCC NP_485474	
Independencia de nutriente	Fijación de nitrógeno		Ferrodoxina específica de fijación de nitrógeno (FdXH)	Nostoc. 7120	Sp PCC NP_485473	
Independencia de nutriente	Fijación de nitrógeno	1.2.7.1	piruvato oxidoreductasa (nifJ)	Nostoc. 7120	Sp PCC NP_486843	
Independencia de nutriente	Fijación de nitrógeno	2.3.3.14	homocitratot sintasa (nifV)	Nostoc. 7120	Sp PCC NP_485450	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifZ)	Nostoc. 7120	Sp PCC NP_485451	
Independencia de nutriente	Fijación de nitrógeno		Proteína accesoria del cofactor FeMo (nifT)	Nostoc. 7120	Sp PCC NP_485452	
Independencia de nutriente	Independencia de vitamina B12	EC 2.1.1.14	5-metiltetrahidropteroltriglutamato-homocisteina metiltransferasa (metE)	Thermotoga maritima	NP_229090	
Independencia de nutriente	Independencia de vitamina B12		Ribonucleósido-difosfato reductasa, subunidad alfa (nrdA)	Synechocystis sp. PCC 6803	NP_441654	
Independencia de nutriente	Independencia de vitamina B12		Ribonucleósido-difosfato reductasa, subunidad beta (nrdB)	Synechocystis sp. PCC 6804	NP_443040	
Independencia de nutriente	Biosíntesis vitamina B12	2.1.1.107	uroporfirin-III metiltransferasa; cobaltoqueletatasa (cysG)	C. typhimurium	NP_462380	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Independencia de nutriente	Biosíntesis de vitamina B12	4.99.1.3	Sirohidroclorina cobaltoquelatasa (cbiK)	<i>S. typhimurium</i>	NP_460970	
Independencia de nutriente	Biosíntesis de vitamina B12	2.1.1.130	precorrina-2 C20-metiltransferasa (cbiL)	<i>S. typhimurium</i>	NP_460969	
Independencia de nutriente	Biosíntesis de vitamina B12	2.1.1.131	Precorrina-3B metilasa (cbiH)	<i>S. typhimurium</i>	NP_460972	
Independencia de nutriente	Biosíntesis de vitamina B12		CbiG/precorrina metiltransferasa funcional (cbiG)	<i>S. typhimurium</i>	NP_460973	
Independencia de nutriente	Biosíntesis de vitamina B12	2.1.1.133	precorrina-4C11-metiltransferasa (cbiF)	<i>S. typhimurium</i>	NP_460974	
Independencia de nutriente	Biosíntesis de vitamina B12		Proteína de biosíntesis de cobalamina (cbiD)	<i>S. typhimurium</i>	NP_460977	
Independencia de nutriente.	Biosíntesis de vitamina B12	1.3.1.54	Precorrina-A6 reductasa NADPH-dependiente (cbiJ)	<i>S. typhimurium</i>	NP_460971	
Independencia de nutriente	Biosíntesis de vitamina B12	2.1.1.132	precorrina-6B C5,15-metiltransferasa (cbiE)	<i>S. typhimurium</i>	NP_460976	
Independencia de nutriente	Biosíntesis de vitamina B12	2.1.1.132	precorrina-6B C12 decarboxilasa (cbiT)	<i>S. typhimurium</i>	NP_460975	
Independencia de nutriente	Biosíntesis de vitamina B12	5.4.1.2	Precorrina-8X-metil mutasa (cbiC)	<i>S. typhimurium</i>	NP_460978	
Independencia de nutriente	Biosíntesis de vitamina B12	6.3.1.-	ácido cobirínico A,C-diamida sintasa (cbiA)	<i>S. typhimurium</i>	NP_460980	
Independencia de nutriente	Biosíntesis de vitamina B12	2.5.1.17	ácido Cob(I)irínico a,c-diamida adenosilitransferasa (btuR)	<i>S. typhimurium</i>	NP_460677	
Independencia de nutriente	Biosíntesis de vitamina B12	6.3.5.10	ácido cobirínico sintasa (cbiP)	<i>S. typhimurium</i>	NP_460964	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Independencia de nutriente	Biosíntesis de vitamina B12	4.1.1.81	ácido cobírico decarboxilasa (cobD)	<i>S. typhimurium</i>	NP_459636	
Independencia de nutriente	Biosíntesis de vitamina B12	6.3.1.10	Adenosilcobinamida-fosfato sintasa (cbiB)	<i>S. typhimurium</i>	NP_460979	
Independencia de nutriente	Biosíntesis de vitamina B12	3.1.3.73	alfa ribazol-5'-P fosfatasa (cobC)	<i>S. typhimurium</i>	NP_459635	
Independencia de nutriente	Biosíntesis de vitamina B12	2.7.8.26	cobalamin(5'-fosfato) sintasa (cobS)	<i>S. typhimurium</i>	Iv P_460962	
Independencia de nutriente	Biosíntesis de vitamina B12	2.7.7.62	cobinamida fosfato guanilil transferasa (cobU)	<i>S. typhimurium</i>	NP_460963	
Independencia de nutriente	Biosíntesis de vitamina B12	2.4.2.21	nicotinato-nucleótido dimetilbenzimidazol-fosforibosil transferasa (cobT)	<i>S. typhimurium</i>	NP_460961	
Independencia de nutriente	Biosíntesis de vitamina B12 - Transporte de Co2+		sistema de transporte de Co2+ tipo ABC, componente permeasa	<i>S. typhimurium</i>	NP_460968	
Independencia de nutriente	Biosíntesis de vitamina B12 - Transporte de Co2+		sistema de transporte de cobalto tipo ABC, componente permeasa	<i>S. typhimurium</i>	NP_460967	
Independencia de nutriente	Biosíntesis de vitamina B12 - Transporte de Co2+		sistema de transporte de cobalto tipo ABC, componente permeasa	<i>S. typhimurium</i>	NP_461989	
Tolerancia al pH		EC 1.15.1.1	superóxido dismutasa (sodA)	<i>Synechocystis PCC 6803</i>	NP_441347	
Tolerancia al pH		EC 4.1.1.15	glutamato decarboxilasa A (GadA),	<i>Escherichia coli</i>	NP_417974	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia al pH		EC 4.1.1.15	Glutamato decarboxilasa beta (GadB)	<i>Escherichia coli</i>	NP_416010	
Tolerancia al pH			Antipotador de glutamato: ácido gamma-aminobutyrico (GadC)	<i>Escherichia coli</i>	NP_416009	
Tolerancia al pH		EC 4.1.1.19	Arginina descarboxilasa biodegradable (AdiA)	<i>Escherichia coli</i>	NP_418541	
Tolerancia al pH			Antipotador de arginina;arginatina (AdiC)	<i>Escherichia coli</i>	NP_418539	
Tolerancia al pH			Proteína del canal de cloruro (EriC)	<i>Escherichia coli</i>	NP_414697	
Tolerancia al pH			Proteína del canal de cloruro (MrtT)	<i>Escherichia coli</i>	NP_416109	
Tolerancia al pH			Proteína chaperona dnaK2 (DnaK)	<i>Synechocystis PCC 6803</i>	NP_441989	
Tolerancia al pH			ARN polimerasa dirigida por ADN, subunidad sigma (sll0306)	<i>Synechocystis PCC 6803</i>	NP_441950	
Tolerancia al pH			Proteasa Zn-dependiente (sll0528)	<i>Synechocystis PCC 6803</i>	NP_442805	
Tolerancia al pH			Fosfoesterasa metal-dependiente (sll0549)	<i>Synechocystis PCC 6803</i>	NP_442414	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (sll0846)	<i>Synechocystis PCC 6803</i>	NP_441124	
Tolerancia al pH			Proteína de membrana relacionada con estrés ácido (sll0939)	<i>Synechocystis PCC 6803</i>	NP_440194	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia al pH		Proteína de tolerancia a estrés ácido (sll086)	Synechocystis PCC 6803	NP_441667		
Tolerancia al pH		Proteína de tolerancia a estrés ácido (sll483)	Synechocystis PCC 6803	NP_442911		
Tolerancia al pH		proteína de choque térmico/pequeña de 16.6 kDa chaperona molecular (sll514)	Synechocystis PCC 6803	NP_440316		
Tolerancia al pH		EC 2.7.7.13 manosa-1-fosfato guaniltransferasa (sll558)	Synechocystis PCC 6803	NP_441699		
Tolerancia al pH		Factor sigma de ARN polimerasa (sll2012)	Synechocystis PCC 6803	NP_441031		
Tolerancia al pH	EC 3.4.21.102	proteasa de procesamiento carboxilo-terminal (sll0008)	Synechocystis PCC 6803	NP_442119		
Tolerancia al pH		chaperona molecular(sll0093)	Synechocystis PCC 6803	NP_442496		
Tolerancia al pH		Proteína de tolerancia a estrés ácido (sll0270),	Synechocystis PCC 6803	NP_441273		
Tolerancia al pH		Geranilgeranil pirofosfato sintasa (sll0611)	Synechocystis PCC 6803	NP_439899		
Tolerancia al pH		Proteína de tolerancia a estrés ácido (sll0967)	Synechocystis PCC 6803	NP_440193		
Tolerancia al pH		Receptor tipo CheY (sll1214)	Synechocystis PCC 6803	NP_440716		
Tolerancia al pH		Histidina quinasa de transducción de señales (sll1285)	Synechocystis PCC 6803	NP_441610		

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr1413)	Synechocystis PCC 6803	NP_440062	
Tolerancia al pH		EC 1.15.1.1	superóxido dismutasa (sin 516)	Synechocystis PCC 6803	NP_441347	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr 1544)	Synechocystis PCC 6803	NP_440790	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr 1573)	Synechocystis PCC 6803	NP_442902	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr 1674)	Synechocystis PCC 6803	NP_441676	
Tolerancia al pH			proteína de expresión/formación de hidrogenasa (slr 675)	Synechocystis PCC 6803	NP_441677	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (sin 1676)	Synechocystis PCC 6803	NP_441678	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr 1687)	Synechocystis PCC 6803	NP_441698	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (slr 1915)	Synechocystis PCC 6803	NP_440459	
Tolerancia al pH			Esterasa (slr 1916)	Synechocystis PCC 6803	NP440460	
Tolerancia al pH			Proteína componente de hidrogenasa (ssl3044)	Synechocystis PCC 6803	NP_44.1697	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (ssl3769)	Synechocystis PCC 6803	NP_441305	

**Figura 1 (Continuación)**

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia al pH			Proteína de tolerancia a estrés ácido (srt2016)	Synechocystis PCC 6803	NP_440709	
Tolerancia al pH			Proteína de tolerancia a estrés ácido (srt2595)	Synechocystis PCC 6803	NP_440789	
Energía reductora NADH	1.1.1.41		isocitrato deshidrogenasa NAD <sup>+</sup> -dependiente idh1	Saccharomyces cerevisiae	YNL037C	
Energía reductora NADH	1.1.1.41		isocitrato deshidrogenasa NAD <sup>+</sup> -dependiente - idh2	Saccharomyces cerevisiae	YOR136W	
Energía reductora NADH	1.1.1.37		malato deshidrogenasa NADH- ubiquinona oxidoreductasa - OPERON (α-η), genes no numerados individualmente	Escherichia coli	JW3205	
Energía reductora NADH				Rhodobacter capsulatus	AF029365	
Energía reductora NADPH	1.1.1.49		glucosa-6-fosfato deshidrogenasa, zwf	Escherichia coli	JW1841	
Energía reductora NADPH	3.1.1.31		S-Fosfogluconolactonasa - pgI	Escherichia coli	JW0750	
Energía reductora NADPH	1.1.1.44		6-Fosfogluconato deshidrogenasa, gnd	Escherichia coli	JW2011	
Energía reductora NADPH	1.1.1.42		Isocitrato deshidrogenasa NADP-dependiente	Escherichia coli	JW1122	
Energía reductora NADPH	1.1.1.40		enzima málica NADP-dependent	Escherichia coli	JW2447	
Energía reductora NADPH	1.6.1.1		piridina nucleótido transhidrogenasa soluble	Escherichia coli	JW551	
					NP_418397.2	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Energía reducida	NADPH		piridina nucleótido transhidrogenasa unida a membrana, subunidad alfa pntA	<i>Escherichia coli</i>	JW1595	
Energía reducida	NADPH		piridina nucleótido transhidrogenasa unida a membrana, subunidad beta, pntB	<i>Escherichia coli</i>	JW1594	
Tolerancia salina			antíportador Na+/H+ (apnhaP)	<i>Aphanotherce halophytica</i>	BAB69459	
Tolerancia salina	EC 1.11.16		catalasa (katG)	<i>Synechocystis PCC 6803</i>	NP_441295	
Tolerancia salina			proteína relacionada con estrés salino y por cadmio (CW80Cd404)	<i>Chlamydomonas sp. W80</i>	BAE53693	
Tolerancia salina			conservación básica de mama (bbc1)	<i>Chlamydomonas sp. W80</i>	BAA23724	
Tolerancia salina	biosíntesis de betaina		glicina sarcosina metiltransferasa (ApGSMT)	<i>Aphanotherce halophytica</i>	BAC56839	
Tolerancia salina	biosíntesis de betaina		dimetil glicina metiltransferasa (ApDMT)	<i>Aphanotherce halophytica</i>	BAC56840	
Tolerancia térmica	Estabilidad y reparación del fotosistema		Polipeptido esabilizador manganeso del fotosistema II (PsbO)	<i>Synechocystis sp. PCC 6803</i>	NP_441796	
Tolerancia térmica	Estabilidad y reparación del fotosistema		Proteína extrínseca de 12 kDa del complejo PS II (PsbU)	<i>Synechocystis sp. PCC 6803</i>	NP_440167	
Tolerancia térmica	Estabilidad y reparación del fotosistema		Citocromo c550 (PsbV)	<i>Synechocystis sp. PCC 6803</i>	NP_441834	
Tolerancia térmica	Plegamiento de proteínas		Proteína de choque térmico pequeña de 16.6 kDa chaperona molecular chaperona (HspA)	<i>Thermosynechococcus elongatus BP-1</i>	NP_681663	

Figura 1 (Continuación)

Módulo	Diana	EC (si es relevante)	Nombre del gen ejemplar	Organismo	Locus/acceso	Alternativos
Tolerancia térmica	Plegamiento de proteínas		chaperonina 1 de 60 kD (GroEL-1),	<i>Thermosynechocoecus elongatus</i> BP-1	NP_680976	
Tolerancia térmica	Plegamiento de proteínas		chaperonina 2 de 60 kD (GroEL-2)	<i>Thermosynechocoecus elongatus</i> BP-1	NP_682202	
Tolerancia térmica	Plegamiento de proteínas		Co-chaperonina (QroES)	<i>Thermosynechocoecus elongatus</i> BP-1	NP_680977	
Tolerancia térmica	Plegamiento de proteínas		Endopeptidasa (ClpB)	<i>Thermosynechocoecus elongatus</i> BP-1	NP_683242	
Tolerancia térmica	biosíntesis de betaina		glicina sarcosina metiltransferasa (ApGSMT)	<i>Aphanothece halophytica</i>	BAC56839	
Tolerancia térmica	biosíntesis de betaina		dimetilglicina metiltransferasa (ApDMT)	<i>Aphanothece halophytica</i>	BAC56840	

Figura 1

TABLA 2: Genes a regular negativamente o eliminar

Módulo	Diana	Relevancia	EC (si es relevante)	Nombre del gen	Organismo	Locus/acceso
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	tlal	Chlamydomonas reinhardtii	AF534570
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA1	Chlamydomonas reinhardtii	AB122114
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA2	Chlamydomonas reinhardtii	AB122115
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA3	Chlamydomonas reinhardtii	AB122116
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA4	Chlamydomonas reinhardtii	AB122117
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA5	Chlamydomonas reinhardtii	AB122118
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA6	Chlamydomonas reinhardtii	AB122119
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA7	Chlamydomonas reinhardtii	AB122120
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA8	Chlamydomonas reinhardtii	scaffold 3:307726-310521
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCA9	Chlamydomonas reinhardtii	AF244524
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM1	Chlamydomonas reinhardtii	AF495473
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM2	Chlamydomonas reinhardtii	AB051209
Recolección de luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM3	Chlamydomonas reinhardtii	AB051208

Figura 2

Módulo	Diana	Relevancia	EC (si es relevante)	Nombre del gen	Organismo	Locus/acceso
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM4	Chlamydomonas reinhardtii	AF104630
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM5	Chlamydomonas reinhardtii	AF104631
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM6	Chlamydomonas reinhardtii	M24072.1
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM7	Chlamydomonas reinhardtii	AF479779
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM8	Chlamydomonas reinhardtii	AF330793
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM9	Chlamydomonas reinhardtii	AF479778
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM10	Chlamydomonas reinhardtii	AF479777
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCBM11	Chlamydomonas reinhardtii	¿No puede localizarse?
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCB4 <cp29>	Chlamydomonas reinhardtii	AB051207
Recolección de clorofila luz	Tamaño de la antena de clorofila	plantas, algas	n/a	LHCB5 (cp26)	Chlamydomonas reinhardtii	AB050007
Recolección de clorofila luz	Tamaño de la antena de bacterioclorofila	verdes del azufre y no del azufre	n/a	bchK, BChl C sintasa	Chlorobium tepidum	CT1992
Recolección de clorofila luz	Tamaño de la antena de bacterioclorofila	verdes del azufre y no del azufre	n/a	bchQ, BChl metiltransferasa	Chlorobium tepidum	CT1777
Recolección de clorofila luz	Tamaño de la antena de bacterioclorofila	verdes del azufre y no del azufre	n/a 1	bchR, BChl c2 metiltransferasa	Chlorobium tepidum	CT1320
Recolección de clorofila luz	Tamaño de la antena de cianobacterias de fioobilisoma	cpcG2	n/a	Synechocystis PCC 6803	SII1471	

Figura 2 (Continuación)

Módulo	Diana	Relevancia	EC (si es relevante)	Nombre del gen	Organismo	Locus/acceso
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	apcE	Synechococcus PCC 6301	YP_171895	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	cpcG1	Synechocystis PCC 6803	slf2051	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcA, subunidad alfa	Synechococcus PCC 6301	syc0495	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcB, subunidad beta	Synechococcus PCC 6301	syc0496	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcC, péptido enlazador a varillas del ficobilisoma	Synechococcus PCC 6301	syc049B	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcD, proteína CpcD, enlazador ficolina 9K	Synechococcus PCC 6301	syc0497	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	cpcE, proteína de maduración ficolisoma	Synechococcus PCC 6301	syc0494	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcG, enlazador a la varilla-núcleo ficobilisoma	Synechococcus PCC 6301	syc2065	
Recolección de luz	Tamaño de la antena cianobacterias de ficobilisoma	n/a	CpcF, ficolinanobilina lisisa de la subunidad alfa de ficolinina	Synechococcus PCC 6301	YP_171203	
Recolección de luz	Tamaño de la antena bacterias púrpura de bacterias púrpura azufre y no azufre	n/a	pufA, cadena alfa B875 de recolección de luz	Rhodobacter sphaerooides	YP_353331	
Recolección de luz	Tamaño de la antena bacterias púrpura de bacterias púrpura azufre y no azufre	n/a	pufB, cadena beta B875 de recolección de luz	Rhodobacter sphaerooides	YP_353332	
Energía reductora	NADPH	Todos los organismos fotoautotróficos	n/a	pgi, fosfogluco isomerasa	Synechococcus PCC 6301	YP_172776
Thermotolerance	Reguladores transcripcional	Todos los organismos fotoautotróficos	n/a	hrcA, represor de la transcripción inducible por calor	Synechocystis PCC 6803	NP_440130

Figura 2 (Continuación)

Módulo	Diana	Relevancia	EC (si es relevante)	Nombre del gen	Organismo	Locus/acceso
Tolerancia térmica	reguladores transcripcionales	cianobacterias	n/a	hik34, quinasa 34	histidina desA,	Synechocystis PCC SIR1285
Tolerancia térmica	perfíles lipídicos de membranas	Todos organismos fotoautotróficos	los 1.14.19	desA, desaturasa	delta-12-desaturasa	Synechocystis PCC NP_441489
Tolerancia térmica	perfíles lipídicos de membranas	Todos organismos fotoautotróficos	los n/a	desB, desaturasa	delta-15 desaturasa	Synechocystis PCC NP_441622
Tolerancia térmica	perfíles lipídicos de membranas	Todos organismos fotoautotróficos	los 1.14.19.1	desC, estearoil-CoA 9-desaturasa	Synechocystis PCC NP_442430	
Tolerancia térmica	perfíles lipídicos de membranas	Todos organismos fotoautotróficos	los 1.14.19.3	desD, desaturasa	delta-6-desaturasa	Synechocystis PCC NP_441824
Fijación del CO <sub>2</sub>	flujo de acetil-CoA	Todos organismos fotoautotróficos	los 1.3.99.3	acil coenzima A deshidrogenasa	Synechococcus PCC 6301	YP_17104S
Fijación del CO <sub>2</sub>	flujo de acetil-CoA	Todos organismos fotoautotróficos	los 1.1.1.94	glicerol 3-fosfato deshidrogenasa, NADP-dependiente	Synechococcus PCC 7942	YP_401539
Fijación del CO <sub>2</sub>	flujo de acetil-CoA	Todos organismos fotoautotróficos	los 1.1.1.28	lactato deshidrogenasa, IdhA	Synechococcus PCC 6301	YP_170916

Figura 2 (Continuación)

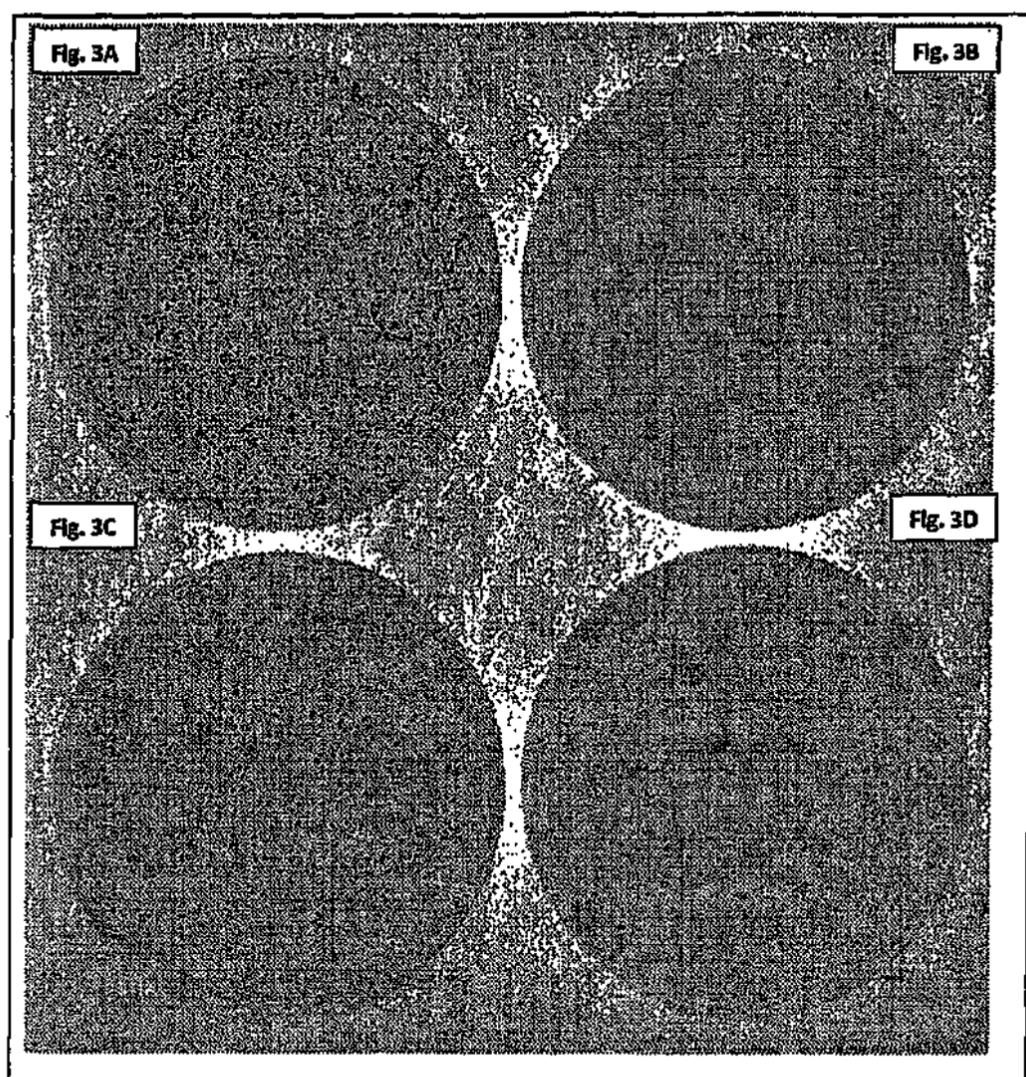


FIG. 3

ES 2 413 482 T3

Figura 4

Met Thr Ile Leu Asn His Thr Leu Gly Phe Pro Arg Val Gly Leu Arg  
1 5 10 15

Arg Glu Leu Lys Lys Ala Gln Glu Ser Tyr Trp Ala Gly Asn Ser Thr  
20 25 30

Arg Glu Glu Leu Leu Ala Val Gly Arg Glu Leu Arg Ala Arg His Trp  
35 40 45

Asp Gln Gln Lys Gin Ala Gly Ile Asp Leu Leu Pro Val Gly Asp Phe  
50 55 60

Ala Trp Tyr Asp His Val Leu Thr Thr Ser Leu Leu Leu Gly Asn Val  
65 70 75 80

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

Phe Arg Ile Gly Arg Gly Arg Ala Pro Thr Gly Glu Pro Ala Ala Ala  
100 105 110

Ala Glu Met Thr Lys Trp Phe Asn Thr Asn Tyr His Tyr Met Val Pro  
115 120 125

Glu Phe Val Lys Gly Gln Gln Phe Lys Leu Thr Trp Thr Gln Leu Leu  
130 135 140

Asp Glu Val Asp Glu Ala Leu Ala Leu Gly His Lys Val Lys Pro Val  
145 150 155 160

Leu Leu Gly Pro Val Thr Trp Leu Trp Leu Gly Lys Val Lys Gly Glu  
165 170 175

Gln Phe Asp Arg Leu Ser Leu Leu Asn Asp Ile Leu Pro Val Tyr Gln  
180 185 190

Gln Val Leu Ala Glu Leu Ala Lys Arg Gly Ile Glu Trp Val Gln Ile  
195 200 205

ES 2 413 482 T3

Asp Glu Pro Ala Leu Val Leu Glu Leu Pro Gln Ala Trp Leu Asp Ala  
210                    215                    220

Tyr Lys Pro Ala Tyr Asp Ala Leu Gln Gly Gln Val Lys Leu Leu Leu  
225                    230                    235                    240

Thr Thr Tyr Phe Glu Gly Val Thr Pro Asn Leu Asp Thr Ile Thr Ala  
245                    250                    255

Leu Pro Val Gln Gly Leu His Val Asp Leu Val His Gly Lys Asp Asp  
260                    265                    270

Val Ala Glu Leu His Lys Arg Leu Pro Ser Asp Trp Leu Leu Ser Ala  
275                    280                    285

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

Tyr Ala Gln Ile Lys Asp Ile Val Gly Lys Arg Asp Leu Trp Val Ala  
305                    310                    315                    320

Ser Ser Cys Ser Leu Leu His Ser Pro Ile Asp Leu Ser Val Glu Thr  
325                    330                    335

Arg Leu Asp Ala Glu Val Lys Ser Trp Phe Ala Phe Ala Leu Gln Lys  
340                    345                    350

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

Ala Ala Leu Ala Glu Trp Ser Ala Pro Ile Gln Ala Arg Arg His Ser  
370                    375                    380

Thr Arg Val His Asn Pro Ala Val Glu Lys Arg Leu Ala Ala Ile Thr  
385                    390                    395                    400

Ala Gln Asp Ser Gln Arg Ala Asn Val Tyr Glu Val Arg Ala Glu Ala  
405                    410                    415

ES 2 413 482 T3

Gln Arg Ala Arg Phe Lys Leu Pro Ala Trp Pro Thr Thr Thr Ile Gly  
420 425 430

Ser Phe Pro Gln Thr Thr Glu Ile Arg Thr Leu Arg Leu Asp Phe Lys  
435 440 445

Lys Gly Asn Leu Asp Ala Asn Asn Tyr Arg Thr Gly Ile Ala Glu His  
450 455 460

Ile Lys Gln Ala Ile Val Glu Gln Glu Arg Leu Gly Leu Asp Val Leu  
465 470 475 480

Val His Gly Glu Ala Glu Arg Asn Asp Met Val Glu Tyr Phe Gly Glu  
485 490 495

His Leu Asp Gly Phe Val Phe Thr Gln Asn Gly Trp Val Gln Ser Tyr  
500 505 510

Gly Ser Arg Cys Val Lys Pro Pro Ile Val Ile Gly Asp Ile Ser Arg  
515 520 525

Pro Ala Pro Ile Thr Val Glu Trp Ala Lys Tyr Ala Gln Ser Leu Thr  
530 535 540

Asp Lys Pro Val Lys Gly Met Leu Thr Gly Pro Val Thr Ile Leu Cys  
545 550 555 560

Trp Ser Phe Pro Arg Glu Asp Val Ser Arg Glu Thr Ile Ala Lys Gln  
565 570 575

Ile Ala Leu Ala Leu Arg Asp Glu Val Ala Asp Leu Glu Ala Ala Gly  
580 585 590

Ile Gly Ile Ile Gln Ile Asp Glu Pro Ala Leu Arg Glu Gly Leu Pro  
595 600 605

Leu Arg Arg Ser Asp Trp Asp Ala Tyr Leu Gln Trp Gly Val Glu Ala  
610 615 620

ES 2 413 482 T3

Phe Arg Ile Asn Ala Ala Val Ala Lys Asp Asp Thr Gln Ile His Thr  
625                       630                       635                       640

His Met Cys Tyr Cys Glu Phe Asn Asp Ile Met Asp Ser Ile Ala Ala  
645                       650                       655

Leu Asp Ala Asp Val Ile Thr Ile Glu Thr Ser Arg Ser Asp Met Glu  
660                       665                       670

Leu Leu Glu Ser Phe Glu Glu Phe Asp Tyr Pro Asn Glu Ile Gly Pro  
675                       680                       685

Gly Val Tyr Asp Ile His Ser Pro Asn Val Pro Ser Val Glu Trp Ile  
690                       695                       700

Glu Ala Leu Leu Lys Lys Ala Ala Lys Arg Ile Pro Ala Glu Arg Leu  
705                       710                       715                       720

Trp Val Asn Pro Asp Cys Gly Leu Lys Thr Arg Gly Trp Pro Glu Thr  
725                       730                       735

Arg Ala Ala Leu Ala Asn Met Val Gln Ala Ala Gln Asn Leu Arg Arg  
740                       745                       750

Gly