

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 633 587**

51 Int. Cl.:

C12P 7/64 (2006.01)

C12N 9/02 (2006.01)

C12N 9/10 (2006.01)

C12N 15/52 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.05.2011 PCT/EP2011/058863**

87 Fecha y número de publicación internacional: **08.12.2011 WO11151298**

96 Fecha de presentación y número de la solicitud europea: **30.05.2011 E 11721794 (3)**

97 Fecha y número de publicación de la concesión europea: **26.04.2017 EP 2576799**

54 Título: **Producción de ácidos grasos por expresión heteróloga de agrupamientos de genes a partir de mixobacterias**

30 Prioridad:

31.05.2010 EP 10005644

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

22.09.2017

73 Titular/es:

**UNIVERSITÄT DES SAARLANDES (100.0%)
Campus Saarbrücken
66123 Saarbrücken, DE**

72 Inventor/es:

**WENZEL, SILKE;
RACHID, SHWAN;
GEMPERLEIN, KATJA y
MÜLLER, ROLF**

74 Agente/Representante:

ARIAS SANZ, Juan

Observaciones :

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 633 587 T3

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

DESCRIPCIÓN

Producción de ácidos grasos por expresión heteróloga de agrupamientos de genes a partir de mixobacterias

- 5 La invención proporciona un método de producción de ácidos grasos poliinsaturados (AGPI) de cadena larga, preferentemente ácidos grasos poliinsaturados omega-3, por expresión heteróloga de agrupamientos de genes que codifican una vía biosintética de AGPI, procedentes de cepas específicas de mixobacterias. Estos agrupamientos de genes únicos pueden clonarse e incorporarse en una amplia serie de organismos que son adecuados para la producción de AGPI en cantidades mejoradas, particularmente AGPI omega-3, por expresión heteróloga.

10

Antecedentes técnicos de la invención

Los ácidos grasos poliinsaturados de cadena larga, incluyendo los de la familia omega-3 que se conocen asimismo como AGPI ω -3 ("omega-3"), son ácidos grasos interesantes en la naturaleza. Son importantes constituyentes de los fosfolípidos que desempeñan un papel en la disminución de la rigidez de la membrana. El ácido eicosapentaenoico (AEP) es un constituyente principal de los fosfolípidos del cerebro humano y sirve como precursor de las prostaglandinas y resolvinas. Otro AGPI destacado de la familia omega-3 es el ácido docosahexaenoico (ADH). La mejora en la función cognitiva y conductual en el desarrollo infantil parece estar correlacionada con altos niveles de este compuesto. Los AGPI omega-3, y en particular para ADH y AEP, han demostrado efectos beneficiosos sobre la salud p. ej., en la prevención del cáncer, artritis reumatoide, enfermedades cardiovasculares, mejora de la función inmunitaria, y salud ocular y cerebral (Teale, MC (ed.) 2006, "*Omega-3 fatty acid research*", Nova Science Publishers. Nueva York, y referencias en el mismo). Debido a estas propiedades beneficiosas, los AGPI omega-3 se están utilizando ampliamente como lípidos nutricionales en complementos de salud y dietéticos y como ingredientes funcionales en un amplio intervalo de alimentos. Los AGPI omega-3 comprenden actualmente uno de los segmentos de mayor y sólido crecimiento del mercado en el sector de la industria de alimentos y bebidas, con un incremento esencial de la demanda en los últimos años. En estos días, el aceite de pescado es la fuente natural más abundante y ampliamente utilizada de ácidos grasos omega-3, pero la fuente designada sufre un exceso de pesca, carece de suministro de aceite de alta calidad con un contenido suficiente de ADH/AEP, y presenta problemas de calidad (olores, desafíos en la fórmula etc.). Procesos alternativos que implican algas y oomicetos como organismos productores se establecen o se están desarrollando (Hinzpeter, *Grasas y Aceites* (2006) 57:336-342; Ward, *Process Biochemistry* (2005) 40:3627-3652). Dado que el suministro de aceite de pescado de alta calidad es cada vez más limitado, se ha tratado de encontrar fuentes biológicas alternativas sostenibles. Se han explorado varios grupos de algas marinas durante más de 20 años y algunos productos basados en la biomasa de algas han llegado entretanto al mercado. Algunos oomicetos pertenecientes al grupo de los stramenopiles (un grupo de organismos eucariotas similares a las algas también conocido como "*Chromophyta*") también se notificaron ocasionalmente para producir los compuestos mencionados anteriormente (p. ej., de los géneros *Achlya* y *Pythium*; (Aki, *J. Ferm Bioeng.* (1998) 86:504-507; Cheng, *Biores. Technol.* (1999) 67:101-110; Athalye, *J. Agric. Food Chem.* (2009) 57:2739-2744). En otros stramenopiles (p. ej., los géneros *Schizochytrium* y *Thraustochytrium*; como se describe en los documentos US 7022512 y WO2007/068997) y en especies de *Amphidinium dinoflagellate* (documento US 2006/0099694), el ADH puede representar hasta el 48 % del contenido de ácidos grasos de las células, que es el contenido más elevado hasta ahora conocido en las eucariotas. No obstante, el cultivo de estos organismos a escala industrial aún supone un reto incluso después de varios años de desarrollo. Otras fuentes biológicas alternativas de AGPI omega-3 hasta el momento descubiertas son eubacterias procariotas [Nichols, *Curr. Opin. Biotechnol.* (1999) 10:240-246; Gentile, *J. Appl. Microbiol.* (2003) 95:1124-1133]. No obstante, la explotación comercial de estos organismos para la producción de AGPI a escala industrial se ve obstaculizada por las características de lento crecimiento de estos microorganismos psicrófilos, así como su rendimiento y productividad inherentemente bajos. La expresión heteróloga de agrupamientos de genes de AGPI omega-3 en los organismos adecuados industriales constituye una alternativa válida a la producción de los AGPI omega-3 deseados a escala industrial, que tiene múltiples ventajas en cuanto a los procesos de producción que utilizan las cepas de tipo natural. Se ha establecido desde hace mucho tiempo que los AGPI omega-3 se biosintetizan de una manera similar a la de los metabolitos secundarios de los polioéctidos tanto en organismos procariotas como eucariotas (véase vista general de Metz, *Science* (2001) 293:290-293), que permite la utilización de técnicas para métodos similares como las que se han establecido en biotecnología microbiana a efectos de mejorar y modificar la producción de agentes antibióticos y anticancerígenos. Para las mixobacterias, este tipo de trabajo en relación con la evaluación de la biosíntesis de los metabolitos secundarios se ha descrito y se explica en revisiones recientes. (Wenzel, *Curr. Opin. Biotechnol.* (2005), 16:594-606; Wenzel, *Curr. Opin. Drug Discov & Develop.* (2009), 12 (2):220-230; Wenzel, *Nat. Prod. Rep.* (2009), 26 (11):1385-1407).

60

La organización de los genes de AGPI en agrupamientos de genes permite su clonación y transferencia a huéspedes heterólogos, que *per se* pueden ser organismos procariotas o eucariotas.

Los agrupamientos de genes que codifican las enzimas de la vía sintética para la biosíntesis de AGPI omega-3 se conocen de diversas bacterias marinas, incluyendo especies de los géneros *Moritella* (Tanaka, *Biotechnol. Lett.* (1999) 21:641-646; *Morita*, *Biochem Soc Trans* 28:943-945 (2000)), *Photobacterium* (Allen, *Microbiology* (2002) 148:1903-1913), y *Shewanella* (Lee, *J. Microbiol. Biotechnol.* (2009) 19:881-887). Tales agrupamientos de genes biosintéticos de AGPI omega-3 bacterianos ya se transfirieron y se expresaron en *Escherichia coli* (Lee, *Biotechnol.*

65

Bioproc. Engin. (2006) 11:510-515; Orikasa, *Biotechnol. Lett.* (2006) 28:1841-1847; Orikasa, *Biotechnol. Lett.* (2007) 29:803-812). Además, la producción heteróloga de AEP se potenció mediante la sustitución de secuencias promotoras en los agrupamientos de genes implicados en la biosíntesis (Lee, *Biotechnol. Lett.* (2008) 30:2139-2142). El agrupamiento de genes de AEP a partir de *Shewanella sp.* también se expresó en la cianobacteria marina transgénica del género *Synechococcus* (Takeyama, *Microbiology* (1997) 143 (Pt 8):2725-2731 Orikasa, *Biotechnol. Lett.* (2007) 29:803-809), se observó una producción heteróloga de AEP potenciada en células de *E. coli* que co-expresan genes implicados en la biosíntesis de AEP y fragmentos de ADN exógeno incluyendo un gen de la catalasa de alto rendimiento. Los rendimientos de AEP aumentaron del 3 % al 12 % del contenido total de ácidos grasos. Jiang (*Methods in Enzymology* (2009) 459,80-96) ha resumido la investigación previa sobre la caracterización de los dominios de la proteína transportadora de acilos en tándem importantes en las sintetasas de ácidos grasos poliinsaturados, también representó diversos estudios que implicaban la expresión heteróloga de los respectivos genes en *E. coli*. Estos ejemplos muestran que es factible alcanzar la producción heteróloga de genes de AGPI omega-3 bacterianos, a pesar de que ningún producto comercial basado en tales técnicas ha resultado hasta el momento. Hasta ahora nunca se han identificado agrupamientos de genes de AGPI omega-3 mixobacterianos y por consiguiente no se han expresado de manera heteróloga. Las vías biosintéticas de productos naturales mixobacterianos se han expresado con éxito en una variedad de huéspedes heterólogos, incluyendo *Pseudomonas putida* (Gross, *Chem. Biol.*, (2006) 13:1253-1264; Wenzel, *Chem. Biol.* (2005) 12 (3):349-356).

Debido a la gran importancia comercial de estos productos, existen varios ejemplos para la viabilidad de la expresión heteróloga de los genes implicados en la biosíntesis de AGPI omega-3 a partir de varios organismos eucariotas y procariontas en plantas y hongos transgénicos. Como se ejemplifica por Domergue (*Plant Physiol.* (2003) 131, 1648-1660) y los genes biosintéticos de ácidos grasos de la alga stramenopile, *Phaeodactylum tricornutum*, incluso la aclaración de la biosíntesis de ciertas fuentes biológicas marinas que no son fáciles de cultivar en el laboratorio han implicado a menudo su expresión heteróloga en *Saccharomyces cerevisiae* u otras levaduras. Cipak (*Free Radic. Biol. & Med.* (2006) 40, 897-906) expresó un gen desaturasa a partir de la *Hevea brasiliensis*, árbol del caucho y también en *S. cerevisiae*. Tonon (2005, *Plant Physiol.* 138, 402-408) utilizó el mismo huésped para expresar y caracterizar una acil-coenzima A (acil-CoA) sintetasa que se descubrió de la minería genómica en la diatomea *Thalassiosira pseudonana*. Hsiao (2007, *Mar. Biotechnol.* 9,154-165) también utilizó *S. cerevisiae* como huésped para la expresión heteróloga y la caracterización funcional de una delta-6 desaturasa a partir de la microalga marina *Glossomastix chryso-plasta*. Lee (2008, *Biotechnol. Bioproc. Engin.*, 13, 524-532) ha demostrado con éxito la actividad de la delta-9 elongasa, una enzima crucial en la biosíntesis de AGPI omega-3 a partir de stramenopile, *Thraustochytrium aureum* mediante la expresión heteróloga en la levadura *Pichia pastoris*. Li (*Biotechnol. Lett.*, 31, 1011-1017) notificó una mejora en la producción de ácido araquidónico y ácido eicosapentaenoico al aumentar el número de copias de los genes que codifican la desaturasa y elongasa de ácidos grasos a partir de la alga *Phaeodactylum tricornutum* en *Pichia pastoris* como huésped heterólogo; no obstante, solo se alcanzaron porcentajes relativamente bajos de AGPI de 0,1 y 0,1 %, respectivamente del contenido de ácidos grasos totales de *Pichia pastoris*. Esta última levadura, así como *Yarrowia lipolytica*, *Hansenula anomala*, y otras levaduras metilotróficas y/u oleaginosas, parecen ideales para la producción de AGPI, ya que son organismos productores industriales bien establecidos (Banlar, *Appl. Microbiol. Biotechnol.* (2009) 84, 47-865; Silva, *J. Food, Agric & Environment.* 2009, 7,268.273). Además, las levaduras oleaginosas, que son especialmente organismos de producción preferentes y que son capaces de crecer en ambientes altamente lipófilos, pueden acumular grandes cantidades de lípidos y pueden utilizar un amplio intervalo de hidrocarburos como sustratos. Otras aplicaciones de hongos semejantes a levaduras en microbiología y biotecnología industrial se han resumido por Porro (2005, *Mol. Biotechnol.* 31, 245-259) e Idiris (2010, *Appl. Microbiol. Biotechnol.* 86:403-417).

Graham (*Curr. Opin. Biotechnol.* (2007) 18 (2), 142-147) ha resumido el estado de la técnica en la ingeniería metabólica de las plantas transgénicas para la biosíntesis de AGPI omega-3. Destacó, en particular, los recientes avances en el diseño racional de los cultivos de semillas oleaginosas con un contenido elevado de AGPI omega-3, pero también discutió varios factores que pueden ser responsables de los hasta ahora bajos rendimientos inherentes observados, que hasta el momento han impedido el desarrollo de procesos de producción competitivos de AGPI mediante el uso de semillas oleaginosas u otras plantas de semillas. Taylor (*Plant Biotechnol. J.* (2009) 7, 925-938) ha identificado y clonado recientemente la 3-cetoacil-CoA sintasa de *Cardamine graeca* y notificó su expresión heteróloga en semillas oleaginosas de *Brassica*, resultando en un aumento de las tasas de producción de los AGPI, ácido nervónico y notificó un aumento de 15 veces del ácido nervónico en el huésped heterólogo. No obstante, el ácido nervónico es un AGPI omega-9, en lugar de un omega-3, y por el momento no ha alcanzado la importancia comercial de ADH y AEP. Además, el ácido nervónico se deriva de la elongación del ácido oleico en la naturaleza.

Lee *et al.* (*Biotech. Bioproc. Eng.* (2006) 11, 510-515) describen la expresión heteróloga de un agrupamiento de genes putativo de polioctétido sintetasa de *Shewanella oneidensis* MR-1 para la producción de AEP en *E. coli*, el agrupamiento de genes se diferencia de la presente invención al menos en que no comprende una acil-glicerol-fosfato aciltransferasa (AGPAT) como se describe en la presente memoria. El agrupamiento de genes se identifica solo como un fragmento de ADN de 20,195 kb obtenido por PCR larga utilizando un par de cebadores, cuyas secuencias nucleotídicas se dan, y por los tamaños del fragmento de restricción generado por BglII y NdeI.

Orikasa *et al.* (*Biotechnol. Lett.* (2006) 28, 1841-1847) describen la síntesis de ADH en *E. coli* mediante la expresión de un agrupamiento de genes a partir de *Moritella marina* MP-1. El agrupamiento de genes utilizado por Orikasa *et*

al., tampoco comprende una acil-glicerol-fosfato aciltransferasa (AGPAT) como se describe en la presente memoria pero en su lugar se describe otra aciltransferasa que es bastante diferente en términos de identidad proteica (Ej. 7).

5 Otros documentos se refieren a la producción de ácidos grasos mediante el cultivo de productores naturales de los mismos. El documento WO 2010/063451A1 enseña en general un método de producción de ácidos grasos poliinsaturados mediante el cultivo de cepas específicas mixobacterianas, pero falla, no obstante, en revelar cualquier secuencia específica que podría utilizarse en la expresión heteróloga. De manera similar, Garcia R *et al.*, (*J Bacteriol.* abril de 2011; 193(8)1930-42) revisa el contenido intracelular de los ácidos grasos en diversas cepas mixobacterianas, pero ni se refiere o desvela las secuencias génicas específicas requeridas para la producción de

10 ácidos grasos.
Schneiker *et al.* (*Nature Biotechnology* (2007) 25, 1281-1289) describen la secuenciación del genoma completo de *Sorangium cellulosum* e identifican agrupamientos de genes de policétido sintetasa implicados en la síntesis de compuestos antibióticos individuales. Una comparación con el genoma de *Myxococcus xanthus* reveló una carencia de sintenia global.

15 Dickschat *et al.* (*Org. Biomol. Chem.* (2005) 3, 2824-2831) describen el análisis del perfil de los ácidos grasos de *Stigmatella aurantiaca* utilizando GC-MS, y las etapas de síntesis de la biosíntesis.

20 Hay una demanda de métodos para la producción de AGPI, particularmente AGPI que contienen 3, 4 o más dobles enlaces, cuyos métodos biosintetizan los AGPI *de novo*, en lugar de modificar, p. ej., alargar, otros ácidos grasos que, a su vez, sirven como precursores.

Objeto de la invención

25 Es un objeto de la invención proporcionar un método para la producción de AGPI que tiene ventajas en comparación con los métodos de la técnica anterior, preferentemente un proceso de producción utilizando fermentación de un microorganismo huésped que expresa las enzimas pertinentes de la vía sintética.

Descripción general de la invención

30 La invención consigue el objeto por el proceso de producción como se define en las reivindicaciones, especialmente al proporcionar las secuencias de ADN que codifican las enzimas de la vía sintética para la síntesis de AGPI, un microorganismo que se manipula genéticamente para contener estas secuencias de ADN para la expresión heteróloga, y un proceso de producción de uno o más ácidos grasos poliinsaturados mediante el cultivo del microorganismo que se manipula genéticamente para contener secuencias de ADN que codifican las enzimas de la vía sintética para la síntesis de AGPI para la expresión génica heteróloga, comprendiendo el proceso preferentemente las etapas que consisten en

40 (i) proporcionar un organismo de producción que comprenda un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados, siendo derivado dicho agrupamiento de genes de un organismo fuente que difiere del organismo de producción; y que abarca una subsecuencia ER que codifica una enoil reductasa y una subsecuencia AT que codifica una aciltransferasa, en la que la subsecuencia AT se ubica aguas abajo con respecto a la subsecuencia ER;

45 (ii) cultivar el organismo de producción de la etapa (i) en presencia de una fuente de carbono fermentable, también denominada cultivo, conforme la cual se producen uno o más ácidos grasos poliinsaturados; y

(iii) opcionalmente, recuperar uno o más ácidos grasos poliinsaturados, preferentemente aislar los ácidos grasos poliinsaturados del caldo de fermentación y/o del organismo de producción.

50 Los inventores han descubierto sorprendentemente que los agrupamientos de genes biosintéticos de AGPI mixobacterianos, p. ej., derivados de *Sorangium cellulosum* y de *Aetherobacter fasciculatus*, que abarcan cada uno los genes pfa1, pfa2 y pfa3 y, por consiguiente, codifican las proteínas correspondientes Pfa1, Pfa2, y Pfa3, pueden utilizarse ventajosamente en la producción de AGPI. El análisis de los loci biosintéticos muestra que la disposición de los genes, así como la organización de los dominios catalíticos en las proteínas biosintéticas difieren significativamente de las secuencias publicadas de los agrupamientos de genes de AGPI de diversos organismos.

55 Para fines de la invención, las expresiones secuencia de aminoácidos, péptido, proteína, y una sección de uno de estos pueden utilizarse de manera intercambiable. La referencia a la función de una secuencia de aminoácidos codificada por una secuencia de ácidos nucleicos puede incluir la referencia a la secuencia de ácidos nucleicos codificante, y además, la referencia a la función catalítica de un dominio puede tomarse como referencia a la sección de secuencia de ácidos nucleicos que codifica la sección del péptido que tiene la función catalítica, cuya sección de

60 secuencia de ácidos nucleicos también puede denominarse dominio de ácidos nucleicos. Las secuencias y secciones de las mismas que se refieren como un dominio pueden incluir uno o más centros catalíticos o dominios catalíticos, que p. ej., se designan de acuerdo con su función catalítica. Las identidades y similitudes de secuencia son preferentemente como se determina por el algoritmo ClustalW o Geneious que se describe en la presente memoria. Debido a la universalidad del código genético, una secuencia de ácidos nucleicos que codifica un dominio se determina por la secuencia de aminoácidos del dominio, en la que preferentemente la secuencia de ácidos nucleicos tiene un uso de codones o un sesgo de codones del microorganismo huésped que se manipula

genéticamente para contener genes heterólogos que codifican el agrupamiento de genes de la enzima sintética de AGPI, p. ej., una secuencia de ácidos nucleicos que codifica el agrupamiento de genes y que tiene el uso del codón del microorganismo huésped.

5 En general, preferentemente el gen pfa 1 contiene un dominio I que comprende una secuencia de nucleótidos que codifica una enoil reductasa (ER), el gen pfa2 contiene un dominio IIa que comprende una secuencia de nucleótidos que codifica un dominio de cetosintasa (CS) y que codifica un dominio de malonil-CoA-transacilasa (MAT), un dominio IIb que comprende una secuencia de nucleótidos que codifica al menos uno, preferentemente tres, cuatro o cinco dominios de la proteína transportadora de acilos (PPA), y un dominio IIc que comprende una secuencia de nucleótidos que codifica un dominio de cetorreductasa (CR) y que codifica un dominio de deshidratasa (DH), y el gen pfa3 contiene un dominio IIIa que comprende una secuencia de nucleótidos que codifica un dominio de cetosintasa (CS) y que codifica un dominio del factor de longitud de cadena (FLC), un dominio IIIb que comprende una secuencia de nucleótidos que codifica uno, preferentemente dos dominios de deshidratasa (DH), y contiene un dominio IV* que comprende una secuencia de nucleótidos que codifica una acil-glicerol-fosfato aciltransferasa (AGPAT), en la que preferentemente los dominios se disponen de 5' a 3' en los respectivos genes. En el agrupamiento de genes sintéticos de *Aetherobacter*, entre la región IIIa y la región IIIb, se ha descubierto un dominio IV que codifica una acil transferasa. Por consiguiente, un dominio IV opcional que comprende una secuencia de nucleótidos que codifica un dominio de aciltransferasa (AT) se dispone en el gen pfa3 gen entre el dominio IIIa y el dominio IIIb, especialmente para la producción de AEP y/o ADH. La acil-glicerol-fosfato aciltransferasa (AGPAT) codificada por el agrupamientos de genes, p. ej., por el dominio IV*, contenida preferentemente en pfa3, es una realización específica de una aciltransferasa (AT).

25 Generalmente, el dominio I codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 12 o 13, SEQ ID NO: 43 o 44, SEQ ID NO: 74 o 75, SEQ ID NO: 108 o 109, SEQ ID NO: 132 o 133 y/o SEQ ID NO: 148 o 149;

el dominio IIa codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 15 o 16, SEQ ID NO: 46 o 47, SEQ ID NO: 77 o 78, SEQ ID NO: 111 o 112 y/o SEQ ID NO: 151 o 152;

30 el dominio IIb codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 18 o 19, SEQ ID NO: 49 o 50, SEQ ID NO: 80 o 81, SEQ ID NO: 114 o 115 y/o SEQ ID NO: 154 o 155;

el dominio IIc codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 21 o 22, SEQ ID NO: 52 o 53, SEQ ID NO: 83 o 84, SEQ ID NO: 117 o 118 y/o SEQ ID NO: 157 o 158;

35 el dominio IIIa codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 24 o 25, SEQ ID NO: 55 o 56, SEQ ID NO: 86 o 87, SEQ ID NO: 120 o 121 y/o SEQ ID NO: 160 o 161;

40 el dominio IIIb codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 27 o 28, SEQ ID NO: 58 o 59, SEQ ID NO: 89 o 90, SEQ ID NO: 123 o 124 y/o SEQ ID NO: 163 o 164.

El dominio IV opcional codifica una secuencia de aminoácidos que tiene una identidad de al menos el 85 % con respecto a una de SEQ ID NO: 92 o 93, SEQ ID NO: 126 o 127 y/o SEQ ID NO: 166 o 167.

45 Concretamente, la invención proporciona un proceso de producción de ácidos grasos poliinsaturados mediante la expresión génica heteróloga de las enzimas de la vía biosintética de ácidos grasos poliinsaturados en un organismo de producción, que comprende el cultivo del organismo de producción en presencia de una fuente de carbono fermentable,

50 - en el que el organismo de producción se manipula genéticamente para comprender un primer gen, un segundo gen y un tercer gen,

- en el que el primer gen contiene un dominio I que codifica una enoil reductasa que tiene una identidad de secuencia de aminoácidos de al menos el 59 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 13, SEQ ID NO: 44, SEQ ID NO: 75, SEQ ID NO: 109, SEQ ID NO: 133 y/o SEQ ID NO: 149, y

55 - en el que el segundo gen contiene un dominio IIa que codifica una cetosintasa y una malonil-CoA-transacilasa que tiene una identidad de secuencia de aminoácidos de al menos el 57 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 16, SEQ ID NO: 47, SEQ ID NO: 78 y/o SEQ ID NO: 112, un dominio IIb que codifica al menos una proteína transportadora de acilos que tiene una identidad de secuencia de aminoácidos de al menos el 26 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 19, SEQ ID NO: 50, SEQ ID NO: 81, SEQ ID NO: 115 y/o SEQ ID NO: 155; un dominio IIc que codifica una cetorreductasa y una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 47 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 22, SEQ ID NO: 53, SEQ ID NO: 84, SEQ ID NO: 118 y/o SEQ ID NO: 158; y

60 - en el que el tercer gen contiene un dominio IIIa que codifica una cetosintasa y un factor de longitud de cadena que tiene una identidad de secuencia de aminoácidos de al menos el 46 % con respecto a una

secuencia de aminoácidos del grupo que contiene SEQ ID NO: 25, SEQ ID NO: 56, SEQ ID NO: 87, SEQ ID NO: 121 y/o SEQ ID NO: 161; y un dominio IIIb que codifica al menos una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 42 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 28, SEQ ID NO: 59, SEQ ID NO: 90, SEQ ID NO: 124 y/o SEQ ID NO: 164, y un dominio que codifica una aciltransferasa (AT) caracterizado por que la aciltransferasa codificada es una acil-glicerol-fosfato aciltransferasa (AGPAT) que tiene una identidad de secuencia de aminoácidos de al menos el 50 % con respecto a una secuencia de aminoácidos codificada por el dominio IV* contenido en el grupo que comprende SEQ ID NO: 30, SEQ ID NO: 61, SEQ ID NO: 95, SEQ ID NO: 129, SEQ ID NO: 135 y/o SEQ ID NO: 169, y secuencias de aminoácidos codificadas por una sección de secuencia de ácidos nucleicos contenida en el grupo que comprende SEQ ID NO: 29, SEQ ID NO: 60, SEQ ID NO: 94, SEQ ID NO: 128, SEQ ID NO: 134 y/o SEQ ID NO: 168.

Preferentemente, cada uno de los genes, regiones y/o dominios codifica una secuencia de aminoácidos que tiene una identidad de al menos 60 %, al menos 85 %, al menos 90 %, y preferentemente al menos 95 % o al menos 98 % respecto a una o más de las secuencias de aminoácidos para cada dominio como se da en la presente memoria, o como se codifica por una secuencia de ácidos nucleicos para cada dominio como se da en la presente memoria.

Preferentemente, el dominio IV* codifica una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 50 %, 75 % o al menos el 84 %, más preferentemente de al menos 90 %, 95 % o al menos 97 % o 98 % con respecto a una de las secuencias de aminoácidos dada para el dominio IV*, p. ej., a una secuencia de aminoácidos del dominio IV* contenida en el grupo que comprende o consiste en SEQ ID NO: 30, SEQ ID NO: 61 y/o una secuencia de aminoácidos del dominio IV* contenida en el grupo que comprende o consiste en SEQ ID NO: 95, SEQ ID NO: 129 o 130, SEQ ID NO: 135 y/o SEQ ID NO: 169, o como se codifica por una sección de secuencia de ácidos nucleicos que codifica un dominio IV*, p. ej., una sección de secuencia de ácidos nucleicos contenida en el grupo que comprende o consiste en SEQ ID NO: 29, SEQ ID NO: 60 y/o en el grupo que comprende o consiste en SEQ ID NO: 94, SEQ ID NO: 128, SEQ ID NO: 134 y/o SEQ ID NO: 168.

Además, el dominio IV* se contiene preferentemente en una secuencia de ácidos nucleicos que codifica el producto génico Pfa3, preferentemente que codifica Pfa3 que tiene una identidad de secuencia de aminoácidos de al menos el 28 %, 30 %, al menos 32,5 %, o al menos 35 %, preferentemente al menos 40 %, al menos 42,5 %, al menos 54 %, más preferente al menos 84 %, 90 % o 95 % con respecto a una de las secuencias de aminoácidos contenida en el grupo que comprende o consiste en SEQ ID NO: 10 y/o SEQ ID NO: 41 y/o en el grupo que comprende o consiste en SEQ ID NO: 72, SEQ ID NO: 106 y/o SEQ ID NO: 146.

35 Breve descripción de los dibujos

La Figura 1 muestra una comparación esquemática de agrupamientos de genes que codifican las vías biosintéticas de ácidos grasos poliinsaturados. Los organismos fuente se mencionan a la izquierda. Las abreviaturas (códigos proteicos) tienen los siguientes significados: ER = enoil reductasa; CS = beta-cetoacil sintasa; MAT = malonil-CoA-transacilasa; CR = cetorreductasa; DH = deshidratasa; FLC = factor de longitud de cadena; AT = aciltransferasa; PPA = proteína transportadora de acilos; AGPAT = acil-glicerol-fosfato aciltransferasa.

La Figura 2 muestra una construcción de ácido nucleico para la expresión heteróloga del agrupamiento de genes biosintéticos de AGPI a partir de *Sorangium cellulosum* So ce56 en otra mixobacteria p. ej., *Myxococcus xanthus* DK1622. Los genes biosintéticos de AGPI (pfa1, pfa2 y pfa3) están bajo el control del promotor Ptet, la construcción se integra al azar en el cromosoma del huésped a través de la transferencia del agrupamiento de genes mediada por transposones, transposasa: gen que codifica la transposasa, RI: repetición invertida, amp^R: gen de resistencia a ampicilina, Bsd^R: gen de resistencia a blasticidina, neo^R: gen de resistencia a kanamicina, regulador tet: gen que codifica el represor del promotor Ptet, Ptet: promotor inducible de tetraciclina, oriT: origen de transferencia.

La Figura 3A muestra un cromatograma de cromatografía de gases (CG) de extractos de ácidos grasos celulares de *S. cellulosum* So ce56 de tipo de natural,

La Figura 3B muestra un cromatograma de cromatografía de gases (CG) de extractos de ácidos grasos celulares de *M. xanthus::p15A-AGPI* que alberga la construcción de expresión mostrada en la Figura 2,

La Figura 3C muestra un cromatograma de cromatografía de gases (CG) de extractos de ácidos grasos celulares de *M. xanthus* DK1622 de tipo natural.

La Figura 4 muestra una construcción de ácidos nucleicos para la expresión heteróloga del agrupamiento de genes biosintéticos de AGPI a partir de *Aetherobacter fasciculatus* DSM 21835 en otra mixobacteria p. ej., *Myxococcus xanthus* DK1622. Los genes biosintéticos de AGPI (pfa1, pfa2 y pfa3) están bajo el control del promotor P_{tet}, la construcción se integra en el cromosoma del huésped a través de sitios MX9 attB. mx9: gen que codifica la MX9 integrasa y que contiene el sitio de unión del fago MX9 (attP), zeo^R: gen de resistencia a zeocina, amp^R: gen de resistencia a ampicilina, p15A ori: origen de replicación de p15A, neo^R: gen de resistencia a kanamicina, P_{tet}: promotor inducible de tetraciclina.

La Figura 5 muestra un análisis por CG de los extractos de ácidos grasos celulares de *M. xanthus* DK1622 de tipo

natural (A) y de extractos de ácidos grasos celulares de *M. xanthus*::pPfaAf-P_{let}-mx9.2 que alberga el agrupamiento de genes que contiene pfa1, pfa2 y pfa3 de la SEQ ID NO 63 (B). Los lípidos celulares se hidrolizaron y los ácidos grasos se derivatizaron para que pudiesen ser analizados por GC-MS. Se muestran derivados de los siguientes ácidos grasos: ácido hexadecanoico (16:0;1), ácido iso-heptadienoico (iso-17:2 ω 5, 11; 2) ácido heptadecenoico (17:1 ω 7; 3), ácido iso-heptadecenoico (iso-17:1 ω 5; 4), ácido iso-heptadecanoico (iso-17:0; 5), ácido iso-3-hidroxi-pentadecanoico (iso-3-OH-15:0; 6), ácido hexadecenoico (16:1 ω 7; 7), ácido hidroxi-hexadecanoico (2-OH-16:0; 3-OH-16:0; 8), ácido hidroxi-heptadecanoico (2-OH-17:0; 3-OH-17:0; 9), ácido eicosapentaenoico (20:5 ω 3, 6, 9, 12, 15; 10), ácido iso-pentadecanoico (iso-15:0; 11), y ácido docosahexaenoico (22:6 ω 3, 6, 9, 12, 15, 18; 12).

10 La Figura 6 muestra un árbol del método de Neighbour-Joining generado a partir de NCBI-BLASTn que muestra las afinidades de las secuencias de ADNr 16S de *Aetherobacter fasciculatus*.

15 La Figura 7 muestra un árbol del método de Neighbour-Joining basado en una secuencia del gen de ADNr 16S mixobacteriano, y que muestra la posición filogenética de AEP y ADH que producen cepas de *Sorangium cellulsum*.

Breve descripción del listado de secuencias

<i>Sorangium cellulsum</i> (Soce56)	
Secuencia de ADN que codifica todo el agrupamiento implicado en la biosíntesis de AGPI: SEQ ID NO: 1;	nt (nucleótidos) 1..1.650 de SEQ ID NO: 1 son codones (CDS) de pfa1 que codifica la proteína: SEQ ID NO: 2
	nt 1.677..9.389 de SEQ ID NO: 1 son CDS de pfa2 que codifica la proteína: SEQ ID NO: 3
	nt 9.386..17.176 de SEQ ID NO: 1 son CDS de pfa3 que codifica la proteína: SEQ ID NO: 4

Secuencia de ADN que codifica pfa1: SEQ ID NO: 5	CDS de pfa1 que codifica la proteína de SEQ ID NO: 6
Secuencia de ADN que codifica pfa2: SEQ ID NO: 7	CDS de pfa2 que codifica la proteína de SEQ ID NO: 8
Secuencia de ADN que codifica pfa3 SEQ ID NO: 9	CDS de pfa3 que codifica la proteína de SEQ ID NO: 10
Secuencia de ADN que codifica el dominio I: SEQ ID NO: 11	CDS de I que codifica la proteína de SEQ ID NO: 12
Proteína del dominio I: SEQ ID NO: 13	
Secuencia de ADN que codifica el dominio IIa: SEQ ID NO: 14	CDS de IIa que codifica la proteína de SEQ ID NO: 15
Proteína del dominio IIa: SEQ ID NO: 16	
Secuencia de ADN que codifica el dominio IIb: SEQ ID NO: 17	CDS de IIb que codifica la proteína de SEQ ID NO: 18
Proteína del dominio de IIb: SEQ ID NO: 19	
Secuencia de ADN que codifica el dominio IIc: SEQ ID NO: 20	CDS de IIc que codifica la proteína de SEQ ID NO: 21
Proteína del dominio de IIc: SEQ ID NO: 22	
Secuencia de ADN que codifica el dominio IIIa: SEQ ID NO: 23	CDS de IIIa que codifica la proteína de SEQ ID NO: 24
Proteína del dominio de IIIa: SEQ ID NO: 25	
Secuencia de ADN que codifica el dominio IIIb: SEQ ID NO: 26	CDS de IIIb que codifica la proteína de SEQ ID NO: 27
Proteína del dominio de IIIb: SEQ ID NO: 28	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 29	CDS de IV* que codifica la proteína de SEQ ID NO: 30
Proteína del dominio de IV*: SEQ ID NO: 31	

20

<i>Sorangium cellulsum</i> Soce 377	
Secuencia de ADN que codifica todo el agrupamiento implicado en la biosíntesis de AGPI: SEQ ID NO: 32	nt 1..1.650 de SEQ ID NO: 32 son CDS de pfa1 que codifica la proteína: SEQ ID NO: 33
	nt 1.677..9.380 de SEQ ID NO: 32 son CDS de pfa2 que codifica la proteína: SEQ ID NO: 34
	nt 9.377..17.221 de SEQ ID NO: 32 son CDS de pfa3 que codifica la proteína: SEQ ID NO: 35
Secuencia de ADN que codifica pfa1: SEQ ID NO: 36	CDS de pfa1 que codifica la proteína de SEQ ID NO: 37
Secuencia de ADN que codifica pfa2: SEQ ID NO: 38	CDS de pfa2 que codifica la proteína de SEQ ID NO: 39
Secuencia de ADN que codifica pfa3 de SEQ ID NO: 40	CDS de pfa3 que codifica la proteína de SEQ ID NO: 41
Secuencia de ADN que codifica el dominio I: SEQ ID	CDS de I que codifica la proteína de SEQ ID NO: 43

ES 2 633 587 T3

<i>Sorangium cellulosum</i> Soce 377	
NO: 42	
Proteína del dominio I: SEQ ID NO: 44	
Secuencia de ADN que codifica el dominio Ila: SEQ ID NO: 45	CDS de Ila que codifica la proteína de SEQ ID NO: 46
Proteína del dominio Ila: SEQ ID NO: 47	
Secuencia de ADN que codifica el dominio IIb: SEQ ID NO: 48	CDS de IIb que codifica la proteína de SEQ ID NO: 49
Proteína del dominio de IIb: SEQ ID NO: 50	
Secuencia de ADN que codifica el dominio IIc: SEQ ID NO: 51	CDS de IIc que codifica la proteína de SEQ ID NO: 52
Proteína del dominio de IIc: SEQ ID NO: 53	
Secuencia de ADN que codifica el dominio IIIa: SEQ ID NO: 54	CDS de IIIa que codifica la proteína de SEQ ID NO: 55
Proteína del dominio de IIIa: SEQ ID NO: 56	
Secuencia de ADN que codifica el dominio IIIb: SEQ ID NO: 57	CDS de IIIb que codifica la proteína de SEQ ID NO: 58
Proteína del dominio de IIIb: SEQ ID NO: 59	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 60	CDS de IV* que codifica la proteína de SEQ ID NO: 61
Proteína del dominio de IV*: SEQ ID NO: 62	

<i>Aethrobacter fasciculatus</i> SBSr002	
Secuencia de ADN que codifica todo el agrupamiento implicado en la biosíntesis de AGPI: SEQ ID NO: 63	nt 1..1.608 de SEQ ID NO: 63 son codones (CDS) de pfa1 que codifica la proteína: SEQ ID NO: 64
	nt 1.644..8.312 de SEQ ID NO: 63 son CDS de pfa2 que codifica la proteína: SEQ ID NO: 65
	nt 8.309..16.219 de SEQ ID NO: 63 son CDS de pfa3 que codifica la proteína: SEQ ID NO: 66
Secuencia de ADN que codifica pfa1: SEQ ID NO: 67	CDS de pfa1 que codifica la proteína de SEQ ID NO: 68
Secuencia de ADN que codifica pfa2: SEQ ID NO: 69	CDS de pfa2 que codifica la proteína de SEQ ID NO: 70
Secuencia de ADN que codifica pfa3 de SEQ ID NO: 71	CDS de pfa3 que codifica la proteína de SEQ ID NO: 72
Secuencia de ADN que codifica el dominio I: SEQ ID NO: 73	CDS de I que codifica la proteína de SEQ ID NO: 74
Proteína del dominio I: SEQ ID NO: 75	
Secuencia de ADN que codifica el dominio Ila: SEQ ID NO: 76	CDS de Ila que codifica la proteína de SEQ ID NO: 77
Proteína del dominio Ila: SEQ ID NO: 78	
Secuencia de ADN que codifica el dominio IIb: SEQ ID NO: 79	CDS de IIb que codifica la proteína de SEQ ID NO: 80
Proteína del dominio de IIb: SEQ ID NO: 81	
Secuencia de ADN que codifica el dominio IIc: SEQ ID NO: 82	CDS de IIc que codifica la proteína de SEQ ID NO: 83
Proteína del dominio de IIc: SEQ ID NO: 84	
Secuencia de ADN que codifica el dominio IIIa: SEQ ID NO: 85	CDS de IIIa que codifica la proteína de SEQ ID NO: 86
Proteína del dominio IIIa: SEQ ID NO: 87	
Secuencia de ADN que codifica el dominio IIIb: SEQ ID NO: 88	CDS de IIIb que codifica la proteína de SEQ ID NO: 89
Proteína del dominio IIIb: SEQ ID NO: 90	
Secuencia de ADN que codifica el dominio IV: SEQ ID NO: 91	CDS de IV que codifica la proteína de SEQ ID NO: 92
Proteína del dominio IV: SEQ ID NO: 93	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 94	CDS de IV* que codifica la proteína de SEQ ID NO: 95
Proteína del dominio de IV*: SEQ ID NO: 96	

<i>Aethrobacter sp.</i> SBSr008	
Secuencia de ADN que codifica todo el agrupamiento implicado en la biosíntesis de AGPI: SEQ ID NO: 97	nt 1..1.608 de SEQ ID NO: 97 son CDS de pfa1 que codifica la proteína: SEQ ID NO: 98
	nt 1.644..8.342 de SEQ ID NO: 97 son CDS de pfa2 que codifica la proteína: SEQ ID NO: 99
	nt 18.339..16.243 de SEQ ID NO: 97 son CDS de pfa3

<i>Aetherobacter sp.</i> SBSr008	
	que codifica la proteína: SEQ ID NO: 100
Secuencia de ADN que codifica pfa1: SEQ ID NO: 101	CDS de pfa1 que codifica la proteína de SEQ ID NO: 102
Secuencia de ADN que codifica pfa2: SEQ ID NO: 103	CDS de pfa2 que codifica la proteína de SEQ ID NO: 104
Secuencia de ADN que codifica pfa3 de SEQ ID NO: 105	CDS de pfa3 que codifica la proteína de SEQ ID NO: 106
Secuencia de ADN que codifica el dominio I: SEQ ID NO: 107	CDS de I que codifica la proteína de SEQ ID NO: 108
Proteína del dominio I: SEQ ID NO: 109	
Secuencia de ADN que codifica el dominio IIa: SEQ ID NO: 110	CDS de IIa que codifica la proteína de SEQ ID NO: 111
Proteína del dominio IIa: SEQ ID NO: 112	
Secuencia de ADN que codifica el dominio IIb: SEQ ID NO: 113	CDS de IIb que codifica la proteína de SEQ ID NO: 114
Proteína del dominio IIb: SEQ ID NO: 115	
Secuencia de ADN que codifica el dominio IIc: SEQ ID NO: 116	CDS de IIc que codifica la proteína de SEQ ID NO: 117
Proteína del dominio IIc: SEQ ID NO: 118	
Secuencia de ADN que codifica el dominio IIIa: SEQ ID NO: 119	CDS de IIIa que codifica la proteína de SEQ ID NO: 120
Proteína del dominio IIIa: SEQ ID NO: 121	
Secuencia de ADN que codifica el dominio IIIb: SEQ ID NO: 122	CDS de IIIb que codifica la proteína de SEQ ID NO: 123
Proteína del dominio IIIb: SEQ ID NO: 124	
Secuencia de ADN que codifica el dominio IV: SEQ ID NO: 125	CDS de IV que codifica la proteína de SEQ ID NO: 126
Proteína del dominio IV: SEQ ID NO: 127	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 128	CDS de IV* que codifica la proteína de SEQ ID NO: 129
Proteína del dominio IV*: SEQ ID NO: 130	

<i>Aetherobacter rufus</i> SBSr003	
Secuencia de ADN que codifica el dominio I: SEQ ID NO: 131	CDS de I que codifica la proteína de SEQ ID NO: 132
Proteína del dominio I: SEQ ID NO: 133	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 134	CDS de IV* que codifica la proteína de SEQ ID NO: 135
Proteína del dominio IV*: SEQ ID NO: 136	

<i>Minicystis rosea</i> SBNa008	
Secuencia de ADN que codifica todo el agrupamiento implicado en la biosíntesis de AGPI: SEQ ID NO: 137	nt 1..1.632 de SEQ ID NO: 137 son CDS de pfa1 que codifica la proteína: SEQ ID NO: 138
	nt 1.659..8.327 de SEQ ID NO: 137 son CDS de pfa2 que codifica la proteína: SEQ ID NO: 139
	nt 18.324..17.191 de SEQ ID NO: 137 son CDS de pfa3 que codifica la proteína: SEQ ID NO: 140
Secuencia de ADN que codifica pfa1: SEQ ID NO: 141	CDS de pfa1 que codifica la proteína de SEQ ID NO: 142
Secuencia de ADN que codifica pfa2: SEQ ID NO: 143	CDS de pfa2 que codifica la proteína de SEQ ID NO: 144
Secuencia de ADN que codifica pfa3 de SEQ ID NO: 145	CDS de pfa3 que codifica la proteína de SEQ ID NO: 146
Secuencia de ADN que codifica el dominio I: SEQ ID NO: 147	CDS de I que codifica la proteína de SEQ ID NO: 148
Proteína del dominio I: SEQ ID NO: 149	
Secuencia de ADN que codifica el dominio IIa: SEQ ID NO: 150	CDS de IIa que codifica la proteína de SEQ ID NO: 151
Proteína del dominio IIa: SEQ ID NO: 152	
Secuencia de ADN que codifica el dominio IIb: SEQ ID NO: 153	CDS de IIb que codifica la proteína de SEQ ID NO: 154
Proteína del dominio IIb: SEQ ID NO: 155	
Secuencia de ADN que codifica el dominio IIc: SEQ ID NO: 157	CDS de IIc que codifica la proteína de SEQ ID NO: 157

<i>Minicystis rosea</i> SBNa008	
NO: 156	
Proteína del dominio IIc: SEQ ID NO: 158	
Secuencia de ADN que codifica el dominio IIIa: SEQ ID NO: 159	CDS de IIIa que codifica la proteína de SEQ ID NO: 160
Proteína del dominio IIIa: SEQ ID NO: 161	
Secuencia de ADN que codifica el dominio IIIb: SEQ ID NO: 162	CDS de IIIb que codifica la proteína de SEQ ID NO: 163
Proteína del dominio IIIb: SEQ ID NO: 164	
Secuencia de ADN que codifica el dominio IV: SEQ ID NO: 165	CDS de IV que codifica la proteína de SEQ ID NO: 166
Proteína del dominio IV: SEQ ID NO: 167	
Secuencia de ADN que codifica el dominio IV*: SEQ ID NO: 168	CDS de IV* que codifica la proteína de SEQ ID NO: 169
Proteína del dominio IV*: SEQ ID NO: 170	

SEQ ID NO: 1 es la secuencia de ADN de todo el agrupamiento de genes *pfa* de *Sorangium cellulosum* Soce56 que comprende 17.176 pb. El gen *pfa1* se ubica desde la posición 1 a la posición 1.650, el gen *pfa2* se ubica desde la posición 1.677 a la posición 9.389, y el gen *pfa3* se ubica desde la posición 9.386 a la posición 17.176 y se solapa con *pfa2*. Las secuencias de proteínas codificadas por los genes *pfa1*, *pfa2* y *pfa3*, respectivamente, también se muestran. SEQ ID NO: 2 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa1*, SEQ ID NO: 3 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa2*, SEQ ID NO: 4 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa3*. El gen *pfa1* se muestra por separado como SEQ ID NO: 5 con la secuencia de aminoácidos codificada como SEQ ID NO: 6. El gen *pfa2* se muestra por separado como SEQ ID NO: 7 con la secuencia de aminoácidos codificada como SEQ ID NO: 8. El gen *pfa3* se muestra por separado como SEQ ID NO: 9 con la secuencia de aminoácidos codificada como SEQ ID NO: 10.

SEQ ID NO: 32 es la secuencia de ADN de *Sorangium cellulosum* Soce377 que comprende 17.221 pb, en la que el gen *pfa1* se ubica desde la posición 1 a la posición 1.650, el gen *pfa2* se ubica desde la posición 1.677 a la posición 9.380, y el gen *pfa3* se ubica desde la posición 9.377 a la posición 17.221 y se solapa con *pfa2*. Las secuencias de proteínas codificadas por los genes *pfa1*, *pfa2* y *pfa3*, respectivamente, también se muestran. SEQ ID NO: 33 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa1*, SEQ ID NO: 34 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa2*, SEQ ID NO: 35 es la secuencia de proteínas de SEQ ID NO: 1 codificada por el gen *pfa3*. El gen *pfa1* se muestra por separado como SEQ ID NO: 36 con la secuencia de aminoácidos codificada como SEQ ID NO: 37. El gen *pfa2* se muestra por separado como SEQ ID NO: 38 con la secuencia de aminoácidos codificada como SEQ ID NO: 39. El gen *pfa3* se muestra por separado como SEQ ID NO: 40 con la secuencia de aminoácidos codificada como SEQ ID NO: 41.

SEQ ID NO: 63 es la secuencia de ADN de todo el agrupamiento de genes *pfa* (16.219 pb) procedente de *Aetherobacter fasciculatus* SBSr002 (DSM 21835). El gen *pfa1* se ubica desde la posición 1 a la posición 1.608, el gen *pfa2* se ubica desde la posición 1.644 a la posición 8.312, y el gen *pfa3* se ubica desde la posición 8.309 a la posición 16.219, superpuesto con *pfa2*. Las secuencias de proteínas codificadas por los genes *pfa1*, *pfa2* y *pfa3*, respectivamente, también se muestran. SEQ ID NO: 64 es la secuencia de proteínas de SEQ ID NO: 63 codificada por el gen *pfa1*, SEQ ID NO: 65 es la secuencia de proteínas de SEQ ID NO: 63 codificada por el gen *pfa2*, SEQ ID NO: 66 es la secuencia de proteínas de SEQ ID NO: 63 codificada por el gen *pfa3*. El gen *pfa1* se muestra por separado como SEQ ID NO: 67 con la secuencia de aminoácidos codificada como SEQ ID NO: 68. El gen *pfa2* se muestra por separado como SEQ ID NO: 69 con la secuencia de aminoácidos codificada como SEQ ID NO: 70. El gen *pfa3* se muestra por separado como SEQ ID NO: 71 con la secuencia de aminoácidos codificada como SEQ ID NO: 72.

SEQ ID NO: 97 es la secuencia de ADN de todo el agrupamiento de genes *pfa* (16.243 pb) procedente de *Aetherobacter* SBSr008. El gen *pfa1* se ubica desde la posición 1 a la posición 1.608, el gen *pfa2* se ubica desde la posición 1.644 a la posición 8.342, y el gen *pfa3* se ubica desde la posición 8.339 a la posición 16.243, superpuesto con *pfa2*. Las secuencias de proteínas codificadas por los genes *pfa1*, *pfa2* y *pfa3*, respectivamente, también se muestran. SEQ ID NO: 98 es la secuencia de proteínas de SEQ ID NO: 97 codificada por el gen *pfa1*, SEQ ID NO: 99 es la secuencia de proteínas de SEQ ID NO: 97 codificada por el gen *pfa2*, SEQ ID NO: 100 es la secuencia de proteínas de SEQ ID NO: 97 codificada por el gen *pfa3*. El gen *pfa1* se muestra por separado como SEQ ID NO: 101 con la secuencia codificada de aminoácidos como SEQ ID NO: 102. El gen *pfa2* se muestra por separado como SEQ ID NO: 103 con la secuencia de aminoácidos codificada como SEQ ID NO: 104. El gen *pfa3* se muestra por separado como SEQ ID NO: 105 con la secuencia de aminoácidos codificada como SEQ ID NO: 106.

SEQ ID NO: 137 es la secuencia de ADN de todo el agrupamiento de genes *pfa* (17.191 pb) procedente de *Aetherobacter* SBNa008. El gen *pfa1* se ubica desde la posición 1 a la posición 1.632, el gen *pfa2* se ubica desde la posición 1.659 a la posición 8.327, y el gen *pfa3* se ubica desde la posición 8.324 a la posición 17.191, superpuesto con *pfa2*. Las secuencias de proteínas codificadas por los genes *pfa1*, *pfa2* y *pfa3*, respectivamente, también se

muestran. SEQ ID NO: 138 es la secuencia de proteínas de SEQ ID NO: 137 codificada por el gen pfa1, SEQ ID NO: 139 es la secuencia de proteínas de SEQ ID NO: 137 codificada por el gen pfa2, SEQ ID NO: 140 es la secuencia de proteínas de SEQ ID NO: 137 codificada por el gen pfa3. El gen pfa1 se muestra por separado como SEQ ID NO: 141 con la secuencia codificada de aminoácidos como SEQ ID NO: 142. El gen pfa2 se muestra por separado como SEQ ID NO: 143 con la secuencia de aminoácidos codificada como SEQ ID NO: 144. El gen pfa3 se muestra por separado como SEQ ID NO: 145 con la secuencia de aminoácidos codificada como SEQ ID NO: 146.

Para *Aetherobacter* SBSr003, se dan las siguientes regiones del agrupamiento implicado en la síntesis de AGPI: gen I como SEQ ID NO: 131 con la proteína codificada como SEQ ID NO: 132 y por separado como SEQ ID NO: 133, dominio IV* como SEQ ID NO: 134 con la proteína codificada como SEQ ID NO: 135 y por separado como SEQ ID NO: 136. Como SBSr003 se deposita bajo el Tratado de Budapest como DSM 23122 (DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Inhoffenstr. 7b, 38124 Braunschweig, Alemania, el 7 de octubre de 2007), el agrupamiento de genes sintéticos de AGPI de esta bacteria se abarca para los fines de la invención. Este agrupamiento de genes contiene las secuencias codificantes dadas para SBSr003, y preferentemente se extiende desde su gen I a su dominio IV*. Preferentemente, este agrupamiento de genes de DSM 23122 codifica una secuencia de aminoácidos que tiene al menos una identidad de 60 %, al menos 85 %, al menos 90 %, y preferentemente al menos 95 % o al menos 98 % con respecto a una o más de las secuencias de aminoácidos de los genes sintéticos de AGPI de *Aetherobacter*, especialmente con respecto a la secuencia de aminoácidos codificada por SEQ ID NO: 63 (secuencias de aminoácidos SEQ ID NO: 64 y 68, SEQ ID NO: 65 y 70, e incluyendo la SEQ ID NO: 64 y 72), por SEQ ID NO: 97 (secuencias de aminoácidos SEQ ID NO: 98 y 102, SEQ ID NO: 99 y 104, e incluyendo la SEQ ID NO: 100 y 106) y/o con respecto a la secuencia de aminoácidos codificada por la SEQ ID NO: 137 (secuencias de aminoácidos SEQ ID NO: 138 y 142, SEQ ID NO: 139 y 144, e incluyendo la SEQ ID NO: 140 y 146).

25 Descripción detallada

Los inventores han identificado agrupamientos de genes biosintéticos de AGPI mixobacterianos a partir de la cepa de *Sorangium cellulosum* 56 (Soce56) y la cepa 377 (Soce377), así como de *Aetherobacter fasciculatus* SBSr002 DSM 21835, *Aetherobacter* sp. SBSr008, *Minicystis rosea* SBNa008, y en parte de *Aetherobacter rufus* SBSr003. El análisis de los loci biosintéticos de estos organismos fuente muestra que la disposición de los genes, así como la organización de los dominios catalíticos en las proteínas biosintéticas son altamente homólogas pero difieren significativamente de las secuencias publicadas de los agrupamientos de genes de AGPI a partir de varios organismos (véase la Figura 1).

Un primer aspecto de la invención se refiere a un proceso de producción de uno o más ácidos grasos poliinsaturados por medio de la expresión génica heteróloga, comprendiendo el proceso las etapas que consisten en:

(i) proporcionar un organismo de producción que comprenda un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados, siendo derivado dicho agrupamiento de genes de un organismo fuente que difiere del organismo de producción; y que abarca una subsecuencia ER que codifica una enoil reductasa y una subsecuencia AT que codifica una aciltransferasa, en la que la subsecuencia AT se ubica aguas abajo con respecto a la subsecuencia ER;

(ii) cultivar el organismo de producción de la etapa (i) en presencia de una fuente de carbono fermentable, conforme la cual se producen uno o más ácidos grasos poliinsaturados; y

(iii) opcionalmente, recuperar uno o más ácidos grasos poliinsaturados.

El proceso de acuerdo con la invención es útil para la producción de ácidos grasos poliinsaturados (AGPI). La expresión "ácido graso poliinsaturado" se utiliza en la presente memoria en el sentido generalmente conocido por el experto. Preferentemente, los ácidos grasos poliinsaturados son ácidos monocarboxílicos que comprenden de 12 a 30 átomos de carbono, más preferentemente 18 a 24 átomos de carbono, particularmente de manera preferente 20 o 22 átomos de carbono. Preferentemente, la expresión "ácido graso poliinsaturado" se refiere a una cadena hidrocarbonada larga compuesta de 18 o más átomos de carbono que tienen al menos 2, preferentemente al menos 3 o al menos 4 dobles enlaces y un grupo carboxilato terminal. Preferentemente, los ácidos grasos poliinsaturados comprenden al menos 2, más preferentemente al menos 3, aún más preferentemente al menos 4, al menos 5 o al menos 6 grupos etilénicamente insaturados (dobles enlaces C=C), que preferentemente no están conjugados entre sí. Para los fines de la invención, la expresión "ácido graso poliinsaturado" también abarca derivados de ácidos carboxílicos poliinsaturados tales como los derivados hidroxilo y/o las sales. Como se conoce en el campo, un ácido graso se puede esterificar para formar triglicéridos y/o fosfolípidos, así como esfingolípidos. De este modo, en una realización, la presente invención también se refiere a tales productos esterificados. Además, el producto de ácidos grasos de la presente invención puede ser libre de ácidos grasos. Los ácidos grasos libres tienen un grupo carboxilo libre, no están conectados químicamente a cualquier otro compuesto incluyendo triacilglicéridos, fosfolípidos o esfingolípidos, y pueden estar presentes libremente en cualquier compartimiento de la célula.

En una realización preferente, al menos 1, más preferentemente al menos 2, aún más preferentemente al menos 3 y lo más preferentemente todos los grupos etilénicamente insaturados tienen una configuración *cis* [(Z)-configuración].

Los ácidos grasos poliinsaturados típicos son ácido oleico, ácido linoleico, ácido alfa-linolénico, ácido gamma-linolénico, ácido dihomo-gamma-linolénico, ácido araquidónico, ácido 5,8,11,14,17-eicosapentaenoico (AEP), ácido docosatetraenoico, ácido estearidónico, ácido eicosatetraenoico (ETE), ácido 7,10,13,16,19-docosapentaenoico y ácido 4,7,10,13,16,19-docosahexaenoico (ADH). Los ácidos grasos poliinsaturados preferentes se seleccionan entre el grupo que consiste en (18:3), (18:4), (18:5), (20:3), (20:4), (20:5), (22:3), (22:4), (22:5), y (22:6).

En una realización particularmente preferente, los ácidos grasos poliinsaturados son ácidos grasos poliinsaturados ω -3. Los ácidos grasos poliinsaturados preferentes se seleccionan entre el grupo que consiste en (18:3 ω -3), (18:4 ω -3), (18:5 ω -3), (20:3 ω -3), (20:4 ω -3), (20:5 ω -3), (22:3 ω -3), (22:4 ω -3), (22:5 ω -3), y (22:6 ω -3). Ejemplos preferentes de ácidos grasos poliinsaturados ω -3 incluyen ácido α -linolénico (18:3 AAL), ácido eicosa-cis-5,8,11,14,17-pentaenoico (20:5 ω -3 AEP) y ácido docosa-cis-4,7,10,13,16,19-hexaenoico (22:6 ω -3 ADH). En una realización preferente adicional, los ácidos grasos insaturados omega-3 se seleccionan entre ácido eicosapentaenoico (AEP), ácido docosahexaenoico (ADH) y mezclas de los mismos. Estos compuestos también se conocen con sus nombres triviales ácido timnodónico para AEP y ácido cervónico para ADH, respectivamente.

En una realización particularmente preferente, uno o más ácidos grasos poliinsaturados se seleccionan entre el grupo que consiste en ácido hexadecadienoico (16:2), ácido hexadecenoico (16:1), ácido hexadecanoico (16:0), ácido heptadecenoico (17:1 ω 7), ácido iso-heptadienoico (iso-17:2 ω 5, 11), ácido iso-heptadecenoico (iso-17:1 ω 5), ácido iso-heptadecanoico (iso-17:0), ácido iso-3-hidroxi-pentadecanoico (iso-3-OH-15:0), ácido γ linolénico (18:3 ω 6), ácido linoleico (18:2 ω 6), ácido octadecenoico (18:1), ácido octadecanoico (18:0), ácido 2-hidroxi-heptadecanoico (2-OH-17:0), ácido 2-hidroxi-heptadecenoico (2-OH-17:1), ácido hidroxí-hexadecanoico (2-OH-16:0, 3-OH-16:0), ácido iso-pentadecanoico (iso-15:0) y ácido eicosadienoico (20:2 ω 6). De acuerdo con el proceso de la invención, se producen uno o más ácidos grasos poliinsaturados. Así, el proceso de acuerdo con la invención produce normalmente una composición a partir de la cual uno o más ácidos grasos poliinsaturados pueden recuperarse y aislarse, respectivamente, como se describe en la etapa opcional (iii).

Preferentemente, el proceso de acuerdo con la invención produce una composición que comprende más de un ácido graso poliinsaturado, p. ej., una composición de al menos 2 o 3 AGPI, en la que un ácido graso poliinsaturado particular es el ingrediente principal.

En una realización preferente, el contenido de un ácido graso poliinsaturado particular asciende a al menos 40 % en peso, más preferentemente al menos 45 % en peso, aún más preferentemente al menos 50 % en peso, aún más preferentemente al menos 55 % en peso, incluso más preferentemente al menos 60 % en peso, lo más preferentemente al menos 65 % en peso y en particular al menos 70 % en peso, basado en el peso total de todos los ácidos grasos poliinsaturados contenidos en la composición. En otra realización preferente, el contenido de dicho ácido graso poliinsaturado particular asciende a al menos 75 % en peso, más preferentemente al menos 80 % en peso, aún más preferentemente al menos 82 % en peso, aún más preferentemente al menos 84 % en peso, incluso más preferentemente al menos 86 % en peso, más preferentemente al menos 88 % en peso y en particular al menos 90 % en peso, basado en el peso total de todos los ácidos grasos poliinsaturados contenidos en la composición.

La invención se refiere a un proceso de producción de uno o más ácidos grasos poliinsaturados por medio de la expresión génica heteróloga. Un experto conoce la expresión "expresión génica heteróloga". Preferentemente, la expresión génica heteróloga se define como la síntesis de proteínas exógenas en un organismo huésped (organismo de producción) después de la transformación de ese organismo por genes portadores de vectores de un organismo diferente (organismo fuente). Por "expresión génica" se entiende la producción de un polipéptido funcional a través de la transcripción de un segmento de ácido nucleico en ARNm y la traducción del ARNm en una proteína. Por "expresión génica heteróloga" se entiende generalmente que un ácido nucleico, no presente de forma natural en el genoma del organismo de producción, está presente en el organismo de producción y se vincula operativamente a las secuencias de ácidos nucleicos promotoras y terminadoras de una manera que se expresa en el organismo de producción. La expresión "vinculado operativamente" se refiere a la asociación de un gen con una secuencia que controla su expresión en un fragmento de ácido nucleico único. Por ejemplo, un promotor se vincula operativamente con una secuencia codificante cuando es capaz de regular la expresión de esa secuencia codificante. Asimismo, en el presente contexto, la expresión génica heteróloga se refiere además a la presencia de un ácido nucleico con una función similar a un ácido nucleico presente de forma natural, en la que la expresión de dicho producto de ácido nucleico heterólogo cambia la composición de ácidos grasos. Por expresión heteróloga de un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados se entiende que varios genes se expresan de forma heteróloga, cuyos productos génicos constituyen etapas en una vía, no presentes de forma natural en el organismo de producción.

De acuerdo con la etapa (i) del proceso de acuerdo con la invención, se proporciona un organismo de producción que comprende un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados. Para los fines de la memoria descriptiva, la expresión "agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados" se refiere a cualquier agrupamiento de genes que codifica una vía biosintética que es capaz de sintetizar uno o más ácidos grasos poliinsaturados por medio del organismo de producción. Por expresión heteróloga del agrupamiento de genes, varios genes se expresan de forma heteróloga, cuyos productos génicos constituyen etapas en una vía, no presentes de forma natural en el organismo de producción. El agrupamiento de

genes puede comprender uno o más genes, cada uno de los cuales puede codificar de forma independiente una proteína y/o enzima mono- o polifuncional. Para los fines de la memoria descriptiva, un gen se considera generalmente como una unidad de la herencia en un organismo vivo. Normalmente es un tramo de ADN que codifica un tipo de proteína o una cadena de ARN que tiene una función en el organismo. Preferentemente, un gen se define como una región localizable de la secuencia genómica, que corresponde a una unidad de herencia, que se asocia con regiones reguladoras, regiones transcritas, y/u otras regiones secuenciales funcionales.

Preferentemente, el agrupamiento de genes abarca al menos dos genes, más preferentemente al menos 3 genes. Preferentemente, el agrupamiento de genes abarca como máximo cuatro genes, más preferentemente como máximo 3 genes. En una realización particularmente preferente, el agrupamiento de genes comprende exactamente tres genes, es decir, consiste en tres genes y, opcionalmente, material genético entre dichos tres genes.

De acuerdo con la invención, el agrupamiento de genes se deriva de un organismo fuente diferente del organismo de producción. En este contexto, "derivado" significa preferentemente que el agrupamiento de genes procede de un organismo de origen natural (tipo natural) o de un organismo modificado genéticamente como organismo fuente. En este sentido, "originar" no significa necesariamente que el agrupamiento de genes haya sido biosintetizado por el organismo fuente, sino que también debe incluir métodos de amplificación biotecnológicos, tales como PCR y síntesis *de novo* a partir de componentes adecuados, tales como fosfoamiditas.

Cuando el agrupamiento de genes se origina a partir de un organismo genéticamente modificado, como organismo fuente, el genoma de dicho organismo genéticamente modificado tiene preferentemente una homología con su correspondiente organismo de origen natural (tipo natural) de al menos 90 %, más preferentemente al menos 92 %, aún más preferentemente al menos 94 %, aún más preferentemente al menos 96 %, incluso más preferentemente al menos 97 %, lo más preferentemente al menos 98 % y en particular al menos 99 %. Para los fines de la invención, el grado de "homología" se basa preferentemente en una comparación de dos secuencias utilizando el algoritmo de BLASTn 2.2.22 de la url <http://blast.ncbi.nlm.nih.gov/Blast.cgi> que se describe adicionalmente en Zhang Z *et al.* (2000) *J Comput Biol* 7: 203-214.

Preferentemente, el número de pares de bases (pb) que forma el agrupamiento de genes de acuerdo con la invención está dentro del intervalo de 5.000 a 30.000 pb, más preferentemente 7.000 a 28.000 pb, aún más preferentemente 9.000 a 26.000 pb, aún más preferentemente 11.000 a 24.000 pb, lo más preferentemente 13.000 a 22.000 pb y en particular 15.000 a 20.000 pb, o 16.200 a 17.300 pb.

De acuerdo con la invención, el agrupamiento de genes abarca una subsecuencia ER que codifica una enoil reductasa y una subsecuencia AT que codifica una aciltransferasa, en la que la subsecuencia AT se ubica aguas abajo con respecto a la subsecuencia ER. En este sentido "subsecuencia" del agrupamiento de genes significa una secuencia de ADN ubicada en el agrupamiento de genes y dentro de un marco de lectura, cuya subsecuencia es más corta que la secuencia de ADN general del agrupamiento de genes. En este sentido "aguas abajo" significa preferentemente más cerca del extremo 3' de la secuencia de ADN. De este modo, la subsecuencia AT se ubica más cerca del extremo 3' que la subsecuencia ER. Un experto sabe que los péptidos o proteínas pueden calificarse como enoil reductasas y aciltransferasas, respectivamente, es decir, cuyos requisitos deben satisfacerse en relación con la actividad biocatalítica.

Generalmente, una enoil reductasa es un tipo de enzima que actúa sobre los grupos enoil. Las enoil reductasas se clasifican preferentemente en la categoría EC 1.3 (oxidorreductasas que actúan sobre el grupo CH-CH de los donantes), subcategoría EC 1.3.1 (con NAD o NADP como aceptor).

Generalmente, una aciltransferasa es un tipo de enzima transferasa que actúa sobre grupos acilo. Las aciltransferasas se clasifican preferentemente en la categoría EC 2.3 y comprenden las subcategorías EC 2.3.1 (aciltransferasas que transfieren grupos distintos de los grupos amino-acilo), EC 2.3.2 (aminoaciltransferasas), y EC 2.3.3 (aciltransferasas que convierten grupos acilo en grupos alquilo tras la transferencia) desde las cuales la aciltransferasa de acuerdo con la invención se seleccionan preferentemente.

De acuerdo con la invención, la subsecuencia ER y/o subsecuencia AT no necesitan codificar proteínas y enzimas independientes separadas, respectivamente. Más bien, es posible que la subsecuencia ER y/o subsecuencia AT codifiquen dominios de una proteína/enzima multifuncional que comprende la actividad biocatalítica respectiva entre otras actividades biocatalíticas. Así, para los fines de la memoria descriptiva, la expresión "que codifica un(a)..." es sinónimo de "que codifica un péptido o proteína que tiene la actividad o funcionalidad de un(a)...., opcionalmente entre otras actividades o funcionalidades".

Preferentemente, el agrupamiento de genes comprende un primer gen, un segundo gen y un tercer gen, en el que el segundo gen se ubica aguas abajo con respecto al primer gen; y el tercer gen se ubica aguas abajo con respecto al segundo gen. En una realización preferente, la subsecuencia ER se ubica en el primer gen y/o la subsecuencia AT, que preferentemente es AGPAT, se ubica en el tercer gen.

En una realización preferente de la invención, la subsecuencia ER codifica una proteína/enzima monofuncional que

no comprende dominios adicionales que exhiban otras actividades biocatalíticas. De acuerdo con la presente realización preferente, la subsecuencia ER se ubica preferentemente en el primer gen, para los fines de la memoria descriptiva, preferentemente, también se refiere como "gen pfa1". De este modo, la expresión heteróloga del gen pfa1 en el organismo de producción produce preferentemente una proteína/enzima preferentemente monofuncional que exhibe actividad enoil reductasa.

En una realización preferente de la invención, la subsecuencia AT codifica un dominio de una proteína/enzima polifuncional que exhibe actividad aciltransferasa en el dominio codificado por la subsecuencia AT y exhibe adicionalmente actividades biocatalíticas adicionales en otros dominios codificados por más subsecuencias. De acuerdo con la presente realización preferente, la subsecuencia AT se ubica preferentemente en el tercer gen, para los fines de la memoria descriptiva, preferentemente, también referido como "gen pfa3". De este modo, la expresión heteróloga del gen pfa3 en el organismo de producción produce preferentemente una proteína/enzima preferentemente polifuncional que exhibe actividad aciltransferasa en el dominio codificado por la subsecuencia AT, así como actividades biocatalíticas adicionales en otros dominios que, a su vez, se codifican por otras subsecuencias que también se ubican en el gen pfa3.

Preferentemente, la enoil reductasa codificada por la subsecuencia ER tiene una identidad proteica con SEQ ID NO: 2 y/o con SEQ ID NO: 6 de al menos 46 %, más preferentemente al menos 48 %, aún más preferentemente al menos 50 %, aún más preferentemente al menos 52 %, incluso más preferentemente al menos 54 %, lo más preferentemente al menos 56 %, y en particular al menos 58 %.

Preferentemente, la aciltransferasa codificada por la subsecuencia AGPAT tiene una identidad proteica con SEQ ID NO: 30 y/o con SEQ ID NO: 31 de al menos 2,5 %, más preferentemente al menos 5,0 %, más preferentemente al menos 7,5 %, aún más preferentemente al menos 10 %, aún más preferentemente al menos 12,5 %, incluso más preferentemente al menos 15 %, lo más preferentemente al menos 17,5 %, y en particular al menos 20 %, una identidad proteica de al menos preferentemente 22 %, p. ej., como se determina por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica en la presente memoria.

Preferentemente, la enoil reductasa codificada por la subsecuencia ER tiene una similitud proteica con SEQ ID NO: 2 y/o con SEQ ID NO: 6 de al menos 66 %, más preferentemente al menos 67 %, aún más preferentemente al menos 68 %, aún más preferentemente al menos 69 %, incluso más preferentemente al menos 70 %, lo más preferentemente al menos 71 %, y en particular al menos 72 %.

Preferentemente, la aciltransferasa codificada por la subsecuencia AGPAT tiene una similitud proteica con SEQ ID NO: 66, SEQ ID NO: 72, SEQ ID NO: 31, SEQ ID NO: 62, SEQ ID NO: 96, SEQ ID NO: 130, SEQ ID NO: 136 y/o con SEQ ID NO: 170 de al menos 2,5 %, más preferentemente al menos 5,0 %, más preferentemente al menos 7,5 %, aún más preferentemente al menos 10 %, aún más preferentemente al menos 15 %, incluso más preferentemente al menos 20 %, preferentemente al menos 22 %, lo más preferentemente al menos 25 %, y en particular al menos 27,5 %, determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica en la presente memoria.

Preferentemente, la subsecuencia ER tiene una identidad por pares de ADN con SEQ ID NO: 67 y/o con SEQ ID NO: 63 nt 1-1.608 de al menos 60 %, más preferentemente al menos 62 %, aún más preferentemente al menos 64 %, aún más preferentemente al menos 66 %, incluso más preferentemente al menos 67 %, lo más preferentemente al menos 68 %, y en particular al menos 69 %.

Preferentemente, la subsecuencia AGPAT tiene una identidad por pares de ADN con SEQ ID NO: 63, nt 8.309-16.219 y SEQ ID NO: 71 y/o con SEQ ID NO: 29, SEQ ID NO: 60, SEQ ID NO: 94, SEQ ID NO: 128, SEQ ID NO: 136 y/o SEQ ID NO: 170 de al menos 46 %, más preferentemente al menos 48 %, más preferentemente al menos 50 %, aún más preferentemente al menos 52 %, aún más preferentemente al menos 54 %, incluso más preferentemente al menos 56 %, más preferentemente al menos 58 %, y en particular al menos 60 %, determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Clustal W, como se especifica en la presente memoria. Para los fines de la memoria descriptiva, identidad proteica es sinónimo de identidad de secuencia de proteínas, identidad por pares de ADN es sinónimo de identidad por pares de la secuencia de ADN y similitud proteica es sinónimo de similitud de la secuencia de proteínas.

La identidad proteica, similitud proteica e identidad por pares de ADN (alineamiento) se determinan en general preferentemente por medio del programa Geneious (<http://www.geneious.com/>) utilizando una de las siguientes herramientas de alineamiento: Geneious Alignment (Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2009) Geneious Pro v5.4.2, disponible en <http://www.geneious.com/>). La identidad por pares de ADN puede determinarse alternativamente utilizando la herramienta Clustal W2.0.11 Alignment (Thompson, *Nucleic Acids Res.* 1994, 4673-4680). Los ajustes de parámetros para la comparación de secuencias proteicas son como indica preferentemente ClustalW Alignment: matriz de coste: BLOSUM, coste de apertura de huecos: 10, coste de extensión de huecos: 0,1. Los ajustes de parámetros para la comparación de secuencias de ADN son como se indica preferentemente (A) Geneious Alignment: matriz de coste: 51 % de similitud (5,0/-3,0), penalización de apertura de huecos: 12, penalización de extensión de huecos: 3, tipo de alineamiento:

alineamiento global con huecos de extremo libres, y (B) Clustal W2.0.11 Alignment: matriz de coste: IUB, coste de apertura de huecos: 15, coste de extensión de huecos: 6,66.

5 En una realización preferente, el agrupamiento de genes de acuerdo con la invención abarca además una subsecuencia CS que codifica una cetosintasa (en lo sucesivo también referida como cetosintasa cs1) y/o al menos una subsecuencia DH que codifica una deshidratasa (en lo sucesivo también referida como deshidratasa dh1). En una realización preferente, el agrupamiento de genes de acuerdo con la invención abarca además una subsecuencia CS que codifica una cetosintasa (en lo sucesivo también referida como cetosintasa cs1), una subsecuencia FLC que codifica un factor de longitud de cadena y/o al menos una subsecuencia DH que codifica una deshidratasa (en lo sucesivo también referida como deshidratasa dh1). Un experto sabe que los péptidos o proteínas pueden calificarse como cetosintasas, factores de longitud de cadena, y deshidratasas, respectivamente, es decir, cuyos requerimientos deben satisfacerse en relación con la actividad biocatalítica. Generalmente, una beta-cetoacil sintasa es un tipo de enzima sintasa que sintetiza la cadena de carbono en un tipo de condensación de Claisen. Las beta-cetoacilintasas se clasifican preferentemente en la categoría EC 2.3.1.41. Generalmente, un factor de longitud de cadena controla la longitud de cadena del ácido graso a sintetizar. El factor de longitud de cadena también desempeña un papel destacado en la síntesis de policétido (cf. p. ej., Y Tang *et al.*, *J Am Chem Soc.* 2003, 125(42), 12708-9).

20 Generalmente, una deshidratasa es un tipo de enzima que cataliza la eliminación de oxígeno e hidrógeno a partir de compuestos orgánicos en forma de agua (deshidratación). Existen cuatro clases de deshidratasas: deshidratasas que actúan en ésteres de 3-hidroxiacil-CoA reductasa (sin cofactores); deshidratasas que actúan en ésteres de 2-hidroxiacil-CoA reductasa (reacción radical, grupo que contiene [4Fe-4S]); deshidratasas que actúan en ésteres 4-hidroxiacil-CoA (que contiene [4Fe-4S] y FAD); y deshidratasas que contienen un agrupamiento [4Fe-4S] como sitio activo (aconitasa, fumarasa, serina deshidratasa), a partir de las cuales se selecciona preferentemente la deshidratasa de acuerdo con la invención. Preferentemente, las deshidratasas se clasifican en EC 4.2.1 (hidrolisas). Los ejemplos preferentes de deshidratasas son hidroxil-acil-deshidratasas, particularmente las β -hidroxil-acil-deshidratasas. En una realización preferente, la deshidratasa se selecciona entre el grupo que consiste en crotonoil-[proteína transportadora de acilos] hidratasa (EC 4.2.1.58), 3-hidroxiocetanoil-[proteína transportadora de acilos] deshidratasa (EC 4.2.1.59), 3-hidroxi-decanoil-[proteína transportadora de acilos] deshidratasa (EC 4.2.1.60), 3-hidroxi-palmitoil-[proteína transportadora de acilos] deshidratasa (EC 4.2.1.61), enoil-CoA hidratasa de cadena larga (EC 4.2.1.74), 3-hidroxi-propionil-CoA deshidratasa (EC 4.2.1.116), y 4-hidroxi-butanil-CoA deshidratasa (EC 4.2.1.120).

35 En una realización preferente, la subsecuencia CS, la subsecuencia FLC, al menos una subsecuencia DH y/o la subsecuencia AT se ubican en el mismo gen, preferentemente en el tercer gen (es decir pfa3), que codifica preferentemente una proteína multifuncional que comprende al menos tres dominios, más preferentemente al menos cuatro dominios y en particular preferentemente al menos cinco dominios.

40 En una realización preferente, el tercer gen codifica una proteína multifuncional que tiene al menos cuatro, preferentemente al menos cinco funcionalidades diferentes, en las que las subsecuencias que codifican las funcionalidades individuales diferentes abarcan la subsecuencia AT que codifica una aciltransferasa y en las que dicha subsecuencia AT se ubica aguas abajo con respecto a las otras subsecuencias.

45 En una realización preferente, el agrupamiento de genes, preferentemente el tercer gen (es decir, pfa3) comprende dos subsecuencias DH que codifican dos deshidratasas, a saber, deshidratasa dh1 y deshidratasa dh2, que pueden ser idénticas o diferentes.

50 Preferentemente, el agrupamiento de genes abarca además una subsecuencia CS que codifica una cetosintasa (en lo sucesivo también referida como cetosintasa cs2), una subsecuencia MAT que codifica una malonil-CoA-transacilasa, al menos una subsecuencia PPA que codifica una proteína transportadora de acilos (en lo sucesivo referida como acp1), una subsecuencia CR que codifica una cetorreductasa y/o una subsecuencia DH que codifica una deshidratasa (en lo sucesivo también referida como deshidratasa dh3). La cetosintasa cs2 puede ser idéntica o diferente de la cetosintasa cs1 anteriormente mencionada.

55 La deshidratasa dh3 puede ser idéntica o diferente de las deshidratasas dh1 y dh2 anteriormente mencionadas, respectivamente.

60 Un experto sabe que los péptidos o proteínas pueden calificarse como malonil-CoA-transacilasa, proteína transportadora de acilos, y cetorreductasas, respectivamente, es decir, que los requerimientos deben satisfacerse en relación con la actividad biocatalítica.

65 Generalmente, una malonil-CoA-transacilasa (también referida como [proteína transportadora de acilos] S-maloniltransferasa) cataliza la transferencia de un grupo malonilo de coenzima A a una proteína transportadora de acilos. Preferentemente, las malonil-CoA-transacilasas se clasifican en la categoría EC 2.3.1.39.

Generalmente, una proteína transportadora de acilos es un pequeño polipéptido con el cual la cadena de carbono/o

acilo creciente se une por medio de un enlace tioéster durante la biosíntesis (TDH Bugg (ed.) en *Introduction to Enzym and Coenzyme Chemistry*, 2ª ed, 2004, Blackwell Publishing).

5 Generalmente, una cetereductasa es una enzima que reduce grupos ceto. Preferentemente, las cetereductasas se clasifican en la categoría EC 1.1.1.184.

10 En una realización preferente, la subsecuencia CS, la subsecuencia MAT, al menos una subsecuencia PPA, la subsecuencia CR y/o la subsecuencia DH se ubican en el mismo gen, preferentemente en el segundo gen (es decir, pfa2).

15 En una realización preferente, el agrupamiento de genes, preferentemente el segundo gen (es decir, pfa2) comprende al menos dos subsecuencias PPA que codifican dos proteínas transportadoras de acilos, a saber, proteínas transportadoras de acilos acp1 y proteína transportadora de acilos acp2, que pueden ser idénticas o diferentes. En una realización preferente, el número total de subsecuencias PPA que codifican las proteínas transportadoras de acilos es 5 o menos (acp1, acp2, acp3, acp4 y acp5). Dichas 5 o menos proteínas transportadoras de acilos pueden ser idénticas o diferentes.

20 En una realización particularmente preferente, el agrupamiento de genes de acuerdo con la invención comprende un primer gen (pfa1), un segundo gen (pfa2) y un tercer gen (pfa3), en el que el segundo gen se ubica opcionalmente aguas abajo con respecto al primer gen y el tercer gen se ubica opcionalmente aguas abajo con respecto al segundo gen; y

25 una subsecuencia ER que codifica un péptido o proteína que exhibe una actividad enoil reductasa, una subsecuencia CS2 que codifica un péptido o proteína que exhibe una actividad cetosintasa (cs2), una subsecuencia MAT que codifica un péptido o proteína que exhibe una actividad malonil-CoA-transacilasa, al menos una subsecuencia PPA que codifica un péptido o proteína que exhibe una funcionalidad de la proteína transportadora de acilos, una subsecuencia CR que codifica un péptido o proteína que exhibe una actividad cetereductasa, una subsecuencia DH3 que codifica un péptido o proteína que exhibe una actividad deshidratasa (dh3), una subsecuencia CS1 que codifica un péptido o proteína que exhibe una actividad cetosintasa (cs1), una subsecuencia FLC que codifica un péptido o proteína que exhibe una funcionalidad del factor de longitud de cadena; dos subsecuencias DH1 y DH2 que codifican un péptido o proteína que exhibe una actividad deshidratasa (dh1 y dh2), y una subsecuencia AT que codifica un péptido o proteína que exhibe una actividad aciltransferasa;

30 en la que la subsecuencia ER se ubica preferentemente en el primer gen (pfa1); las subsecuencias CS2, MAT, PPA, CR y DH3 se ubican en el segundo gen (es decir, pfa2); y las subsecuencias CS1, FLC, DH1, DH2, y AT se ubican en el tercer gen (es decir, pfa3).

35 Preferentemente, las subsecuencias tienen el siguiente orden en el agrupamiento de genes (aguas arriba-aguas abajo, es decir, 5'-3'): ER, CS2, MAT, nPPA, CR, DH3, CS1, FLC, DH1, DH2, AT, en el que n se refiere al número de PPA, preferentemente 1, 2, 3, 4 o 5.

40 En una realización preferente, el agrupamiento de genes de acuerdo con la invención comprende una subsecuencia que codifica una proteína Pfa1 que tiene una identidad proteica con la SEQ ID NO: 2, SEQ ID NO: 6 y/o con SEQ ID NO: 33, SEQ ID NO: 37 (Soce) y/o SEQ ID NO: 64, SEQ ID NO: 68 (*Aetherobacter*) y/o SEQ ID NO: 98, SEQ ID NO: 102 (*Aetherobacter*) y/o SEQ ID NO: 138, SEQ ID NO: 142 (*Aetherobacter*) de al menos 30 %, al menos 32,5 %, al menos 35 %, al menos 37,5 %, al menos 40 %, más preferentemente al menos 42,5 %, al menos 44 %, aún más preferentemente al menos 45 %, al menos 46,5 %, aún más preferentemente al menos 47,5 %, incluso más preferentemente al menos 49 %, al menos 50 %, al menos 51,5 %, lo más preferentemente al menos 52,5 %, y en particular al menos 54 %, al menos 55 %, al menos 56 %, al menos 56,5 % aún más preferentemente al menos 58 % o, al menos 59 %, aún más preferentemente al menos 60 %, incluso más preferentemente al menos 62 %, lo más preferentemente al menos 64 %, y en particular al menos 66 %, determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica en la presente memoria. La proteína Pfa1 puede codificarse por una secuencia de ADN que tiene al menos 54 %, más preferentemente al menos 56 %, aún más preferentemente al menos 58 %, aún más preferentemente al menos 60 %, incluso más preferentemente al menos 62 %, lo más preferentemente al menos 64 %, y en particular al menos 66 % con la SEQ ID NO: 5 y/o con la SEQ ID NO: 36 (Soce) y/o SEQ ID NO: 67 y/o SEQ ID NO: 101 y/o SEQ ID NO: 141 (*Aetherobacter*), determinada preferentemente por medio del software Geneious y la herramienta de alineamiento ClustalW, como se especifica en la presente memoria.

60 En una realización preferente, el agrupamiento de genes de acuerdo con la invención comprende una subsecuencia que codifica una proteína Pfa2 que tiene una identidad proteica con SEQ ID NO: 3, SEQ ID NO: 8 y/o con SEQ ID NO: 34, SEQ ID NO: 39 (Soce) y/o SEQ ID NO: 65, SEQ ID NO: 70 (*Aetherobacter*) y/o SEQ ID NO: 99, SEQ ID NO: 104 (*Aetherobacter*) y/o SEQ ID NO: 139, SEQ ID NO: 144 (*Aetherobacter*) de al menos 30 %, al menos 32,5 %, al menos 35 %, al menos 37,5 %, al menos 40 %, más preferentemente al menos 42,5 %, al menos 44 %, aún más preferentemente al menos 45 %, al menos 46,5 %, aún más preferentemente al menos 47,5 %, incluso más preferentemente al menos 49 %, al menos 50 %, al menos 51,5 %, lo más preferentemente al menos 52,5 %, y en particular al menos 54 %, al menos 55 %, al menos 56 %, al menos 56,5 % aún más preferentemente al menos 58 %, al menos 59 %, aún más preferentemente al menos 60 %, incluso más preferentemente al menos 62 %, lo más preferentemente al menos 64 %, y en particular al menos 66 %, determinada preferentemente por medio del

software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica en la presente memoria. La proteína Pfa2 puede codificarse por una secuencia de ADN que tiene al menos 54 %, más preferentemente al menos 56 %, aún más preferentemente al menos 58 %, aún más preferentemente al menos 60 %, incluso más preferentemente al menos 62 %, lo más preferentemente al menos 64 %, y en particular al menos 66 % con la SEQ ID NO: 7 y/o con SEQ ID NO: 38 (Soce) y/o SEQ ID NO: 69 y/o SEQ ID NO: 103 y/o SEQ ID NO: 143 (*Aetherobacter*), determinada preferentemente por medio del software Geneious y la herramienta de alineamiento ClustalW, como se especifica en la presente memoria.

En una realización preferente, el agrupamiento de genes de acuerdo con la invención comprende una subsecuencia que codifica una proteína Pfa3 que tiene una identidad proteica con la SEQ ID NO: 4, SEQ ID NO: 10 y/o con SEQ ID NO: 35, SEQ ID NO: 41 (Soce) y/o SEQ ID NO: 66, SEQ ID NO: 72 (*Aetherobacter*) y/o SEQ ID NO: 100, SEQ ID NO: 106 (*Aetherobacter*) y/o SEQ ID NO: 140, SEQ ID NO: 146 (*Aetherobacter*) de al menos 28 %, al menos 31 %, al menos 34 %, al menos 35 %, al menos 37 %, al menos 40 %, más preferentemente al menos 43 %, al menos 46 %, aún más preferentemente al menos 48 %, al menos 50 %, aún más preferentemente al menos 52 %, incluso más preferentemente al menos 54 % o al menos 56 %, aún más preferentemente al menos 58 %, al menos 59 %, aún más preferentemente al menos 60 %, incluso más preferentemente al menos 62 %, lo más preferentemente al menos 64 %, y en particular al menos 66 %, determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica en la presente memoria. La proteína Pfa3 puede codificarse por una secuencia de ADN que tiene al menos 49 %, más preferentemente al menos 51 %, al menos 53,5 %, aún más preferentemente al menos 56 %, al menos 57 %, al menos 58,5 %, aún más preferentemente al menos 59 %, incluso más preferentemente al menos 61 %, lo más preferentemente al menos 63,5 %, y en particular al menos 66 % con la SEQ ID NO: 9 y/o con SEQ ID NO: 40 (Soce) y/o SEQ ID NO: 71 y/o SEQ ID NO: 105 y/o SEQ ID NO: 145 (*Aetherobacter*), determinada preferentemente por medio del software Geneious y la herramienta de alineamiento ClustalW, como se especifica en la presente memoria.

En una realización preferente, el agrupamiento de genes de acuerdo con la invención comprende una subsecuencia que tiene una identidad por pares de ADN con SEQ ID NO: 1 y/o con SEQ ID NO: 32 (Soce) y/o con SEQ ID NO: 63, con SEQ ID NO: 97 y/o con SEQ ID NO: 137 (*Aetherobacter*) de al menos 49 %, más preferentemente al menos 51 %, aún más preferentemente al menos 53 %, más preferentemente al menos 53,5 %, aún más preferentemente al menos 55 %, aún más preferentemente al menos 56 %, incluso más preferentemente al menos 57 %, aún más preferentemente al menos 58,5 %, lo más preferentemente al menos 59 %, incluso más preferentemente al menos 61 %, lo más preferentemente al menos 63,5 %, y en particular al menos 66 %, como se determina preferentemente por medio de el software Geneious y la herramienta de alineamiento ClustalW, como se especifica en la presente memoria.

Preferentemente, el número de pb del agrupamiento de genes que se abarca por cualquiera de las subsecuencias anteriores es de al menos 50 %, más preferentemente al menos 55 %, aún más preferentemente al menos 60 %, aún más preferentemente al menos 65 %, incluso más preferentemente al menos 70 %, lo más preferentemente al menos 75 % y en particular al menos 80 %, del número total de pb del agrupamiento de genes.

En general, el tipo de organismo de producción que se emplea preferentemente en el proceso de acuerdo con la invención no está particularmente limitado.

Las mixobacterias se clasifican científicamente en el reino de manera preferente: bacterias, filo, proteobacterias, clase: deltaproteobacterias, orden: Myxococcales.

Para los fines de la invención, los términos "mixobacterias" y "cepa mixobacteriana" se refieren preferentemente a cualquier especie u otros miembros de microorganismos que pertenecen al orden de Myxococcales.

El término utilizado "miembros" se utiliza preferentemente para cualquier cepa mixobacteriana de la invención que se define como un vecino homólogo de *Aetherobacter*.

Para los fines de la invención, todas las relaciones taxonómicas utilizadas son acordes a Brenner DJ, *et al.* (eds.) *Bergey's Manual of Systematic Bacteriology*, 2ª ed., Nueva York: Springer, especialmente los métodos de la invención se refieren a esta noma.

Preferentemente, el genoma del organismo de producción tiene un contenido GC de al menos 50,0 %, más preferentemente al menos 52,5 %, aún más preferentemente al menos 55,0 %, aún más preferentemente al menos 57,5 %, incluso más preferentemente al menos 60,0 %, lo más preferentemente al menos 62,5 %, y en particular al menos 65,0 %.

Preferentemente, el organismo de producción se selecciona entre el grupo que consiste en eubacterias gramnegativas, eubacterias grampositivas, hongos y levaduras.

De acuerdo con el proceso de acuerdo con la invención, la producción transcurre por medio de la expresión génica heteróloga en un organismo de producción que comprende un agrupamiento de genes derivado de un organismo

fuelle que difiere del organismo de producci3n.

El requisito previo m3nimo para satisfacer este requerimiento es que el organismo de producci3n no sea id3ntico al organismo de tipo natural que sirve como organismo fuente, es decir, el organismo fuente del que se origina el agrupamiento de genes identificado.

En una realizaci3n preferente, el organismo de producci3n y el organismo fuente son microorganismos.

En lo que se refiere al grado taxon3mico de la relaci3n, el organismo de producci3n y el organismo fuente pertenecen preferentemente a diferentes especies o g3neros; m3s preferentemente a diferentes infratribus, subtribus o tribus; a3n m3s preferentemente a diferentes infrafamilias, subfamilias o familias; a3n m3s preferentemente a diferentes sub3rdenes u3rdenes; incluso m3s preferentemente a diferentes subgrupos o grupos; lo m3s preferentemente a diferentes subclases o clases; y en particular a diferentes subfilos, filos, reinos o dominios.

En una realizaci3n preferente, el organismo de producci3n es un organismo procari3tico, preferentemente seleccionado entre eubacterias gramnegativas y eubacterias grampositivas.

Las eubacterias gramnegativas preferentes induyen bacterias, proteobacterias, gammaproteobacterias (incluye *Pseudomonas*, *Escherichia*, *Myxococcus*); particularmente preferentes son *Pseudomonas putida* (proteobacterias, gammaproteobacterias), *pseudomonadales*, *pseudomonadaceae*, *Pseudomonas*, *Escherichia coli* (proteobacterias, gammaproteobacterias), *Enterobacteriales*, *Enterobacteriaceae*, *Escherichia*, *Myxococcus xanthus* (proteobacterias), *deltaproteobacterias*, *Myxococcales*, *cystobacterineae*, *myxococcaceae*, *myxococcus*.

Las eubacterias grampositivas preferentes incluyen actinobacterias; son particularmente preferentes los miembros de los g3neros *Corynebacterium*, *Rhodococcus*, *Streptomyces*, *Bacillus*, *Corynebacterium glutamicum* (bacterias), actinobacterias, *actinobacteridae*, *actinomycetales*, *corynebacterineae*, *corynebacteriaceae*, *corynebacterium*, *Rhodococcus jostii* (bacterias, actinobacterias, *actinobacteridae*, *actinomycetales*, *corynebacterineae*), *nocardiaceae*, *Rhodococcus*, *Streptomyces coelicolor* (bacterias, actinobacterias, *actinobacteridae*, *actinomycetales*, *streptomycineae*, *streptomycetaceae*, *streptomyces*).

En otra realizaci3n preferente, el organismo de producci3n es un organismo eucariota, preferentemente seleccionado entre hongos y levaduras.

Los hongos y las levaduras preferentes induyen *Pichia pastoris* (*ascomycota*, *saccharomycetes*, *saccharomycetidae*, *pichiaceae*, *pichia*), *Saccharomyces cerevisiae* (*ascomycota*, *saccharomycetes*, *saccharomycetidae*), *saccharomycetaceae*, *Saccharomyces*, *hansenula* spp. (*ascomycota*, *saccharomycetes*, *saccharomycetidae*, *saccharomycetaceae*, *Torulaspora* spp. (*ascomycota*, *saccharomycetes*, *saccharomycetidae*, *saccharomycetaceae*, *Yarrowia lipolytica* hongos; *dikarya*; *ascomycota*; *saccharomyceta*; *saccharomycotina*; *saccharomycetes*; *saccharomycetales*; *dipodascaceae* (incluyendo su etapa asexual, *Candida lipolytica*).

En una realizaci3n preferente, el organismo de producci3n es una levadura metilotr3fica u oleaginosa, es decir, la levadura que tiene una fisiolog3a que le permite crecer en metanol y que por lo general tambi3n acumula una gran cantidad de 3cidos grasos para conservar energ3a. Las preferentes son *Hansenula*, *Pichia*, *Candida*, *Torulopsis*, *yarrowia* y tambi3n la levadura *basidiomycete Rhodotorula*.

Normalmente, el organismo de producci3n tiene un estado GRAS, es decir, est3 generalmente reconocido como seguro.

De acuerdo con el proceso de la invenci3n, el agrupamiento de genes se deriva de un organismo fuente. El organismo fuente puede ser cualquier organismo que comprende un agrupamiento de genes que codifica una v3a biosint3tica del 3cido graso poliinsaturado y abarca una subsecuencia ER que codifica una enoil reductasa y una subsecuencia AT que codifica una aciltransferasa, en la que la subsecuencia AT se ubica aguas abajo con respecto a la subsecuencia ER.

Preferentemente, el organismo fuente se selecciona entre mixobacterias.

Se cree que las mixobacterias han evolucionado como un grupo monofil3tico de organismos en el orden *Myxococcales*, un subgrupo delta en proteobacterias. En la actualidad, 3 sub3rdenes (*Cystobacterineae*, *Nannocystineae*, y *Sorangiiineae*) se reconocen en las mixobacterias [Reichenbach H (2005) *Orden VIII. Myxococcales* Tchan, Pochon y Pr3vot 1948, 398AL. En Brenner DJ, et al. (eds.) *Bergey's manual of Systematic Bacteriology*, 2ª ed., vol. 2, parte C, p3gs. 1.059-1.072, Nueva York: Springer]. Estos sub3rdenes se dividen en seis familias, a saber, *Cystobacteraceae*, *Myxococcaceae*, *Nannocystaceae*, *Kofleriaceae*, *Polyangiaceae* y *Phaselicystidaceae*.

La familia *Myxococcaceae* se compone de los g3neros *Myxococcus*, *Corallococcus* y *Pyxidicoccus*. En la familia relacionada *Cystobacteraceae*, se conocen cinco g3neros (*Cystobacter*, *Archangium*, *Hyalangium*, *Meliittangium* y

Stigmatella). *Nannocystaceae* del suborden *Nannocystineae* se componen de *Nannocystis* y dos géneros marinos (*Enhygromyxa* y *Plesiocystis*). Su familia relacionada *Koffleriaceae* se compone del género terrestre *Koffleria* y el género marino *Haliangium*. La familia *Polyangiaceae* abarca los géneros *Jahnella*, *Chondromyces*, *Polyangium*, *Byssovorax* y *Sorangium*. Hasta el momento, los dos últimos son los únicos géneros conocidos de microorganismos degradadores de celulosa entre el orden; la mayoría de los otros grupos taxonómicos son difíciles de aislar y cultivar. El género recientemente descubierto, *Phaselicystis*, es el único género de la familia recientemente erigido *Phaselicystidaceae* [García RO *et al.* (2009) *Int J Syst Evol Microbiol* 59:1524-1530]. En la actualidad, 20 géneros son reconocibles y válidamente descritos en mixobacterias para contener todos los aislados terrestres y marinos conocidos.

Myxococcales incluyen los siguientes subórdenes ejemplificados que, a su vez, incluyen las siguientes familias ejemplificadas, géneros ejemplificados y especies ejemplificadas. El organismo fuente se selecciona preferentemente entre cualquiera de las siguientes especies:

suborden	familia	género	especies (ejemplos preferentes)
<i>Cystobacteriineae</i>	<i>Cystobacteraceae</i>	<i>Archangium</i> <i>Cystobacter</i> <i>Hyalangium</i> <i>Melittangium</i> <i>Stigmatella</i>	
	<i>Myxococcaceae</i>	<i>Anaeromyxobacter</i> <i>Corallocooccus</i> <i>Myxococcus</i> <i>Pyxidiooccus</i>	
<i>Nannocystineae</i>	<i>Haliangiaceae</i> (<i>Koffleriaceae</i>)	<i>Haliangium</i>	
	<i>Nannocystaceae</i>	<i>Nannocystis</i> <i>Plesiocystis</i> <i>Enhygromyxa</i>	<i>Enhygromyxa salina</i>
<i>Sorangiiinae</i>	<i>Phaselicystidaceae</i>	<i>Phaselicystis</i>	
	<i>Polyangiaceae</i>	<i>Polyangium</i> <i>Jahnella</i> <i>Chondromyces</i> <i>Byssovorax</i> <i>Aetherobacter</i>	<i>Byssovorax cruenta</i> <i>Aetherobacter fasciculatus</i> (DSM 21835) <i>Aetherobacter rufus</i> (DSM 23122) <i>Aetherobacter sp.</i> (DSM 23098)
	<i>Sorangium</i>		<i>Sorangium cellulosum</i>

Las relaciones filogenéticas como se infiere a partir de una comparación de datos de la secuencia de ADNr 16S de los géneros y especies representativos cuyos cultivos viables son existentes, se presentan en la Fig. 5.

En una realización preferente, el agrupamiento de genes se deriva de una cepa mixobacteriana que pertenece al suborden *Sorangiiinae* del orden *Myxococcales*. En otra realización preferente, el agrupamiento de genes se deriva de una cepa mixobacteriana que pertenece al suborden *Polyangiaceae* del orden *Myxococcales*. En una realización preferente, la cepa mixobacteriana se selecciona entre las cepas *Aetherobacter fasciculatus* DSM 21835, *Aetherobacter rufus* DSM 23122, y *Aetherobacter sp.* DSM 23098. En otra realización preferente, la cepa mixobacteriana es *Sorangium cellulosum* So ce56.

En una realización preferente, el agrupamiento de genes se deriva de una cepa mixobacteriana que produce una composición que comprende ácidos grasos poliinsaturados, en la que el contenido de ácidos grasos poliinsaturados omega-3 es al menos el 10 %, preferentemente al menos 15 % en peso de contenido de ácidos grasos celulares totales de la composición.

Una vez que un agrupamiento de genes deseado de un organismo fuente se ha identificado y aislado, se puede expresar en un organismo de producción. Para los fines de expresar de forma heteróloga tal agrupamiento de genes en un organismo de producción, éste se vincula operativamente a una secuencia promotora y terminadora utilizando técnicas o patrones de clonación convencionales en los procedimientos *in vitro*, tales como fusión por PCR. El término "promotor" se refiere a una secuencia de ADN capaz de controlar la expresión de un gen. Las construcciones que contienen agrupamientos de genes de interés pueden introducirse en el organismo de producción por técnicas convencionales. Estas técnicas incluyen la transformación, tal como, por ejemplo, en *S. cerevisiae*, la transformación de acetato de litio, formación de esferoplastos, y el uso de un mutante *kar1* n.º 15 (Georgieva, B. *et al.*, (2002) *Meth. Enzymol.* 350:278-89), fusión de protoplastos, lipofección, transfección, transducción, conjugación, infección, impacto balístico, electroporación, o cualquier otro método que introduce el ADN exógeno en la célula del organismo de producción. Por razones de simplicidad, un organismo de producción manipulado de esta manera se referirá como "transformado", "recombinante" o "genéticamente modificado". La construcción que se introduce en el organismo de producción contiene preferentemente además del casete de expresión un gen marcador, que permite la identificación de células transformadas y, en el caso de la expresión extracromosómica, también evita que la célula pierda la construcción.

La etapa (ii) del proceso de acuerdo con la invención comprende el cultivo del organismo de producción de la etapa (i) en presencia de una fuente de carbono fermentable por la cual se producen uno o varios ácidos grasos poliinsaturados.

Normalmente, los organismos de producción se fermentan en un medio nutriente que contiene una fuente de carbono fermentable y un material proteínico. Las fuentes de carbono preferentes incluyen glucosa, azúcar marrón, sacarosa, glicerol, almidón, almidón de maíz, lactosa, dextrina, melaza, y similares. Las fuentes de nitrógeno preferentes incluyen harina de semilla de algodón, licor de maceración del maíz, levadura, levadura de cerveza autolizada con sólidos lácteos, soja triturada, harina de semillas de algodón, harina de maíz, sólidos lácteos, digerido pancreático de caseína, sólidos de destilería, licores de peptona de origen animal, carne y trozos de hueso, y similares. Las combinaciones de estas fuentes de carbono y nitrógeno se pueden utilizar ventajosamente. No hay necesidad de añadir metales traza, p. ej., cinc, magnesio, manganeso, cobalto, hierro y similares, al medio de fermentación ya que el agua del grifo y los ingredientes no purificados se utilizan como componentes del medio.

Preferentemente, la fuente de carbono fermentable está libre de ácidos grasos, preferentemente se selecciona entre el grupo que consiste en monosacáridos, oligosacáridos, polisacáridos, y aminas que contienen metanol y carbono.

La fermentación a gran escala para cultivos de producción se puede inducir a cualquier temperatura conductora para un crecimiento satisfactorio de los organismos de producción entre aproximadamente 18 y 32 °C y preferentemente a aproximadamente 28 °C. Habitualmente, la producción óptima de los compuestos se obtiene en aproximadamente 2 a 8 días de fermentación, y preferentemente en aproximadamente 4 a 5 días de fermentación.

La producción puede llevarse a cabo en matraces de agitación pero también en medios sólidos y fermentadores agitados. Cuando el crecimiento se lleva a cabo en matraces de agitación o en grandes recipientes y tanques, es preferible utilizar la forma vegetativa, en lugar de la forma de esporas de los organismos de producción para la inoculación. Esto evita un retraso pronunciado en la producción de los compuestos de AGPI y la utilización ineficaz asistente del equipo. En consecuencia, es deseable producir un inóculo vegetativo en un medio nutriente acuoso mediante la inoculación de este medio con una alícuota de un cultivo terrestre o indinado. Cuando un inóculo vegetativo activo joven ha sido por lo tanto asegurado, se transfiere asépticamente a otros matraces de agitación u otros dispositivos adecuados para la fermentación de organismos de producción. El medio en el que se produce el inóculo vegetativo puede ser el mismo, o diferente, al utilizado para la producción de compuestos, siempre que se obtenga el crecimiento adecuado del microorganismo.

En general, se utiliza la siembra de las cepas bacterianas y la fermentación y la producción de compuestos en la fermentación aerobia sumergida en recipientes agitados. La producción es independiente de los contenedores, fermentadores y procedimientos iniciadores utilizados. Los compuestos también se pueden obtener por cultivo en matraz de agitación, o en otros recipientes especialmente diseñados, tales como tanques de fermentación aerolevado o Biowave. Para fermentaciones en grandes volúmenes es preferible utilizar un inóculo vegetativo. El inóculo vegetativo se prepara inoculando un pequeño volumen del medio de cultivo con la forma de esporas o un sedimento liofilizado del organismo. El inóculo vegetativo se transfiere entonces a un recipiente de fermentación, en el que, después de un tiempo de incubación adecuado, los compuestos se producen con un rendimiento óptimo.

Como es habitual en el proceso de cultivo aerobio sumergido, el aire estéril se dispersa por el medio de cultivo. Para un crecimiento eficiente del organismo, el volumen del aire utilizado se encuentra en el intervalo de aproximadamente 0,25 a aproximadamente 0,5 vvm. Una tasa óptima en un recipiente de 10 l es de aproximadamente 0,3 vvm con agitación proporcionada por impulsores convencionales que giran a aproximadamente 240 rpm. La adición de una pequeña cantidad (es decir, 1 ml/l) de un agente antiespumante, tal como silicona para medios de fermentación es necesaria si la formación de espuma se convierte en un problema. Para los organismos microaerófilos puede ser favorable reducir aún más la aireación a efectos de soportar la

producción de biomasa. La fermentación se lleva a cabo por lo general en el modo por lotes, pero para lograr un mejor crecimiento y un mayor rendimiento del producto, fermentaciones por lotes semicontinuos pueden llevarse a cabo mediante el suministro de la fuente de nutrientes requerida para cultivar un cultivo una vez que se ha deplecionado en el medio de cultivo original.

5 Los productos deseados estarán normalmente presentes en su mayoría en la biomasa de las cepas bacterianas fermentadas, pero en caso de su sobreproducción, pueden ubicarse también en el filtrado de cultivo del caldo de fermentación. El caldo de cultivo se puede separar por filtración en un filtro prensa.

10 La etapa (iii) del proceso de acuerdo con la invención solo es opcional y comprende la recuperación de uno o más ácidos grasos poliinsaturados producidos en la etapa (ii).

15 Los métodos adecuados para la recuperación de ácidos grasos son conocidos por el experto e incluyen extracción, cromatografía y similares. En particular, una variedad de procedimientos se puede emplear para aislar y purificar los compuestos de AGPI a partir del caldo de fermentación, por ejemplo, mediante procedimientos cromatográficos de adsorción seguidos de elución con un disolvente adecuado, cromatografía en columna, cromatografía de partición, por extracción con fluidos supercríticos, y combinaciones de los métodos antes mencionados.

20 Un aspecto adicional de la invención se refiere a un método para producir uno o más ácidos grasos poliinsaturados por medio de la expresión génica heteróloga de genes que codifican las enzimas de la vía biosintética para la síntesis de AGPI como se describe en la presente memoria, comprendiendo el método preferentemente las etapas de:

25 (i) proporcionar un organismo de producción que comprende un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados, derivando dicho agrupamiento de genes de un organismo fuente que difiere del organismo de producción; y que abarca una primera subsecuencia y/o una segunda subsecuencia, en la que la proteína codificada por dicha primera subsecuencia tiene una identidad proteica con la secuencia de aminoácidos de Pfa1 de al menos 46 %, más preferentemente al menos 48 %, aún más preferentemente al menos 50 %, aún más preferentemente al menos 52 %, incluso más preferentemente al menos 54 %, lo más preferentemente al menos 56 %, y en particular al menos 58 %; determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica con más detalle en la presente memoria; y/o

30 que codifica una proteína que tiene una identidad proteica con la secuencia de aminoácidos de Pfa2 de al menos 2,5 %, más preferentemente al menos 5,0 %, más preferentemente al menos 7,5 %, aún más preferentemente al menos 10 %, aún más preferentemente al menos 12,5 %, incluso más preferentemente al menos 15 %, lo más preferentemente al menos 17,5 %, y en particular al menos 20 %; y/o

35 la proteína codificada por dicha segunda subsecuencia tiene una identidad proteica con la secuencia de aminoácidos de Pfa3 de al menos 2,5 %, más preferentemente al menos 5,0 %, más preferentemente al menos 7,5 %, aún más preferentemente al menos 10 %, aún más preferentemente al menos 12,5 %, incluso más preferentemente al menos 15 %, lo más preferentemente al menos 17,5 %, y en particular al menos 20 %; determinada preferentemente por medio del software Geneious y la herramienta de alineamiento Geneious Alignment, como se especifica con más detalle en la presente memoria;

40 (ii) cultivar el organismo de producción de la etapa (i) en presencia de una fuente de carbono fermentable por medio de la cual se producen uno o más ácidos grasos poliinsaturados; y

45 (iii) opcionalmente, recuperar uno o más ácidos grasos poliinsaturados.

50 Preferentemente, la primera subsecuencia es una subsecuencia ER como se describe anteriormente y/o la segunda subsecuencia es una subsecuencia AT como se describe anteriormente.

55 Todas las realizaciones preferentes que se han descrito en relación con el primer aspecto de la invención (proceso de acuerdo con la invención) también se refieren de manera análoga a este aspecto de la invención (método de acuerdo con la invención) y por lo tanto, no se repiten de aquí en adelante.

Se desvela adicionalmente en la presente memoria una composición que comprende al menos dos ácidos grasos poliinsaturados diferentes obtenibles por el proceso de acuerdo con la invención y el método de acuerdo con la invención, respectivamente.

60 La filogenia es conforme con las características morfológicas y fisiológicas de las mixobacterias. Lo más importante, los perfiles de ácidos grasos como se infiere por análisis por GC-MS del contenido de ácidos grasos celulares se utilizan generalmente y se consideran aceptables para la segregación taxonómica de mixobacterias, así como muchos otros grupos de organismos bacterianos, ya que se descubrió que era una característica constante, al menos cuando se aplica una metodología estandarizada. Las primeras aplicaciones de esta técnica se realizaron en la década de 1980 [Tornabene G (1985) *Methods in Microbiology* 18, 209-234]. Por lo tanto, tales perfiles de ácidos grasos basados en GC-MS (o CG) se han utilizado ampliamente en la filogenia y la taxonomía bacteriana.

Así, los agrupamientos de genes derivados de diferentes especies competentes individuales, es decir especies diferentes que son capaces de producir AGPI, dan productos de diferente composición, es decir, productos que contienen diferentes AGPI y/o AGPI en diferentes cantidades. De este modo, los productos que pueden obtenerse por los diferentes procesos tienen características individuales que los distinguen entre sí.

Un aspecto adicional de la invención se refiere a un organismo de producción transformado que comprende un agrupamiento de genes que codifica una vía biosintética de ácidos grasos poliinsaturados como se ha definido anteriormente.

La invención se describe ahora con mayor detalle con referencia a las figuras y por medio de ejemplos que, sin embargo, no deben interpretarse como limitantes del alcance de las reivindicaciones.

Aetherobacter fasciculatus, actualmente también denominada SBSr002, se depositó de acuerdo con el Tratado de Budapest en el DMSZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Inhoffenstr. 7 B, 38124 Braunschweig, Alemania, el 27 de agosto de 2008, con el número de depósito DSM 21835. *Aetherobacter rufus*, actualmente también denominada SBSr003, se depositó ante el DSMZ con el número de depósito DSM 23122, y *Aetherobacter sp.*, actualmente también denominada SBSr008, se depositó ante el DSMZ con el número de depósito DSM 23098 ante DSMZ el 12 de noviembre de 2009.

A. fasciculatus DSM 21835 se identificó como un miembro de la clase mixobacterias, orden *Myxococcales* mostrando el enjambre característico de las células vegetativas gramnegativas, con forma de bastón finas, formación de cuerpos de fructificación, y actividad bacteriolítica. La cepa es aerobia con respecto a aerobio facultativo, quimioheterotrófico, y también exhibe resistencia a diversos antibióticos. Los ácidos grasos principales son C22:6 (ácido docosahexaenoico), iso-C15:0, ante-iso C17:0 y C20:5 (ácido eicosapentaenoico). El contenido de G + C del ADN genómico es 68,9 % en moles. La secuencia de ADNr 16S muestra una identidad del 96 % con respecto a *Byssovorax cruenta* de degradación de celulosa y 95 % con respecto a *Sorangium cellulosum*. Esto demuestra claramente que la cepa pertenece al suborden *Sorangineae* del orden *Myxococcales*. Además, la singularidad en las fases de crecimiento morfológico y el innovador perfil de ácidos grasos, implican claramente que DSM 21835 pertenece a un nuevo taxón que se clasifica de forma propuesta como perteneciente al género innovador *Aetherobacter* y a la especie innovadora *A. fasciculatus*.

Ejemplo 1: Generación de una construcción de expresión para un agrupamiento de genes biosintéticos para la codificación de la producción de AGPI

a) Subclonación de los genes de AGPI y construcción de las construcciones de expresión:

Los genes de AGPI se subdonaron a partir de los cromosomas de las cepas de mixobacterias productoras de AGPI mediante la generación de bibliotecas genómicas (p. ej., utilizando un cósmido, fósido, vectores YAC o BAC) o mediante la amplificación de genes de AGPI o fragmentos de los mismos por PCR seguido de subclonación de los productos de PCR. Alternativamente, los genes de AGPI se subdonaron utilizando métodos que se basan en la recombinación, p. ej., la tecnología de recombinación Red/ET (Zhang Y, *Nat. Genet.* 1998, 20:123-128; Zhang Y, *Nat. Biotechnol.* 2000, 18:1314-1317) o la recombinación en levaduras (Kouprina N, *Nature Protocols* 2008, 3:371-377; Larionov V, *P. Natl. Acad. Sci. EE.UU.* 1996, 93:491-496; Wolfgang MC, *Proc Natl. Acad. Sci. EE.UU.* 2003, 100:8484-8489). En otro enfoque, las secuencias de ADN nativas o artificiales de genes de AGPI se sintetizaron químicamente *de novo* (p. ej., mediante síntesis génica, que está disponible comercialmente). Un plásmido preferente que contiene todo el agrupamiento de genes de pfa1, pfa2, pfa3 se muestra en la Fig. 2. Los elementos de RI que flanquean el agrupamiento de genes son opcionales y preferentes para la co-transformación de la sección del plásmido que contiene el agrupamiento de genes y los elementos de RI de flaqueo con un gen que codifica una transposasa, que puede contenerse en la misma o en una construcción de ácido nucleico separada.

Para movilizar los genes de AGPI, la construcción de expresión se modificó con elementos genéticos que permiten la transferencia, propagación y control de la expresión en el huésped heterólogo. Para la transferencia a través de la conjugación, se añade un origen de transferencia (oriT) a la construcción de expresión. Para la propagación de los genes de AGPI en la célula huésped, las construcciones de expresión se equiparon con un origen de replicación, que es funcional en la cepa huésped, o la construcción de expresión se integró en el cromosoma huésped utilizando diversos métodos. Para integrar al azar los genes de AGPI en el cromosoma huésped por transposición (Fu J, *Nucleic Acids Res.* 2008, 36:e113), el agrupamiento de genes biosintéticos de AGPI junto con un marcador de selección adecuado se flanquea por elementos de repetición invertida (RIs) y la construcción de expresión contiene la transposasa del transposón marino pMycoMar (véase la Figura 2) (Rubin EJ, *P. Natl. Acad. Sci. EE.UU.* 1999, 96:1645-1650). La integración dirigida en el cromosoma huésped se logró mediante la recombinación homóloga o por medio de sitios de unión al fago específico del huésped catalizados por integrasas dedicadas (p. ej., sitios de unión mX8 o MX9 en mixobacterias (Magrini V, *J. Bacteriol.* 1999, 181:4050-4061; Julien B, *J. Bacteriol.* 2003, 185:6325-6330), o el sitio de unión Φ C31 en actinomicetos (Kieser T: *Practical Streptomyces Genetics*. Norwich, Inglaterra: Fundación John Innes, 2000)).

b) Modificación del agrupamiento de genes biosintéticos de AGPI:

Para permitir y/u optimizar la expresión heteróloga, el agrupamiento de genes biosintéticos de AGPI se modificó por elementos genéticos de modificación por ingeniería genética implicados en la transcripción/traducción (p. ej., modificación por ingeniería genética de promotores o sitios de unión al ribosoma) o mediante la adaptación de los codones de genes de AGPI para el uso del codón del huésped heterólogo. Para la expresión heteróloga en mixobacterias, las estructuras promotoras nativas se intercambiaron con PT7A1, Pm, Ptet (véase la Figura 2), el promotor npt derivado de tn5 u otros promotores (p. ej., promotores que estimulan la expresión de las vías del metabolito secundario nativo). Para optimizar el perfil de producción de AGPI, se modificaron por ingeniería genética las maquinarias biosintéticas de AGPI mixobacterianas (p. ej., mutación/inserción/delección/intercambio de dominios catalíticos) mediante la modificación de los genes biosintéticos.

Ejemplo 2: Transferencia del agrupamiento de genes en un organismo huésped para la producción heteróloga de AGPI

Las construcciones de expresión de los agrupamientos de genes biosintéticos de AGPI mixobacterianos se transfirieron a las cepas huésped heterólogas mediante electroporación, conjugación u otros métodos de transformación (p. ej., transformación de protoplastos/esferoblastos o utilizando la competencia natural). Los transformantes se analizaron genotípicamente para la presencia (y la integración correcta) de las construcciones que contienen el gen de AGPI. Las cepas huésped que contienen los agrupamientos de genes biosintéticos de AGPI mixobacterianos se cultivaron en condiciones apropiadas (p. ej., para *M. xanthus*: medio CTT a 30 °C durante 2-4 días). La producción de AGPI se analizó por cromatografía de gases acoplada a espectrometría de masas (GC-MS) utilizando el método FAME, p. ej., análisis de esterificación metílica de ácidos grasos.

Ejemplo 3: Generación de una construcción de expresión para el agrupamiento de genes biosintéticos de AGPI a partir de *Aetherobacter fasciculatus* DSM21835 mediante métodos de clonación convencionales

El agrupamiento de genes biosintéticos de AGPI a partir de *Aetherobacter fasciculatus* DSM 21835 se amplificó por PCR a partir de ADN cromosómico. Para ello, el agrupamiento de genes se dividió en varias regiones, que se amplificaron individualmente mientras que los sitios de restricción adecuados se introdujeron con las secuencias de cebadores. En total, 7 fragmentos se amplificaron por PCR: el gen completo pfa1 sin su secuencia promotora nativa (1, Swal-Spel-pfa1-HpaI), el extremo 5' del gen pfa2 (2, HpaI-5'pfa2-HindIII), la sección central del gen pfa2 (3, HindIII-CR-pfa2-NheI), el extremo 3' del gen pfa2 (4, NheI -3'pfa2-AseI), el extremo 5' del gen pfa3 (5, AseI-5'pfa3-HindIII), la sección central del gen pfa3 (6, HindIII-CR-pfa3-NheI), y el extremo 3' del gen pfa3 (7, NheI-3'pfa3-SspI-PacI). Los productos de PCR se subclonaron mediante métodos de restricción y ligamiento convencionales en un vector de clonación de copias elevado (p. ej., derivado de pUC18). La subclonación del fragmento 1 condujo al plásmido pPfa1, la subclonación de fragmentos 2-4 en el mismo vector generó el plásmido pPfa2 y la subclonación de los fragmentos 5-7 en el mismo vector proporcionó el plásmido pPfa3. En la siguiente etapa, los tres genes biosintéticos de AGPI (pfa1, pfa2 y pfa3) se subdonaron en el mismo plásmido. Para ello, los genes pfa2 y pfa3 se ligaron en el plásmido pPfa1 que alberga el gen pfa1 que genera la construcción pPfa1-2-3. Los elementos genéticos requeridos para la expresión de los genes pfa en el huésped heterólogo (secuencias promotoras, genes reguladores, sitios de unión al ribosoma) se ligaron a través de los sitios de restricción Swal-Spel derivados del fragmento 1. Los casetes requeridos para la transferencia y la replicación o integración en el cromosoma huésped se introdujeron a través de sitios de restricción únicos del plásmido pPfa1-2-3.

Ejemplo 4: Generación de una construcción de expresión para el agrupamiento de genes biosintéticos de AGPI a partir de *Aetherobacter fasciculatus* DSM21835 mediante técnicas de recombinación genética

El agrupamiento de genes biosintéticos de AGPI se subdonó directamente a partir de ADN cromosómico de *A. fasciculatus* DSM 21835 mediante técnicas de recombinación. Para ello, el ADN genómico se digirió con enzimas de restricción de corte en las regiones flanqueantes del agrupamiento de genes biosintéticos de AGPI revelando una mezcla de fragmentos de ADN genómico. Un casete que contiene un origen de replicación, así como un marcador de selección (p. ej., p15Aori-cmR de pACYC184 o p15Aori-ampR de pACYC177) se amplificó por PCR y brazos de homología a ambos extremos del fragmento de ADN genómico que alberga el agrupamiento de genes biosintéticos de AGPI se introdujeron con la secuencia del cebador. Además, se introdujeron sitios de restricción únicos (p. ej., PacI) para la subclonación posterior de casetes de transferencia junto con el brazo de homología del extremo 3' del agrupamiento de genes. El fragmento de ADN genómico que alberga el agrupamiento de genes biosintéticos de AGPI se subdonó a continuación en el producto de PCR por recombinación homóloga doble *in vivo*. La construcción recombinante se modificó adicionalmente mediante una segunda etapa de doble recombinación homóloga para integrar un gen de resistencia amplificado por PCR y flanqueado por dos brazos de homología y sitios de restricción (p. ej., Swal y Spel) en el extremo 5' del agrupamiento de genes biosintéticos de AGPI. A continuación, la construcción obtenida se modificó por ingeniería genética adicionalmente mediante métodos de clonación convencionales para introducir casetes requeridos para la transferencia y la expresión heteróloga: elementos genéticos requeridos para la expresión de los genes pfa en el huésped heterólogo (secuencias promotoras, genes reguladores, sitios de unión a ribosomas) se ligaron a través de sitios de restricción Swal-Spel y casetes requeridos para la transferencia y replicación o integración en el cromosoma del huésped se introdujeron a través de sitios de

restricción únicos en el extremo 3' del agrupamiento de genes biosintéticos de AGPI.

Ejemplo 5: Transformación del agrupamiento de genes biosintéticos de AGPI a partir de *Aetherobacter fasciculatus* DSM 21835 en organismos huésped y expresión heteróloga

5 M. xanthus se cultivó en medio CTT (10 g/l de casitona, 10 ml/l de tampón Tris, 1 ml/l de tampón fosfato de potasio (pH 7,6; 86 ml de K₂HPO₄ 1 M + 13,4 ml de KH₂PO₄ 1 M), 10 ml/l de sulfato de magnesio 0,8 M) a 30 °C hasta que se alcanzó la DO₆₀₀ de 0,5. 10 ml del cultivo bacteriano se centrifugaron a 13.000 rpm a 4 °C, el sedimento celular se lavó dos veces con agua destilada helada y finalmente se resuspendió en 40 µl de agua helada. La suspensión
10 celular se sometió a electroporación con la construcción de expresión en una cubeta de electroporación de 0,1 cm helada bajo las siguientes condiciones: 400 Ω, 25 µF y 650 V para lograr una longitud de pulso de 8-9 ms. Después, se añadió 1 ml de medio CTT y se incubaron las células durante 18 h a 30 °C. 3 ml de agar suave CTT (medio CTT más 7,5 g/l de agar) que contiene 25 µl de 50 mg/ml de kanamicina se añadieron al cultivo y la mezcla se vertió sobre la superficie de placas de agar CTT (medio CTT más 15 g/l de agar). Las placas se incubaron durante 3 días a
15 30 °C hasta que las colonias mutantes aparecieron. Los transformantes se analizaron luego genotípica y fenotípicamente.

Ejemplo 6: Producción heteróloga de AGPI

20 Mutantes de *M. xanthus* que albergan el agrupamiento de genes biosintéticos de AGPI se cultivaron en medio CTT suplementado con 50 µg/ml de kanamicina a 30 °C durante 3-5 días. El cultivo se centrifugó y se extrajeron los AGPI. Los ácidos grasos celulares se extrajeron utilizando el método FAME [Bode HB *et al.* (2006) *J. Bacteriol* 188:6524-6528; Ring MW *et al.* (2006), *J Biol Chem* 281:36691-36700 (2006)]. Alícuotas (1 µl) de los extractos se analizaron mediante GC-MS. Identificación de los ácidos grasos celulares incluyendo AEP y ADH: ácidos grasos
25 celulares incluyendo los AGPI omega-3 (AEP y ADH) se identificaron basándose en los patrones de fragmentación y el tiempo de retención. Estos ácidos grasos (AG) se compararon con un patrón de referencia de la mezcla por FAME (Sigma-Aldrich) que contiene 37 ésteres metílicos de ácidos grasos. La presencia de ADH y AEP se confirmó utilizando estándares de referencia de Sigma-Aldrich (cis-4,7,10,13,16,19-ADH, cis-5,8,11,14,17-AEP).

30 La construcción de ácidos nucleicos para la expresión heteróloga del agrupamiento de genes biosintéticos de AGPI a partir de *Sorangium cellulosum* So ce56 puede utilizarse para la expresión heteróloga de los productos génicos sintéticos en las eubacterias o en otras mixobacterias, p. ej., *Myxococcus xanthus* DK1622.

35 La Fig. 3 muestra el resultado analítico de la producción de AGPI por *M. xanthus* que contiene el agrupamiento de genes pfa1, pfa2, pfa3 de Soce56 (SEQ ID NO: 1) de una construcción de ácido nucleico de acuerdo con la Fig. 2.

Como un ejemplo adicional para la producción de AGPI utilizando un microorganismo productor que está genéticamente manipulado para contener el agrupamiento de genes que codifica Pfa1, Pfa2 y Pfa3 para la expresión heteróloga, *M. xanthus* se transformó con una construcción de ácido nucleico como se muestra en la Fig. 4, que contenía pfa1, pfa2, pfa3 clonado a partir de *Aetherobacter* SBSr002 (SEQ ID NO: 63, o alternativamente, SEQ ID NO: 67, SEQ ID NO: 69 y SEQ ID NO: 71). En detalle, el ADN genómico de DSM 21835 se digirió con el corte de Scal en las regiones flanqueantes del agrupamiento de genes biosintéticos de AGPI revelando una mezcla de fragmentos de ADN genómico. Un casete que contiene un origen de replicación p15A, así como un gen de resistencia a ampicilina (p15A ori-amp^R de pACYC177) se amplificó por PCR y brazos de homología a ambos
45 extremos del fragmento de ADN genómico que alberga el agrupamiento de genes biosintéticos de AGPI se introdujeron con la secuencia del cebador. Además, un sitio de restricción único PaeI para la subdonación posterior de casetes de transferencia se introdujo junto con el brazo de homología del extremo 3' del agrupamiento de genes. El fragmento de ADN genómico que alberga el agrupamiento de genes biosintéticos de AGPI se subclonó a continuación en el producto de PCR por doble recombinación homóloga *in vivo*. La construcción recombinante se modificó adicionalmente mediante una segunda etapa de doble recombinación homóloga para integrar un gen de resistencia a kanamicina, así como un promotor P_{tet} amplificado por PCR y flanqueado por dos brazos de homología en el extremo 5' del agrupamiento de genes biosintéticos de AGPI. A continuación, la construcción se modificó
50 adicionalmente por ingeniería genética mediante métodos de clonación convencionales: El casete *zeo^R-mx9* que contiene un gen de resistencia a zeocina y un gen que codifica la integrasa Mx9 con sitios de unión correspondiente a fagos (attP) para la integración en el cromosoma del huésped se introdujo a través del sitio de restricción único PaeI en el extremo 3' del agrupamiento de genes biosintéticos de AGPI. La construcción de expresión de AGPI se transforma entonces en *M. xanthus* DK1622. Los transformantes se analizaron genotípica y fenotípicamente. Los transformantes de *M. xanthus* que albergan el agrupamiento de genes biosintéticos de AGPI se cultivaron en medio CTT suplementado con 60 µg/ml de kanamicina y 20 µg/ml de zeocina a 30 °C durante 2-3 días. El cultivo se centrifugó y se extrajeron los AGPI, se esterificaron y se analizaron por GC-MS.
60

Como un ejemplo adicional para la producción de AGPI mediante el cultivo de un organismo de producción en un proceso de la invención, se utilizó un biorreactor agitado de 10 l (Biostat E, Braun Melsungen) para cultivar la cepa de *M. xanthus* que alberga el agrupamiento de genes de AGPI de Soce56, y por separado *M. xanthus* que alberga el agrupamiento de genes de AGPI de *Aetherobacter fasciculatus*, cada una como se describe en la presente memoria, en el medio apropiado a 2 l/min de aireación, agitación a 120 rpm a 28 °C durante 5 a 7 días. Para la inoculación, se
65

utilizaron precultivos en matraces de agitación durante 72 h. Los análisis de AGPI metilado mostraron 0,4 a 1,6 % de ácido γ -linolénico para cada cepa y 0,6 % de ADH para *M. xanthus* que expresa el agrupamiento de genes de *A. fasciculatus* en la biomasa seca.

Ejemplo 7: Comparación de secuencias

5 Los dominios catalíticos en las nuevas sintasas de AGPI mixobacterianas se anotaron utilizando diversas herramientas bioinformáticas que se enumeran en la sección siguiente, donde ER = enoil reductasa, CS = β -cetoacil sintasa, MAT = malonil/acetil transferasa, PPA = proteína transportadora de acilos, CR = β -cetoacilreductasa, DH = β -hidroxiacil deshidratasa:

10 el dominio Pfa1_ER se anotó por alineamientos con las secuencias correspondientes de *Shewanella* sp. SCRC-2738, *Moritella marina* MP-1, *Photobacterium proiundum* TC 9, y *Schizochytrium* sp. CACT_20888. Pfa2_CS-MAT-PPAs-CR se anotó con el análisis PKS/NRPS de Jacques Ravel, sitio web (<http://nrps.igs.umaryland.edu/nrps/>).

15 Pfa2_DH se anotó por alineamientos con las secuencias correspondientes de *Shewanella* sp. SCRC-2738, *Moritella marina* MP-1, *Photobacterium proiundum* SS9, y *Schizochytrium* sp. CACT_20888. Pfa3_CS-FLC-DH-DH se anotó con la búsqueda de secuencias Pfam (<http://pfam.sanger.ac.uk/>).

20 Pfa3_AT se anotó por alineamientos con las secuencias correspondientes de *Shewanella* sp. SCRC-2738, *Moritella marina* MP-1, *Photobacterium proiundum* SS9, y *Schizochytrium* sp. CACT_20888. En una búsqueda de secuencias BlastP (Instituto Nacional del Cáncer, EE. UU.: <http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) y Pfam (versión 24.0; The Wellcome Trust Sanger Institute, Cambridge, RU: <http://pfam.sanger.ac.uk/>), el acierto más significativo representa un dominio AT, pero puede ser que sirva a otra función. Los dominios actualmente adicionales no identificados pueden codificarse en los genes descritos en la presente memoria. Además de las diferencias de organización entre los agrupamientos de genes, las regiones funcionales comparables en los agrupamientos de genes mixobacterianos, como se define en la Figura 1, muestran un nivel inesperadamente bajo

25 de identidad de secuencias en la proteína, así como en el nivel de ADN con respecto a los genes biosintéticos de AGPI conocidos (tabla 1) como se muestra en las tablas 2 y 3. La similitud de secuencias entre los genes de AGPI mixobacterianos de la invención es en general más alta (véase la tabla 4).

Tabla 1: Agrupamientos de genes biosintéticos conocidos comparativos que codifican AGPI-sintasas

Cepa	N.º de acceso a GenBank	Producto	Referencia
<i>Shewanella</i> sp. SCRC-2738	U73935.1	APE/ADP	Metz, <i>Science</i> 2001, 293: 290-293
<i>Moritella marina</i> MP-1 CACT_15381	AB025342.2	ADH	Takeyama, <i>Microbiol.</i> 1997, 143:2725-2731 Tanaka, <i>Biochem. Soc. Trans.</i> 2000, 28:943-945: Morita, <i>Biotechnol. Lett.</i> 1999, 21:641-646
<i>Photobacterium profundum</i> SS9	AF409100.1	APE	Allen, <i>Microbiol.</i> 2002, 148:1903-1913
<i>Schizochytrium</i> sp. CACT_20888	AF378327.2 AF378328.2 AF378329.2	ADH/ADP	Metz, <i>Science</i> 2001

Número de acceso a BLAST de *Sorangium cellulosum* Soce56: YP001611455.1

30 La siguiente comparación de secuencias de las secuencias (AGPAT) del dominio IV* mixobacteriano del agrupamiento de genes de AGPI de *Sorangium cellulosum* So CE56 (= So ce56) y AGPAT de *Aetherobacter fasciculatus* DSM 21835 (= *A. fasciculatus*) con las secuencias correspondientes de los microorganismos comparativos *Methylibium petroleiphilum* PM1 (1ª línea de cada cuadro) *Clostridium* sp. 7_2_43FAA (2ª línea de cada cuadro), *Talaromyces stipitatus* CACT 10500 (3ª línea de cada cuadro), y *Pseudomonas putida* GB-1 (4ª línea de cada cuadro) muestran cada una importantes diferencias en las secuencias, tanto a nivel de ADN como proteico.

Tabla 2: Comparación de secuencias del dominio IV* con dominios conocidos

40 Bloque superior: So ce56 frente a microorganismos comparativos
Bloque inferior: *A. fasciculatus* frente a microorganismos comparativos

Tipo	herramienta	Dominio IV*
Identidad por pares de ADN [%]	Geneious	55,1
		15,8
		49,4
		53,7

Tipo	herramienta	Dominio IV*
	ClustalW	50,2 35,4 43,4 47,6
Identidad proteica [%]	ClustalW	9 8 8 8
Similitud proteica [%]	ClustalW	14 15 16 13
tipo	herramienta	Dominio IV*
Identidad por pares de ADN [%]	Geneious	53,2 57,1 48,5 52,0
	ClustalW	49,5 33,2 39,4 47,7
Identidad proteica [%]	ClustalW	14 12 16 17
Similitud proteica [%]	ClustalW	23 22 28 26

Las identidades y similitudes más bien bajas del dominio IV* de *Sorangium* y *Aetherobacter* respecto a los dominios funcionalmente similares de organismos conocidos muestran que los genes sintéticos utilizados en la invención difieren significativamente de los genes previamente conocidos.

5 Comparación de los dominios génicos de AGPI mixobacterianos y dominios de proteínas con genes/proteínas de AGPI bacterianos comparativos

10 La comparación de la secuencia de la tabla 3 de los dominios de genes de AGPI de la invención con dominios homólogos de microorganismos conocidos muestra importantes diferencias de secuencia entre los genes de la invención y los genes conocidos, tanto a nivel de ADN como proteico.

Tabla 3: Comparación de secuencias con genes comparativos

tipo	herramienta	Regiones de agrupamientos de AGPI						
		I	Ila	Ilb	Ilc	Illa	IIIb	IV* vs. IV
Identidad por pares de ADN [%]	Geneious	53,6	50,2	55,2	48,1	47,0	50,3	46,8
		53,5	50,0	50,9	46,3	44,6	46,7	45,0
		52,4	51,1	53,7	47,2	46,5	49,5	45,8
	ClustalW	51,8	46,6	53,0	43,4	42,8	45,9	40,1
		50,1	45,9	49,4	41,8	40,4	41,5	37,8
		50,3	47,1	50,3	43,1	42,2	42,6	37,6
Identidad proteica [%]	ClustalW	45	37	41	21	23	31	9
		44	35	39	26	22	28	11
		44	36	41	25	24	31	12
Similitud proteica [%]	ClustalW	65	51	56	34	36	44	23
		63	51	56	40	35	42	25
		64	51	56	38	38	44	26

15

tipo	herramienta	I	Ila	Ilb	Ilc	Illa	IIIb	IV	IV* vs. IV
Identidad por pares de ADN [%]	Geneious	53,1	50,2	58,6	48,5	47,6	49,4	48,2	46,8

tipo	herramienta	I	Ila	IIb	IIc	IIIa	IIIb	IV	IV* vs. IV
		51,8	49,9	54,8	46,9	45,8	47,2	45,4	46,7
		53,6	50,9	57,5	47,2	46,4	49,1	47,6	45,9
	ClustalW	50,5	46,0	54,6	42,9	43,9	45,3	43,4	40,4
		49,1	47,2	52,5	40,9	42,0	41,7	39,7	38,5
		51,7	47,3	54,3	42,1	42,3	45,0	40,9	39,9
Identidad proteica [%]	ClustalW	45	36	26	22	24	23	8	8
		44	36	30	24	23	23	8	8
		44	37	29	24	25	25	9	9
Similitud proteica [%]	ClustalW	64	50	34	35	38	36	16	16
		63	50	41	38	37	36	14	15
		63	51	37	36	39	35	16	18

Bloque superior: So ce56 comparada con los genes de AGPI de *Shewanella* (1ª línea de cada cuadro), *Moritella marina* (2ª línea de cada cuadro), *Photobacterium profundum* (3ª línea de cada cuadro),

5 Bloque inferior: *A. fasciculatus* comparada con los genes de AGPI de *Shewanella* (1ª línea de cada cuadro), *Moritella marina* (2ª línea de cada cuadro), *Photobacterium profundum* (3ª línea de cada cuadro.)

† Los alineamientos se realizaron con el programa Geneious como se especifica en la presente memoria.

‡ Definición de las regiones alineadas:

10 Generalmente, los dominios del agrupamiento de genes de Soce y *Aetherobacter* (SBSr002) son como se describen en la breve descripción del listado de secuencias.

15 Dominio I (comparación de secuencias de ADN): la secuencia completa del gen de pfa1 de Soce56 y pfa1 de *A. fasciculatus* (SBSr002) se alineó contra las secuencias completas de genes de pfaD (= orf8) de *Shewanella sp.* SCRC-2738 (n.º acceso: U73935.1/30.730-32.361 nt), orf11 de *Moritella marina* MP-1 (n.º acceso: AB025342.2/27.119-28.735 nt) y pfaD de *Photobacterium profundum* SS9 (n.º acceso: AF409100.1/23.166-24.800 nt).

20 Dominio I (comparación de secuencias proteicas): las secuencias de proteínas completas de Pfa1 de Soce56 o Pfa1 de *A. fasciculatus* se alinearon contra las secuencias de proteínas completas de PfaD (= orf8) de *Shewanella sp.* SCRC-2738 (n.º acceso: AAB81126.1/1-543 aa), Orf11 de *Moritella marina* MP-1 (n.º acceso: BAA89385.1/1-538 aa) y PfaD de *Photobacterium profundum* SS9 (n.º de acceso: AAL01063.1/1-544 aa).

25 Dominio IIa (comparación de secuencias de ADN): el gen pfa2 de Soce56 y el gen pfa2 de *A. fasciculatus* se alinearon con el gen pfaA (= orf5) de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1), con respecto al gen orf de *Moritella marina* MP-1 (n.º de acceso: AB025342.2) y al gen pfaA de *Photobacterium profundum* SS9 (n.º de acceso: AF409100.1).

30 Dominio IIa (comparación de secuencias proteicas): la proteína Pfa2 de So ce56 y la proteína Pfa2 de *A. fasciculatus* se alinearon contra la proteína PfaA (= orf5) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81123.1), el extremo N-terminal de la proteína Orf8 de *Moritella marina* MP-1 (n.º de acceso: BAA89382.2) y el extremo N-terminal de la proteína PfaA de *Photobacterium profundum* SS9 (n.º de acceso: AAL01060.1).

35 Dominio IIb de (comparación de secuencias de ADN): el gen pfa2 de So ce56 se alineó contra la sección central del gen pfaA (= orf5) de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1), el gen orf8 de *Moritella marina* MP-1 (n.º de acceso: AB025342.2) el gen pfaA de *Photobacterium profundum* SS9 (n.º de acceso: AF409100.1).

40 Dominio IIb (comparación de secuencias proteicas): la proteína Pfa2 de So ce56 se alineó contra la sección central de la proteína PfaA (= orf5) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81123.1/1.244-1.747 aa), la sección central de la proteína Orf8 de *Moritella marina* MP-1 (n.º de acceso: BAA89382.2/1.251-1734 aa) y la sección central de la proteína PfaA a partir de *Photobacterium profundum* SS9 (n.º de acceso: AAL01060.1/1.227-1.700 aa).

45 Dominio IIc (comparación de secuencias de ADN): el gen pfa2 de So ce56 y el gen pfa2 de *A. fasciculatus* se alinearon contra el extremo 3' del gen pfaA (= orf5) a partir de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1/19.147-22.176 nt), el extremo 3' del gen orf8 de *Moritella marina* MP-1 (n.º de acceso: AB025342.2/15.213-17.969 nt) y el extremo 3' del gen pfaA de *Photobacterium profundum* SS9 (n.º de acceso: AF409100.1/12.554-15.175 nt).

50 Dominio IIc (comparación de secuencias proteicas): la proteína Pfa2 de So y la proteína Pfa2 de *A. fasciculatus* se alinearon contra el extremo C-terminal de la proteína PfaA (= orf5) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81123.1/1.748-2.756 aa), el extremo C-terminal de la proteína orf8 de *Moritella marina* MP-1 (n.º de acceso: BAA89382.2/1.735-2.652 aa) y el extremo C-terminal de la proteína PfaA de *Photobacterium profundum* SS9 (n.º de

acceso: AAL01060.1/1.701-2.573 aa).

5 Dominio IIIa (comparación de secuencias de ADN): el gen pfa3 de So ce56 y el extremo 5' del gen pfa3 de *A. fasciculatus* se alinearon contra el extremo 5' del gen de pfaC (= orf7) de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1/24.518-27.868 nt), el extremo 5' del gen orf10 de *Moritella marina* MP-1 (n.º de acceso: AB025342.2/20.847-24.044 nt) y el extremo 5' del gen pfaC de *Photobacterium profundum* SS9 (n.º de acceso: AF409100.1/17.271-20.528 nt).

10 Dominio IIIa (comparación de secuencias proteicas): la proteína Pfa3 de So ce56 y la proteína Pfa3 de *A. fasciculatus* se alinearon contra el extremo N-terminal de la proteína pfaC (= orf7) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81125.1/1-1.117 aa), el extremo N-terminal de la proteína Orf10 de *Moritella marina* MP-1 (n.º de acceso: BAA89384.1/1-1.066 aa) y el extremo N-terminal de la proteína de PfaC de *Photobacterium profundum* SS9 (n.º de acceso: AAL01062.1/1-1.086 aa).

15 Dominio IIIb (comparación de secuencias de ADN): el gen pfa3 de So ce56 y el gen pfa3 de *A. fasciculatus* se alinearon contra el extremo 3' del gen de pfaC (= orf7) de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1/27.869-30.532 nt), el extremo 3' del gen orf10 gen de *Moritella marina* MP-1 (n.º de acceso: AB025342.2/24.045-26.882 nt) y el extremo 3' del gen pfaC de *Photobacterium profundum* SS9 (n.º de acceso: AF409100.1/20.529-23.147 nt).

20 Dominio IIIb (comparación de secuencias proteicas): la proteína Pfa3 de So ce56 y la sección central de la proteína Pfa3 de *A. fasciculatus* se alinearon contra el extremo C-terminal de la proteína PfaC (= orf7) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81125.1/1.118-2.004 aa), el extremo C-terminal de la proteína Orf10 de *Moritella marina* MP-1 (n.º de acceso: BAA89384.1/1.067-2.011 aa) y el extremo C-terminal de la proteína PfaC de *Photobacterium profundum* SS9 (n.º de acceso: AAL01062.1/1.087-1.958 aa).

30 Dominio IV* (comparación de secuencias de ADN): el pfa3 de So ce56 y el extremo 3' del gen pfa3 de *A. fasciculatus* se alinearon contra las secuencias de genes completos de pfaB (= orf6) de *Shewanella sp.* SCRC-2738 (n.º de acceso: U73935.1/22.176-24.518 nt), orf9 de *Moritella marina* MP-1 (n.º de acceso: AB025342.2/18.126-20.723 nt) y pfaB de *Photobacterium Profundum* ES9 (n.º de acceso: AF409100.1/15.175-17.274 nt).

35 Dominio IV* (comparación de secuencias proteicas): la proteína Pfa3 de So ce56 y la proteína Pfa3 de *A. fasciculatus* se alinearon contra las secuencias de proteínas completas de PfaB (= orf6) de *Shewanella sp.* SCRC-2738 (n.º de acceso: AAB81124.1/1-780 aa), de Orf9 *Moritella marina* MP-1 (n.º de acceso: BAA89383.1/1-865 aa) y PfaB de *Photobacterium Profundum* ES9 (n.º de acceso: AAL01061.1/1-699 aa).

La Figura 1 muestra las regiones antes definidas I, IIa, IIb, IIc, IIIa, IV si es aplicable, IIIb y IV*.

40 De nuevo, la comparación de secuencias de dominios de genes de AGPI de So ce56 y *Aetherobacter* con las de *Schizochytrium* muestra la homología de secuencia muy baja de los genes de la invención con respecto a los genes conocidos de la síntesis de AGPI.

Tabla 4: Comparación de secuencias de genes/proteínas de AGPI mixobacterianos con genes/proteínas de AGPI a partir de *Schizochytrium sp.*

45 Bloque superior: So ce56 frente a *Schizochytrium sp.*
Bloque inferior: *A. fasciculatus* frente a *Schizochytrium sp.**

tipo	herramienta	Dominio del agrupamiento de AGPI‡							
		la	lb	IIa	IIb	IIc	IIIa	IIIb	IV* vs. IV
Identidad por pares de ADN [%]	Geneious	59,2	61,4	55,6	60,9	56,7	52,9	54,1	51,0
	ClustalW	55,8	58,8	51,5	57,4	52,6	48,5	50,0	43,3
Identidad proteica [%]	ClustalW	37	38	30	32	26	22	26	10
Similitud proteica [%]	ClustalW	57	57	44	42	39	34	38	20

tipo	herramienta	la	lb	IIa	IIb	IIc	IIIa	IIIb	IV	IV* vs. IV
Identidad por pares de ADN [%]	Geneious	58,3	61,3	55,9	62,8	56,7	52,8	54,9	53,2	51,1
	ClustalW	55,6	59,0	52,2	61,7	51,9	49,7	49,3	50,1	45,5
Identidad proteica [%]	ClustalW	40	41	29	15	26	23	21	16	11
Similitud proteica [%]	ClustalW	57	58	43	21	42	35	31	26	22

50 ‡ Definición de los dominios alineados: Los dominios del agrupamiento de genes de So ce56 y *Aetherobacter* (SBSr002) son como se describen en la breve descripción del listado de secuencias.

Los dominios comparativos de *Schizochytrium* fueron:

- 5 Dominio Ia (comparación de secuencias de ADN): extremo 3' del gen orfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378328.2/4.500-6.180 nt).
- 10 Dominio Ia (comparación de secuencias proteicas): extremo C-terminal de la proteína OrfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72880.2/1500-2059 aa).
- 15 Dominio Ib (comparación de secuencias de ADN): extremo 3' del gen orfC (= subunidadC) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378329.2/2.848-4.509 nt).
- 20 Dominio Ib (comparación de secuencias proteicas): extremo C-terminal de la proteína OrfC (= SubunidadC) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72881.2/950-1.502 aa).
- 25 Dominio IIa (comparación de secuencias de ADN): extremo 5' del gen orfA (= subunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378327.2/1-3.699 nt).
- 30 Dominio IIa (comparación de secuencias proteicas): sección del extremo N-terminal de la proteína OrfA (= SubunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72879.2/1-1.233 aa).
- 35 Dominio IIb (comparación de secuencias de ADN): Gen orfA (= subunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378327.2/3.700-6.198 nt).
- 40 Dominio IIb (comparación de secuencias proteicas): la proteína OrfA (= SubunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72879.2/1.234-2.066 aa).
- 45 Dominio IIc (comparación de secuencias de ADN): extremo 3' del gen orfA (= subunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378327.2/6.199-8.733 nt).
- 50 Dominio IIc (comparación de secuencias proteicas): sección del extremo C-terminal de la proteína OrfA (= subunidadA) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72879.2/2.067-2.910 aa).
- 55 Dominio IIIa (comparación de secuencias de ADN): extremo 5' del gen orfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378328.2/1-3.150 nt).
- 60 Dominio IIIa (comparación de secuencias proteicas): sección del extremo N-terminal de la proteína OrfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72880.2/1-1.050 aa).
- 65 Dominio IIIb (comparación de secuencias de ADN): extremo 5' del gen orfC (= subunidadC) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378329.2/1-2.847 nt).
- 70 Dominio IIIb (comparación de secuencias proteicas): sección del extremo N-terminal de la proteína OrfC (= subunidadC) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72881.2/1-949 aa).
- 75 Dominio IV (comparación de secuencias de ADN): la región central del gen orfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AF378328.2/3.151-4.497 nt) se comparó con el dominio IV* de *Sorangium*, y con los dominios IV y IV* de *Aethero bacter*, respectivamente.
- 80 Dominio IV (comparación de secuencias proteicas): la región central de la proteína OrfB (= subunidadB) de *Schizochytrium sp.* CACT_20888 (número de acceso: AAK72880.2/1.051-1.499 aa) se comparó con el dominio IV* de *Sorangium*, y con los dominios IV y IV* de *Aethero bacter*, respectivamente.
- 85 Una comparación de secuencias entre los genes de AGPI en el nivel de ADN y de proteína muestra importantes homologías de secuencia entre los dominios, genes y agrupamientos de genes de la invención, en especial en el grupo de los dominios, genes y agrupamientos de genes de *Sorangium*, y en el grupo de los dominios, genes y agrupamientos de genes de *Aethero bacter*, respectivamente.
- 90 La Tabla 5 muestra a continuación el resultado de la comparación de secuencias para los dominios de *Aethero bacter* SBSr002 y SBSr008 (bloque superior), para los dominios de *Aethero bacter* SBSr002 y SBSr003 (bloque medio), y para los dominios de *Aethero bacter* SBSr008 y SBSr003 (bloque inferior).
- 95 La Tabla 6 muestra a continuación el resultado de la comparación de secuencias para los dominios de *Sorangium*, Soce56 y Soce377.
- 100 Los dominios utilizados en estas comparaciones son los descritos en el listado de secuencias.

Tabla 5: Comparación de genes/proteínas de AGPI micobacterianos de *Aetherobacter*

tipo	herramienta	Dominios de los agrupamientos de AGPI							
		I	Ila	IIb	IIc	IIIa	IIIb	IV	IV* vs. IV
Identidad por pares de ADN [%]	Geneious	99,0	97,8	95,9	96,3	98,6	96,9	98,3	96,8
	ClustalW	99,0	97,8	95,9	96,3	98,6	96,9	98,3	96,8
Identidad proteica [%]	ClustalW	99	97	99	96	98	96	98	95
Similitud proteica [%]	ClustalW	99	98	99	97	98	97	99	98

tipo	herramienta	I							IV*
Identidad por pares de ADN [%]	Geneious	92,3							85,6
	ClustalW	92,3							85,4
Identidad proteica [%]	ClustalW	94							84
Similitud proteica [%]	ClustalW	96							88
tipo	herramienta	I							IV*
Identidad por pares de ADN [%]	Geneious	92,5							85,9
	ClustalW	92,5							85,5
Identidad proteica [%]	ClustalW	94							84
Similitud proteica [%]	ClustalW	97							89

5 Tabla 6: Comparación de genes/proteínas de AGPI micobacterianos de *Sorangium*

tipo	herramienta	Dominios de los agrupamientos de AGPI						
		I	Ila	IIb	IIc	IIIa	IIIb	IV*
Identidad por pares de ADN [%]	Geneious	98,2	95,8	91,1	97,8	96,1	97,0	96,9
	ClustalW	98,2	95,8	91,1	97,8	95,9	97,0	96,9
Identidad proteica [%]	ClustalW	98	94	90	97	94	95	97
Similitud proteica [%]	ClustalW	99	96	91	97	95	96	98

10 Estos análisis muestran los grados significativos de identidad de los dominios del agrupamiento de genes de AGPI para los grupos de genes de *Sorangium* y para el grupo de genes de *Aetherobacter*, respectivamente, en combinación con un grado significativo de identidad de secuencia encontrado entre *Sorangium* y *Aetherobacter*, especialmente con respecto al dominio IV*.

15 Además, los dominios del agrupamiento de genes sintéticos de AGPI de *Sorangium* y *Aetherobacter* tienen una alta identidad y similitud de proteínas como se muestra en la Tabla 7 a continuación, que es, p. ej., significativamente mayor que la identidad y similitud de proteínas, respectivamente, de *Sorangium* y *Aetherobacter* hacia las regiones y dominios de organismos conocidos pertinentes, p. ej., hacia los dominios funcionalmente similares de *Shewanella*, *Moritella marina*, y *Photobacterium profundum* como se muestra en la tabla 3, y hacia *Schizochytrium* como se muestra en la tabla 4.

20 Tabla 7: Comparación de genes/proteínas de AGPI mixobacterianos de *Sorangium* y *Aetherobacter*

tipo	herramienta	Dominios de los agrupamientos de AGPI						
		I	Ila	IIb	IIc	IIIa	IIIb	IV*
Identidad por pares de ADN [%]	Geneious	70,4	69,8	67,5	65,9	65,3	69,0	63,5
	ClustalW	69,2	68,0	66,1	63,7	62,3	67,1	62,0
Identidad proteica [%]	Geneious	59	58	27	47	46	42	22
	ClustalW	59	57	26	47	47	42	24
Similitud proteica [%]	Geneious	73	67	34	61	56	49	29
	ClustalW	73	68	33	60	56	50	31

25 Por consiguiente, las enzimas de la vía sintética de AGPI de la invención contienen un dominio que tiene actividad AGPAT y una secuencia de aminoácidos que es al menos 50 %, 85 %, 90 % o al menos 95 % idéntica a al menos una secuencia de aminoácidos de un dominio IV* de *Sorangium* y/o de *Aetherobacter* como se describe en la presente memoria.

REVINDICACIONES

1. Proceso de producción de ácidos grasos poliinsaturados por expresión de genes heterólogos de enzimas de la vía biosintética de ácidos grasos poliinsaturados en un organismo de producción, que comprende el cultivo del organismo de producción en presencia de una fuente de carbono fermentable,
- en el que el organismo de producción se manipula genéticamente para comprender un primer gen, un segundo gen y un tercer gen,
 - en el que el primer gen contiene un dominio I que codifica una enoil reductasa que tiene una identidad de secuencia de aminoácidos de al menos el 59 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 13, SEQ ID NO: 44, SEQ ID NO: 75, SEQ ID NO: 109, SEQ ID NO: 133 y/o SEQ ID NO: 149, y
 - en el que el segundo gen contiene un dominio IIa que codifica una cetosintasa y una malonil-CoA-transacilasa que tiene una identidad de secuencia de aminoácidos de al menos el 57 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 16, SEQ ID NO: 47, SEQ ID NO: 78 y/o SEQ ID NO: 112, un dominio IIb que codifica al menos una proteína transportadora de acilos que tiene una identidad de secuencia de aminoácidos de al menos el 26 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 19, SEQ ID NO: 50, SEQ ID NO: 81, SEQ ID NO: 115 y/o SEQ ID NO: 155; un dominio IIc que codifica una cetorreductasa y una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 47 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 22, SEQ ID NO: 53, SEQ ID NO: 84, SEQ ID NO: 118 y/o SEQ ID NO: 158; y
 - en el que el tercer gen contiene un dominio IIIa que codifica una cetosintasa y un factor de longitud de cadena que tiene una identidad de secuencia de aminoácidos de al menos el 46 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 25, SEQ ID NO: 56, SEQ ID NO: 87, SEQ ID NO: 121 y/o SEQ ID NO: 161; y un dominio IIIb que codifica al menos una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 42 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 28, SEQ ID NO: 59, SEQ ID NO: 90, SEQ ID NO: 124 y/o SEQ ID NO: 164, y un dominio que codifica una aciltransferasa (AT) caracterizado por que la aciltransferasa codificada es una acil-glicerol-fosfato aciltransferasa (AGPAT) que tiene una identidad de secuencia de aminoácidos de al menos el 50 % con respecto a una secuencia de aminoácidos codificada por el dominio IV* contenido en el grupo que comprende SEQ ID NO: 30, SEQ ID NO: 61, SEQ ID NO: 95, SEQ ID NO: 129, SEQ ID NO: 135 y/o SEQ ID NO: 169, y secuencias de aminoácidos codificadas por una sección de secuencia de ácidos nucleicos contenida en el grupo que comprende SEQ ID NO: 29, SEQ ID NO: 60, SEQ ID NO: 94, SEQ ID NO: 128, SEQ ID NO: 134 y/o SEQ ID NO: 168.
2. Proceso de acuerdo con la reivindicación 1, caracterizado por que el dominio IV* está contenido en una secuencia de ácidos nucleicos que codifica el producto génico Pfa3 que tiene una identidad de secuencia de aminoácidos de al menos el 28 % con respecto a una secuencia de aminoácidos contenida en el grupo que comprende SEQ ID NO: 10 y/o SEQ ID NO: 41, SEQ ID NO: 72, SEQ ID NO: 106 y/o SEQ ID NO: 146.
3. Proceso de acuerdo con la reivindicación 1 o 2, caracterizado por que el tercer gen contiene un dominio IV que codifica una aciltransferasa (AT) que tiene una identidad de secuencia de aminoácidos de al menos el 98 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 93, SEQ ID NO: 127 y/o SEQ ID NO: 167, cuyo dominio IV se dispone entre el dominio IIIa y el dominio IIIb.
4. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el primer gen codifica una enoil reductasa que tiene una identidad de secuencia de aminoácidos de al menos el 30 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 6, SEQ ID NO: 37, SEQ ID NO: 68, SEQ ID NO: 102, SEQ ID NO: 132 y/o SEQ ID NO: 142.
5. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el segundo gen codifica una cetosintasa, una malonil-CoA-transacilasa, una proteína transportadora de acilos, una cetorreductasa y una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 30 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 8, SEQ ID NO: 39, SEQ ID NO: 70, SEQ ID NO: 104 y/o SEQ ID NO: 144.
6. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el tercer gen codifica una cetosintasa, un factor de longitud de cadena, una deshidratasa, y la acil-glicerol-fosfato aciltransferasa (AGPAT) que tiene una identidad de secuencia de aminoácidos de al menos el 28 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 10, SEQ ID NO: 41, SEQ ID NO: 72, SEQ ID NO: 106 y/o SEQ ID NO: 146.
7. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el organismo de producción se manipula genéticamente para codificar una proteína que tiene una identidad de secuencia de aminoácidos de al menos el 30 % con respecto a una secuencia de aminoácidos codificada por una secuencia de

ADN del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 32, SEQ ID NO: 63, SEQ ID NO: 97, SEQ ID NO: 137.

8. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que la identidad de secuencia de aminoácidos es de al menos el 60 %, de al menos el 85 %, de al menos el 90 % o de al menos el 95 %.

9. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que la fuente de carbono fermentable está libre de ácidos grasos.

10. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el ácido graso poliinsaturado comprende al menos 3 grupos etilénicamente insaturados.

11. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el organismo de producción se selecciona entre el grupo que consiste en eubacterias gramnegativas, eubacterias grampositivas, mixobacterias, hongos y levaduras

12. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que el organismo de producción se selecciona entre el grupo que consiste en *Yarrowia lipolytica* o *Pseudomonas putida*.

13. Proceso de acuerdo con una de las reivindicaciones precedentes, caracterizado por que

- i) un ácido graso poliinsaturado está presente en al menos 40 % en peso basado en el peso total de todos los ácidos grasos poliinsaturados contenidos en la composición y/o
- ii) el ácido graso poliinsaturado es ácido eicosapentanoico (AEP), ácido docosahexanoico (ADH), ácido linoleico y/o ácido linolénico.

14. Uso de un organismo de producción, como se define en una cualquiera de las reivindicaciones 1 a 13, para la producción de ácidos grasos poliinsaturados.

15. Uso de un primer gen, de un segundo gen y de un tercer gen, para construir un organismo de producción, en el que

- el primer gen contiene un dominio I que codifica una enoil reductasa que tiene una identidad de secuencia de aminoácidos de al menos el 59 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 13, SEQ ID NO: 44, SEQ ID NO: 75, SEQ ID NO: 109, SEQ ID NO: 133 y/o SEQ ID NO: 149, y
- el segundo gen contiene un dominio IIa que codifica una cetosintasa y una malonil-CoA-transacilasa que tiene una identidad de secuencia de aminoácidos de al menos el 57 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 16, SEQ ID NO: 47, SEQ ID NO: 78 y/o SEQ ID NO: 112, un dominio IIb que codifica al menos una proteína transportadora de acilos que tiene una identidad de secuencia de aminoácidos de al menos el 26 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 19, SEQ ID NO: 50, SEQ ID NO: 81, SEQ ID NO: 115 y/o SEQ ID NO: 155; y un dominio IIc que codifica una cetorreductasa y una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 47 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 22, SEQ ID NO: 53, SEQ ID NO: 84, SEQ ID NO: 118 y/o SEQ ID NO: 158; y
- el tercer gen contiene un dominio IIIa que codifica una cetosintasa y un factor de longitud de cadena que tiene una identidad de secuencia de aminoácidos de al menos el 46 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 25, SEQ ID NO: 56, SEQ ID NO: 87, SEQ ID NO: 121 y/o SEQ ID NO: 161; y un dominio IIIb que codifica al menos una deshidratasa que tiene una identidad de secuencia de aminoácidos de al menos el 42 % con respecto a una secuencia de aminoácidos del grupo que contiene SEQ ID NO: 28, SEQ ID NO: 59, SEQ ID NO: 90, SEQ ID NO: 124 y/o SEQ ID NO: 164, y un dominio que codifica una aciltransferasa (AT), caracterizado por que la aciltransferasa codificada es una acilglicerol-fosfato aciltransferasa (AGPAT) que tiene una identidad de secuencia de aminoácidos de al menos el 50 % con respecto a una secuencia de aminoácidos codificada por el dominio IV* contenido en el grupo que comprende SEQ ID NO: 30, SEQ ID NO: 61, SEQ ID NO: 95, SEQ ID NO: 129, SEQ ID NO: 135 y/o SEQ ID NO: 169, y secuencias de aminoácidos codificadas por una sección de secuencia de ácidos nucleicos contenida en el grupo que comprende SEQ ID NO: 29, SEQ ID NO: 60, SEQ ID NO: 94, SEQ ID NO: 128, SEQ ID NO: 134 y/o SEQ ID NO: 168.

Figura 1

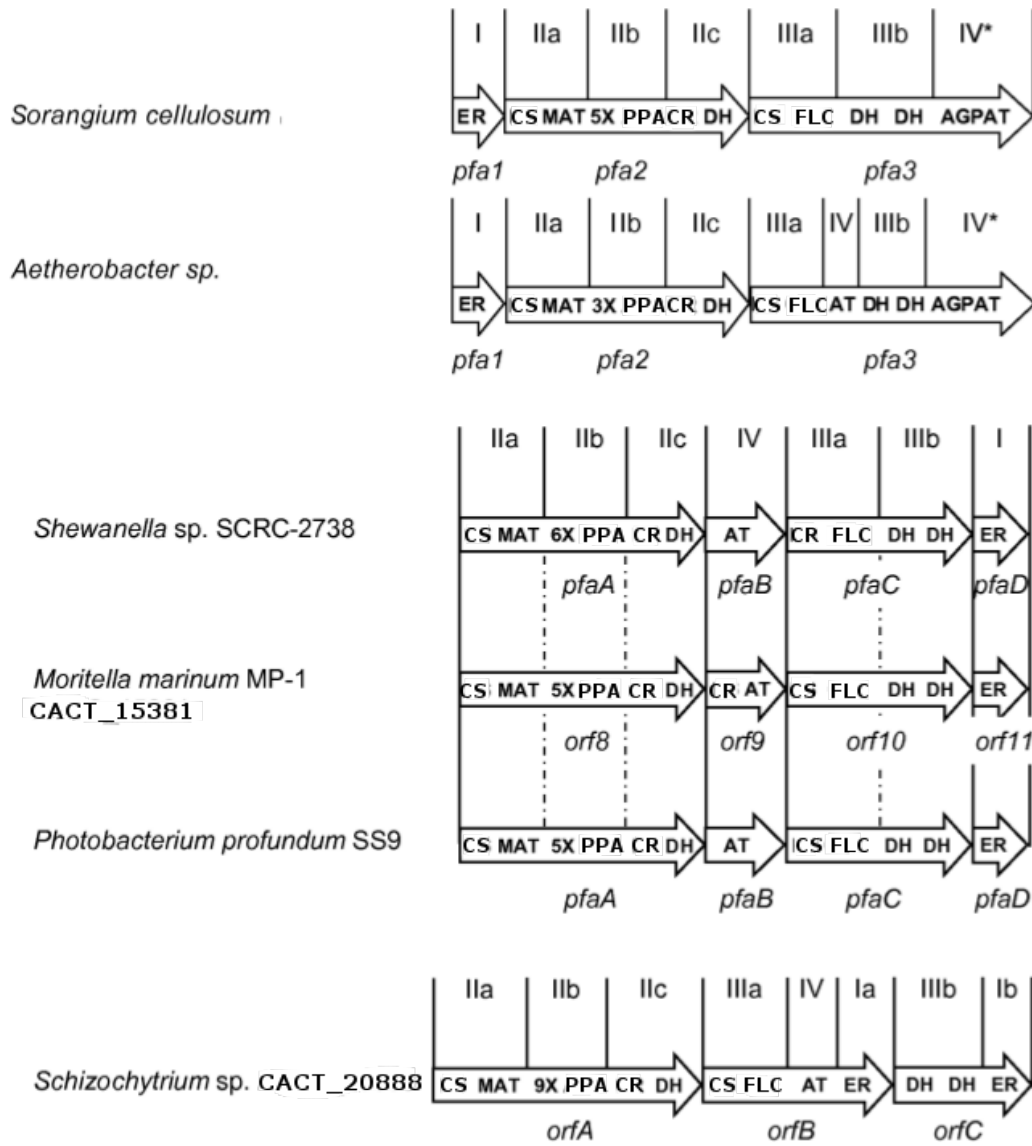


Figura 2

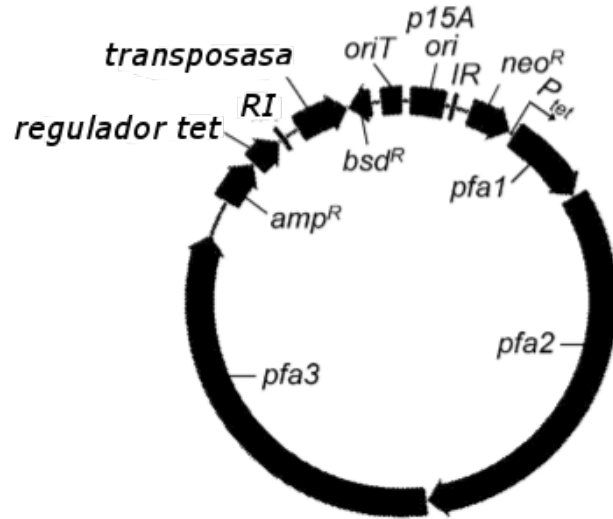


Figura 3

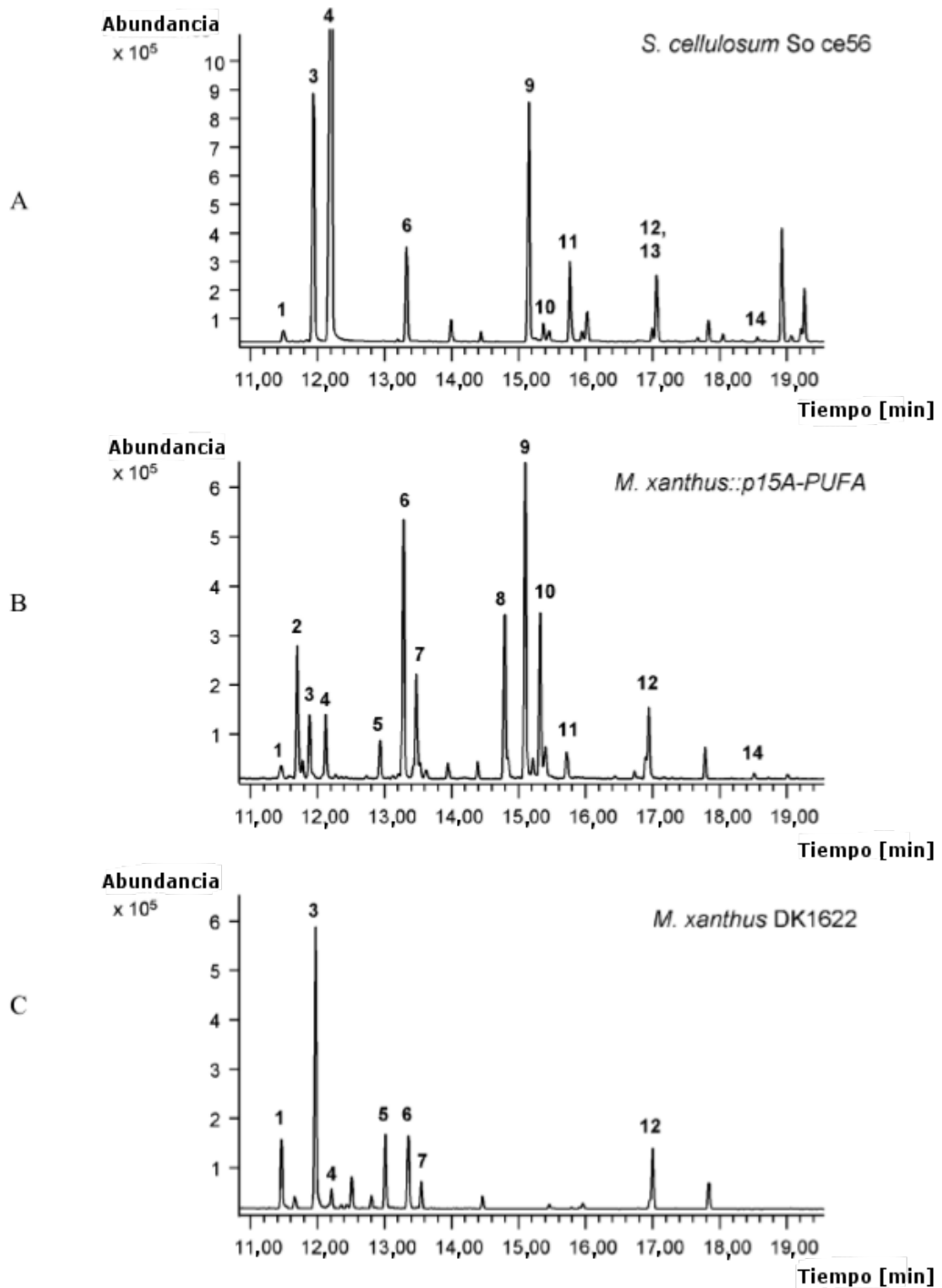


Figura 4

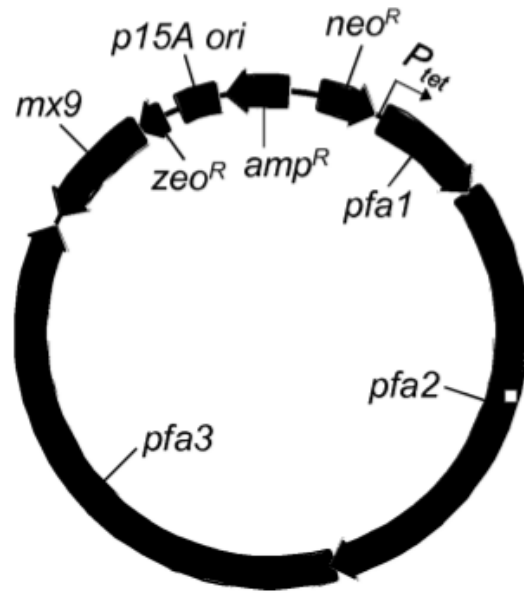


Figura 5

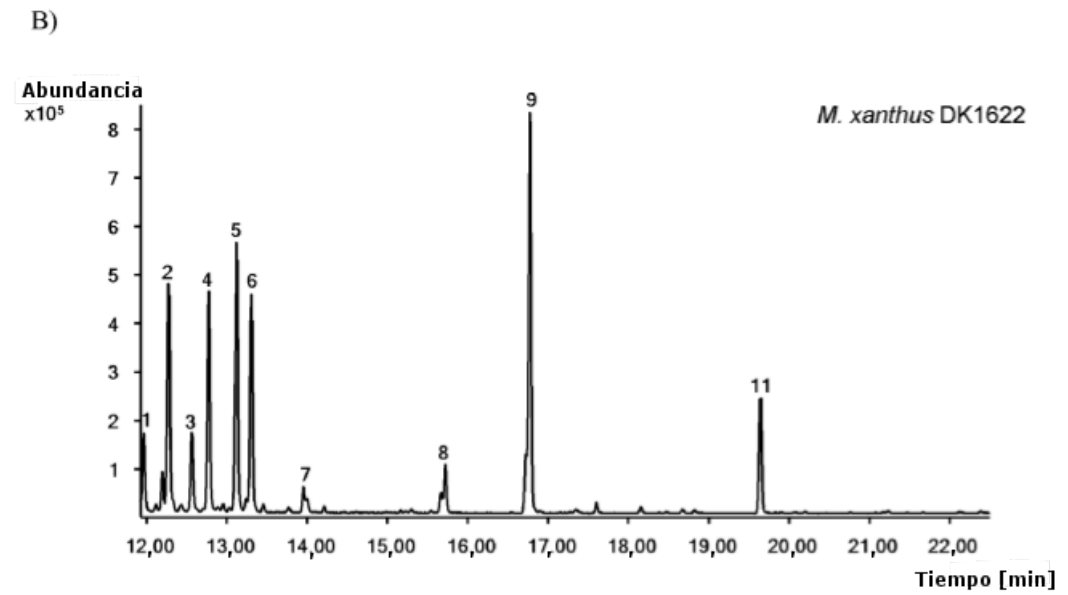
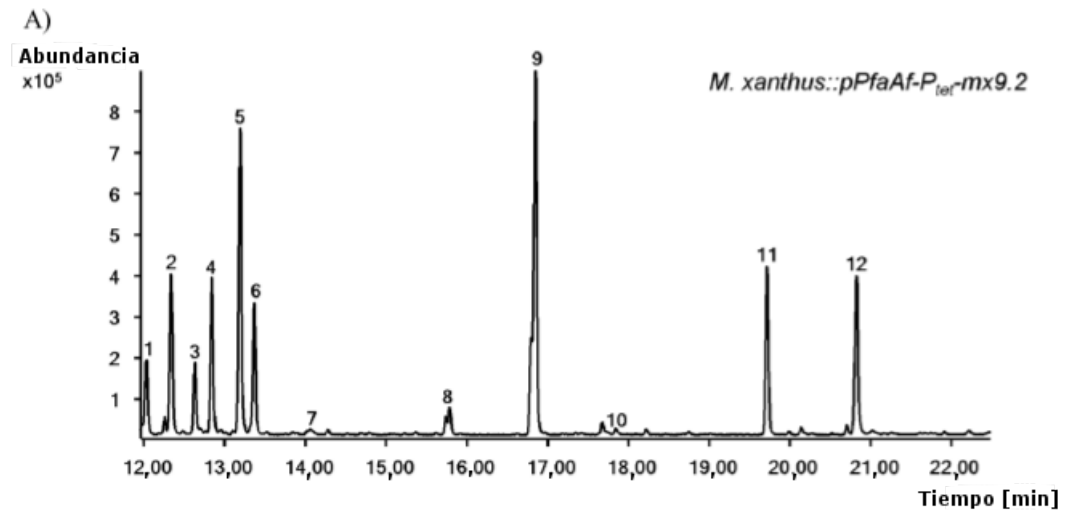
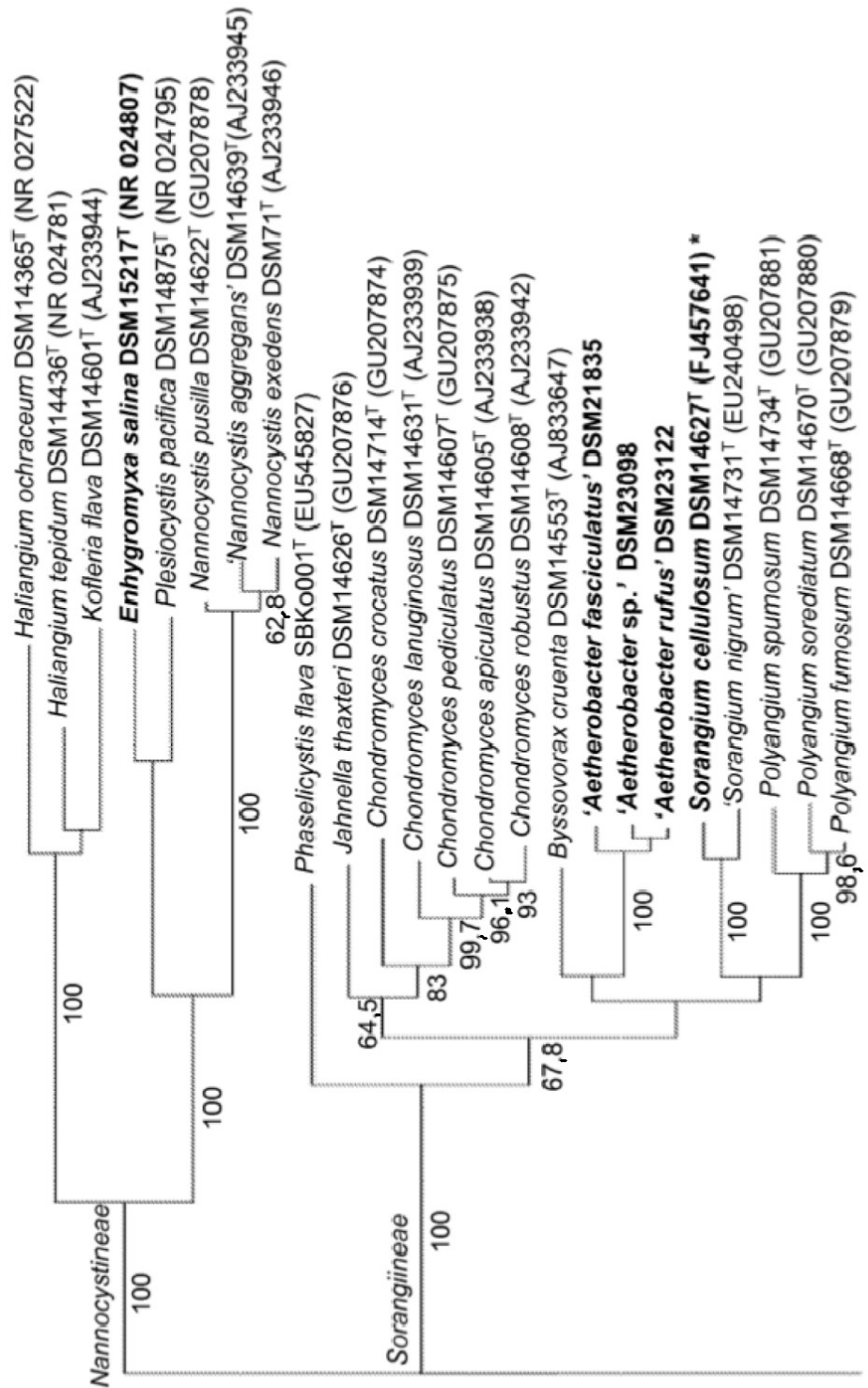


Figura 5



Cont. de la Figura 5

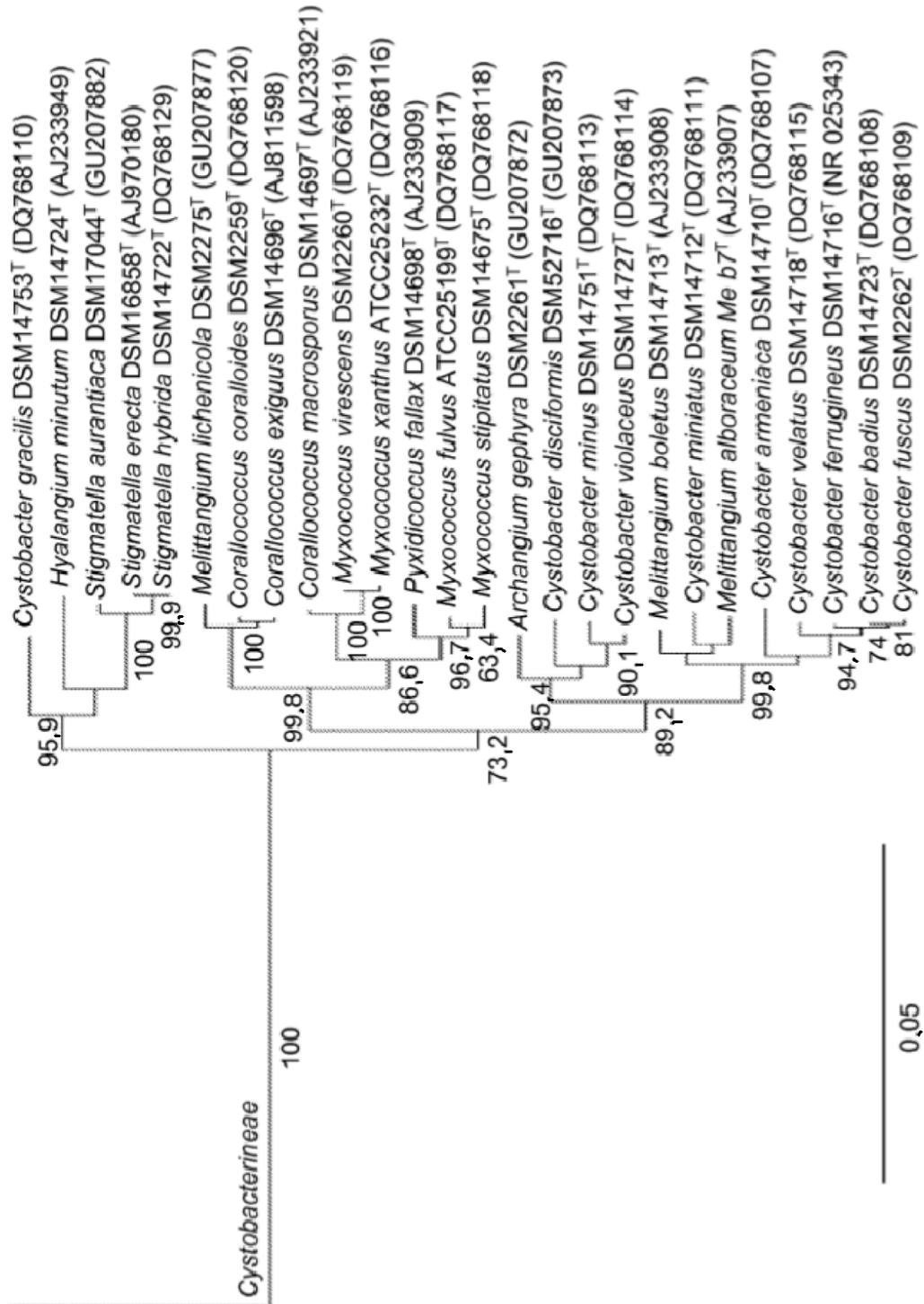
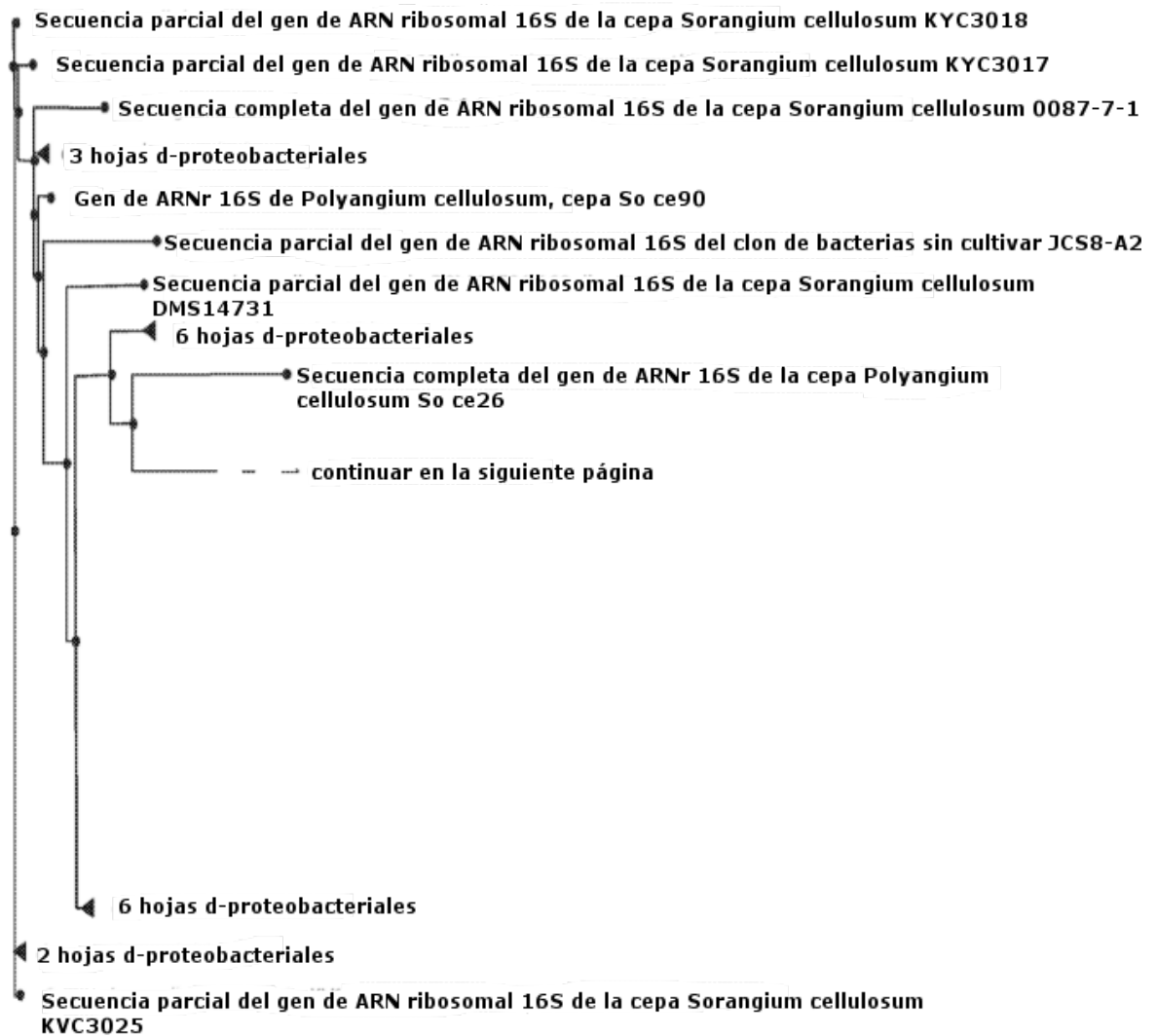


Figura 6



Cont. de la Figura 6

