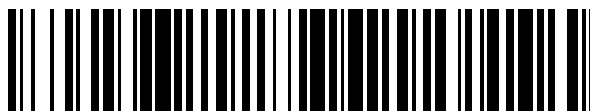


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 410 906**

51 Int. Cl.:

C12N 15/52 (2006.01)

C12N 9/88 (2006.01)

C12P 7/22 (2006.01)

C12P 7/26 (2006.01)

C12P 15/00 (2006.01)

C12P 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.03.2009 E 09717549 (1)**

97 Fecha y número de publicación de la concesión europea: **27.03.2013 EP 2252691**

54 Título: **Procedimiento de producción de alfa-santaleno**

30 Prioridad:

06.03.2008 EP 08102357

03.04.2008 EP 08103362

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

03.07.2013

73 Titular/es:

FIRMENICH S.A. (100.0%)
1, route des Jeunes P.O. Box 239
1211 Geneva 8, CH

72 Inventor/es:

SCHALK, MICHEL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 410 906 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

Procedimiento de producción de alfa-santaleno

Campo técnico

5 La presente invención proporciona un procedimiento de producción de α -santaleno, comprendiendo dicho procedimiento poner en contacto al menos un polipéptido con farnesil pirofosfato (FPP). En particular, dicho procedimiento puede llevarse a cabo *in vitro* o *in vivo* para producir α -santaleno, un compuesto muy útil en los campos de la perfumería y los aromatizantes. La presente invención también proporciona la secuencia de aminoácidos de un polipéptido útil en el procedimiento de la invención. Un ácido nucleico que codifica el polipéptido de la invención y un vector de expresión que contiene dicho ácido nucleico también son parte de la presente invención. Un organismo huésped no humano o una célula transformada para usarse en el procedimiento de producir α -santaleno también son un objeto de la presente invención.

Técnica anterior

15 Los terpenos se encuentran en la mayoría de los organismos (microorganismos, animales y vegetales). Estos compuestos se componen de cinco unidades de carbono llamadas unidades de isopreno y se clasifican por el número de estas unidades presentes en su estructura. Por lo tanto los monoterpenos, sesquiterpenos y diterpenos son terpenos que contienen 10, 15 y 20 átomos de carbono respectivamente. Los sesquiterpenos, por ejemplo, se encuentran ampliamente en el reino vegetal. Muchas moléculas de sesquiterpeno se conocen por sus propiedades saporíferas y de fragancia y sus efectos cosméticos, medicinales y antimicrobianos. Se han identificado más de 300 hidrocarburos de sesquiterpenos y 3000 sesquiterpenoides y se identifican muchas estructuras nuevas cada año. Se usan como fuente de terpenos extractos vegetales obtenidos por diferentes medios tales como destilación por vapor o extracción de disolvente. Las moléculas de terpeno se usan con frecuencia tal cual, pero en algunos casos se usan reacciones químicas para transformar los terpenos en otras moléculas de alto valor.

25 La producción biosintética de terpenos implica enzimas llamadas terpeno sintasas. Hay prácticamente una infinitud de sesquiterpenos sintasas presentes en el reino vegetal, que usan todas el mismo sustrato (farnesil pirofosfato, FPP) pero que tienen diferentes perfiles de producto. Se han clonado genes y ADNc que codifican sesquiterpeno sintasas y se han caracterizado las enzimas recombinantes correspondientes. La biosíntesis de terpenos en plantas y otros organismos se ha estudiado exhaustivamente y no se detalla adicionalmente en el presente documento, pero se hace referencia a Dewick, Nat. Prod. Rep., 2002, 19, 181-222, que revisa el estado de la técnica de las rutas biosintéticas del terpeno.

30 α -santaleno es una molécula de sesquiterpeno de origen natural. El isómero (+) puede usarse como material de partida para la síntesis química o la biosíntesis de (Z)-(+)- α -santalol, que es un constituyente importante del aceite de sándalo. El aceite de sándalo es un ingrediente de perfumería importante obtenido por destilación del duramen de especies de *Santalum*. El sándalo también se usa en gran medida para inciensos y medicina tradicional. El aceite contiene 90% de alcoholes de sesquiterpeno. (Z)-(+)- α -santalol y (Z)-(-)- β -santalol representan los constituyentes principales (respectivamente 45-47% y 20-30%) y son principalmente responsables del olor balsámico y de madera dulce típico del aceite de sándalo. Otros constituyentes tales como epi- β -santalol y trans- α -bergamotol también están presentes y pueden contribuir a la importancia del sándalo.

40 En general, el precio y disponibilidad de extractos naturales vegetales depende de la abundancia, producción de aceite y origen geográfico de las plantas. Además, la disponibilidad y calidad de los extractos naturales depende en gran medida del clima y otras condiciones locales que conducen a variabilidad de un año a otro, haciendo el uso de dichos ingredientes en perfumería de alta calidad muy difícil o incluso imposible algunos años. Debido a la sobreexplotación de los recursos naturales, dificultades de cultivo, crecimiento lento de las plantas de *Santalum*, la disponibilidad de materia prima de sándalo se ha reducido drásticamente durante las últimas décadas. Por lo tanto, sería una ventaja proporcionar una fuente de (Z)-(+)- α -santalol, que esté menos sujeta a fluctuaciones de disponibilidad y calidad. No está disponible hasta la fecha una síntesis química de los constituyentes de sesquiterpeno de sándalo. Sería por lo tanto de gran interés una ruta bioquímica que condujera a la síntesis de (+)- α -santaleno, que podría después usarse para producir (Z)-(+)- α -santalol. Dada la dificultad para controlar la producción de sesquiterpeno en especies de *Santalum*, se han buscado fuentes vegetales alternativas.

50 El sesquiterpeno de tipo santalano, y particularmente sesquiterpenos con el esqueleto de α -santalano, se ha identificado en varias especies vegetales. Se ha indicado que *Clausena lansium*, una planta de la familia Rutaceae contiene grandes cantidades de sesquiterpenos de santalano en las hojas. Zhao y colaboradores (Zhao y col, Z. Naturforsch., 2004, 59c, 153-156) han analizado las hojas de *C. lansium* de China y han detectado la presencia de α -santalol y β -santalol. El análisis de las hojas de *C. lansium* de Cuba ha revelado la presencia de (Z)- α -santalol, epi- β -santalol, (Z)- β -santalol y (E)- β -santalol (Pino y col., J. Essent. Oil Res., 2006, 18, 139-141). Sorprendentemente el análisis de diferentes partes de *C. lansium* de origen tailandés no mostró la presencia de sesquiterpenos con esqueletos de santalano (Chokeprasert y col, Journal of Food Composition and Analysis, 2007, 20(1), 52-56).

- Una sesquiterpeno sintasa capaz de sintetizar al menos un sesquiterpeno bicíclico y/o tricíclico que tiene un esqueleto de carbono de santalano, el ácido nucleico correspondiente y un procedimiento de producción de dichos compuestos que tienen un esqueleto de carbono de santalano, se desvelan en la solicitud de patente Internacional WO 2006/134523. Se citan (+)-epi- β -santaleno, (-)- β -santaleno, (+)- β -santaleno, (+)- α -santaleno y (-)- α -santaleno como ejemplos de compuestos que tienen un esqueleto de carbono de santalano. No obstante, la sesquiterpeno sintasa proporcionada en los ejemplos no produce α -santaleno. Solamente se produce epi- β -santaleno. Las propiedades de este compuesto son muy diferentes de las del α -santaleno. En particular, epi- β -santaleno no es de interés en la síntesis de (Z)-(+)- α -santalol. Además, la sesquiterpeno sintasa desvelada en el documento WO 2006/134523 comparte solamente el 37% de identidad con la secuencia de la invención.
- También se han encontrado terpenos sintasas que tienen un cierto porcentaje de identidad de secuencia con la secuencia de la α -santaleno sintasa de la presente invención en las bases de datos de secuencias. No obstante, el porcentaje de identidad entre las sesquiterpeno sintasas conocidas y el polipéptido de la invención es muy bajo. La secuencia proteica más cercana a la (+)- α -santaleno sintasa de la invención es una (E)- β -farneseno sintasa de *Citrus junos* (Nº de referencia del NCBI AAK54279; Maruyama y col, Biol. Pharm. Bull., 2001, 24(10), 1171-1175) que comparte del 67 al 68% de identidad de secuencia de aminoácidos con la α -santaleno sintasa de la invención.
- Además de la diferencia entre las secuencias en sí mismas, también debe indicarse que la estructura y las propiedades de los productos sintetizados por la enzima anteriormente mencionada son muy diferentes de las de α -santaleno. En particular (E)- β -farneseno no es adecuado como un material de partida para la síntesis de (Z)-(+)- α -santalol, que es un ingrediente muy útil en el campo de la perfumería.
- Se desvela una α -santaleno sintasa en el documento WO 2008/142318. Este documento no se ha publicado a la fecha de la presente solicitud de prioridad. Describe una enzima capaz de catalizar la transformación de Z,Z-farnesil pirofosfato a α -santaleno. Por lo tanto la reacción catalizada por la enzima de la técnica anterior es diferente de la catalizada por la sintasa de la presente invención, que comienza a partir de E,E-farnesil pirofosfato. Además, la α -santaleno sintasa de la invención comparte solamente el 23,8% de identidad de secuencia con la descrita en el documento WO 2008/142318.
- A pesar de estudios exhaustivos de la ciclación de terpenos, el aislamiento y caracterización de las terpeno sintasas es aún difícil, particularmente en plantas, debido a su baja abundancia, sus patrones de expresión con frecuencia transitorios, y la complejidad de purificarlas a partir de las mezclas de resinas y compuestos fenólicos en tejidos en los que se expresan.
- Es un objetivo de la presente invención proporcionar procedimientos para preparar (+)- α -santaleno de una manera económica, como se ha indicado anteriormente. En consecuencia, la presente invención tiene el objetivo de producir (+)- α -santaleno teniendo a la vez pocos residuos, un procedimiento más eficaz con respecto a energía y recursos y reduciendo a la vez la dependencia de combustibles fósiles. Es un objetivo adicional proporcionar enzimas capaces de sintetizar α -santaleno, que sea útil como ingrediente de perfumería y/o aroma.

Abreviaturas Usadas

pb	par de bases
kb	kilobase
BSA	albúmina de suero bovino
DMAPP	dimetilalil difosfato
ADN	ácido desoxirribonucleico
ADNc	ADN complementario
dT	desoxi timina
dNTP	desoxi nucleótido trifosfato
DTT	ditiotreitól
FPP	farnesil pirofosfato
CG	cromatografía gaseosa
idi	isopentenil difosfato isomerasa
IPP	isopentenil difosfato
IPTG	isopropil-D-tiogalacto-piranósido
LB	caldo de lisogenia
MOPSO	ácido 3-(N-morfolino)-2-hidroxiopropanosulfónico
EM	espectrometría de masas
mvaK1	mevalonato quinasa
mvaK2	mevalonato difosfato quinasa
RMN	resonancia magnética nuclear
PCR	reacción en cadena de la polimerasa

RMCE	Intercambio de casetes mediado por recombinasa
3'-/5'-RACE	amplificación rápida de extremos 3' y 5' de ADNc
ARN	ácido ribonucleico
ARNm	ácido ribonucleico mensajero

Descripción de la invención

La presente invención proporciona un procedimiento para producir de forma biosintética α -santaleno de una manera económica, fiable y reproducible.

5 Una "sesquiterpeno sintasa" o un "polipéptido que tiene una actividad sesquiterpeno sintasa", se entiende en el presente documento como un polipéptido capaz de catalizar la síntesis de una molécula de sesquiterpeno o de una mezcla de moléculas de sesquiterpeno del precursor de terpeno acíclico FPP.

10 Como una " α -santaleno sintasa" o como un "polipéptido que tiene una actividad α -santaleno sintasa", los inventores entienden en el presente documento un polipéptido capaz de catalizar la síntesis de α -santaleno, en forma de cualquiera de sus estereoisómeros o una mezcla de los mismos, comenzando a partir de FPP. α -santaleno puede ser el único producto o puede ser parte de una mezcla de sesquiterpenos.

Como una "(+)- α -santaleno sintasa" o como un "polipéptido que tiene una actividad (+)- α -santaleno sintasa" los inventores entienden en el presente documento capaz de catalizar la síntesis de (+)- α -santaleno comenzando a partir de FPP. (+)- α -santaleno puede ser el único producto o puede ser parte de una mezcla de sesquiterpenos. La (+)- α -santaleno sintasa es un ejemplo particular de α -santaleno sintasa.

15 La capacidad de un polipéptido para catalizar la síntesis de un sesquiterpeno particular (por ejemplo (+)- α -santaleno) puede confirmarse de forma sencilla realizando el ensayo enzimático como se detalla en el Ejemplo 4.

De acuerdo con una realización preferida de la invención, FPP está en la forma de (2E,6E)-FPP.

20 De acuerdo con la presente invención, se pretende que los polipéptidos también incluyan polipéptido truncados siempre que mantengan su actividad sesquiterpeno sintasa como se define en cualquiera de las realizaciones anteriores y que compartan al menos el porcentaje de identidad definido con el correspondiente fragmento de SEC ID N°: 1.

25 Como se pretende posteriormente en el presente documento, "una secuencia de nucleótidos obtenida modificando SEC ID N°: 2" abarca cualquier secuencia que se haya obtenido cambiando la secuencia de SEC ID N°: 2 usando cualquier procedimiento conocido en la técnica, por ejemplo, introduciendo cualquier tipo de mutación tal como mutaciones de delección, inserción o sustitución. Se citan ejemplos de dichos procedimientos en la parte de la descripción relativa a los polipéptidos variantes y los procedimientos para prepararlos.

30 El porcentaje de identidad entre dos secuencias peptídicas o nucleotídicas está en función del número de aminoácidos o restos de nucleótidos que son idénticos en las dos secuencias cuando se ha generado un alineamiento de estas dos secuencias. Los restos idénticos se definen como restos que son iguales en las dos secuencias en una posición dada del alineamiento. El porcentaje de identidad de secuencia, como se usa en el presente documento, se calcula a partir del alineamiento óptimo tomando el número de restos idénticos entre dos secuencias, dividiéndolo por el número total de restos en la secuencia más corta y multiplicando por 100. El alineamiento óptimo es el alineamiento en el que el porcentaje de identidad es el mayor posible. Pueden introducirse huecos en una o más secuencias en una o más posiciones del alineamiento para obtener el alineamiento óptimo.

35 Estos huecos se tienen en cuenta después como restos no idénticos para el cálculo del porcentaje de identidad de secuencia.

40 Puede conseguirse alineamiento con el fin de determinar el porcentaje de identidad de secuencia de aminoácidos o ácido nucleico de diversas maneras usando programas de ordenador y por ejemplo programas de ordenador públicamente disponibles en Internet. Preferentemente, el programa BLAST (Tatiana y col, FEMS Microbiol Lett., 1999, 174: 247-250, 1999) ajustado a los parámetros por defecto, disponible del Centro Nacional para la Información Biotecnológica (NCBI) en <http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi>, puede usarse para obtener un alineamiento óptimo de secuencias peptídicas o nucleotídicas y para calcular el porcentaje de identidad de secuencia.

Un objeto de la presente invención es por lo tanto un procedimiento de producción de α -santaleno que comprende

- 45 a) poner en contacto FPP con al menos un polipéptido que tenga una actividad α -santaleno sintasa y que comprende una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1;
- b) opcionalmente, aislar el α -santaleno producido en la etapa a).

De acuerdo con una realización preferida, el procedimiento es un procedimiento de producción de α -santaleno como un producto principal. De acuerdo con una realización aún más preferida, el α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% del producto

producido por el procedimiento de la invención.

De acuerdo con una realización más preferida, el procedimiento es un procedimiento de producción de (+)- α -santaleno y el polipéptido que tiene una actividad α -santaleno sintasa tiene una actividad (+)- α -santaleno sintasa.

5 De acuerdo con una realización aún más preferida, el procedimiento es un procedimiento de producción de (+)- α -santaleno como un producto principal. De acuerdo con una realización más preferida, (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% de los productos producidos por el procedimiento de la invención.

El procedimiento puede llevarse a cabo *in vitro* así como *in vivo*, como se explicará en detalle posteriormente.

10 El polipéptido para poner en contacto con FPP *in vitro* puede obtenerse mediante extracción de cualquier organismo que lo exprese, usando tecnologías de extracción de enzimas o proteínas convencional. Si el organismo huésped es un organismo unicelular o célula que libera el polipéptido de la invención al medio de cultivo, el polipéptido puede simplemente recogerse del medio de cultivo, por ejemplo mediante centrifugación, opcionalmente seguido de etapas de lavado y resuspensión en soluciones de tampón adecuadas. Si el organismo o célula acumula el polipéptido dentro de sus células, el polipéptido puede obtenerse por rotura o lisis de las células y extracción adicional del polipéptido del lisado celular.

15 El polipéptido que tiene una actividad α -santaleno sintasa, en una forma aislada o junto con otras proteínas, por ejemplo en un extracto proteico en bruto obtenido de células o microorganismos cultivados, puede después suspenderse en una solución de tampón a pH óptimo. Si es adecuado, pueden añadirse sales, BSA y otros tipos de cofactores enzimáticos para optimizar la actividad enzimática. Se describen condiciones apropiadas en más detalle en los Ejemplos más adelante.

20 El FPP precursor puede después añadirse a la suspensión o solución, que después se incuba a temperatura óptima, por ejemplo entre 15 y 40 °C, preferentemente entre 25 y 35 °C, más preferentemente a 30 °C. Después de la incubación, el α -santaleno producido puede aislarse de la solución incubada por procedimientos de aislamiento convencionales, tales como extracción de disolvente y destilación, opcionalmente después de la retirada de los polipéptidos de la solución.

25 De acuerdo con otra realización preferida, el procedimiento de cualquiera de las realizaciones anteriormente descritas se lleva a cabo *in vivo*. En este caso, la etapa a) comprende cultivar un organismo o célula huésped no humano capaz de producir FPP y transformado para expresar al menos un polipéptido que comprenda una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1 y que tenga una actividad α -santaleno sintasa, en condiciones que conduzcan a la producción de α -santaleno.

30 De acuerdo con una realización más preferida, el procedimiento comprende además, antes de la etapa a), transformar un organismo o célula no humano capaz de producir FPP con al menos un ácido nucleico que codifique un polipéptido que comprenda una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1 y que tenga una actividad α -santaleno sintasa, de modo que dicho organismo exprese dicho polipéptido.

35 Estas realizaciones de la invención son particularmente ventajosas puesto que es posible llevar a cabo el procedimiento *in vivo* sin aislar previamente el polipéptido. La reacción se produce directamente dentro del organismo o célula transformada para expresar dicho polipéptido.

40 De acuerdo con una realización particular de la invención, el al menos un ácido nucleico que codifica la α -santaleno sintasa comprende una secuencia de nucleótidos al menos 50%, preferentemente al menos 55%, preferentemente al menos 60%, preferentemente al menos 65%, preferentemente al menos 70%, preferentemente al menos 75%, preferentemente al menos 80%, preferentemente al menos 85%, preferentemente al menos 90%, más preferentemente al menos 95% y aún más preferentemente al menos 98% idéntica a SEC ID N°: 2 o el complemento de la misma. De acuerdo con una realización más preferida, dicho ácido nucleico comprende la secuencia de nucleótidos SEC ID N°: 2 o el complemento de la misma. En una realización aún más preferida, dicho ácido nucleico

45 consiste en SEC ID N°: 2 o el complemento de la misma.

De acuerdo con una realización más preferida, el al menos un ácido nucleico usado en cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2. De acuerdo con una realización aún más preferida, dicho al menos un ácido nucleico consiste en una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2.

50 De acuerdo con otra realización, el al menos un ácido nucleico se aísla de *Clausena lansium*.

55 Se pretende que el organismo o célula “expresen” un polipéptido, siempre que el organismo o célula se transforme para albergar un ácido nucleico que codifique dicho polipéptido, este ácido nucleico se transcriba a ARNm y el polipéptido se encuentre en el organismo o célula huésped. El término “expresar” abarca “expresar de forma heteróloga” y “sobrexpresar”, refiriéndose este último a niveles de ARNm, polipéptidos y/o actividad enzimática por encima de la que se mide en un organismo o célula no transformado. Una descripción más detallada de

procedimientos adecuados para transformar un organismo o célula huésped no humano se describirá posteriormente o en la parte de la memoria descriptiva que está dedicada a dichos organismos o células huésped no humanos transformados como objetos específicos de la presente invención y en los ejemplos.

5 Se entiende que un organismo o célula particular es “capaz de producir FPP” cuando produce FPP de forma natural o cuando no produce FPP de forma natural pero se transforma para producir FPP, antes de la transformación con un ácido nucleico como se ha descrito en el presente documento o junto con dicho ácido nucleico. Los organismos o células transformados para producir una cantidad mayor de FPP que el organismo o célula de origen natural también están abarcados por los “organismos o células capaces de producir FPP”. Ya se conocen en la técnica procedimientos para transformar organismos, por ejemplo microorganismos, de modo que produzcan FPP. Dichos procedimientos pueden por ejemplo hallarse en la bibliografía, por ejemplo en las siguientes publicaciones Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D., y Keasling, J. D. *Nat Biotechnol.*, 2003, 21(7), 796-802 (transformación de *E. coli*); Wu, S., Schalk, M., Clark, A., Miles, R. B., Coates, R., y Chappell, J., *Nat Biotechnol.*, 2006, 24 (11), 1441-1447 (transformación de plantas); Takahashi, S., Yeo, Y., Greenhagen, B. T., McMullin, T., Song, L., Maurina-Brunker, J., Rosson, R., Noel, J., Chappell, J, *Biotechnology and Bioengineering*, 2007, 97(1), 170-181 (transformación de levadura).

15 Para llevar a cabo la invención *in vivo*, el organismo o célula huésped se cultiva en condiciones que conduzcan a la producción de α -santaleno. En consecuencia, si el huésped es una planta transgénica, se proporcionan condiciones de crecimiento óptimas, tales como condiciones de luz, agua y nutrientes óptimas, por ejemplo. Si el huésped es un organismo unicelular, las condiciones que conducen a la producción de α -santaleno pueden comprender adición de cofactores adecuados al medio de cultivo del huésped. Además, puede seleccionarse un medio de cultivo, para maximizar la síntesis de α -santaleno. Se describen condiciones de cultivo óptimas de una manera más detallada en los siguientes Ejemplos.

20 Los organismos huésped no humanos adecuados para llevar a cabo el procedimiento de la invención *in vivo* pueden ser cualquier organismo multicelular o unicelular no humano. En una realización preferida, el organismo huésped no humano usado para llevar a cabo la invención *in vivo* es una planta, un procarionte o un hongo. Puede usarse cualquier planta, procarionte u hongo. Son plantas particularmente útiles las que producen de forma natural grandes cantidades de terpenos. En una realización más preferida, la planta se selecciona de la familia de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Madvaceae*, *Asteraceae* o *Lamiaceae*. Por ejemplo, la planta se selecciona de los géneros *Nicotiana*, *Solanum*, *Sorghum*, *Arabidopsis*, *Brassica* (colza), *Medicago* (alfalfa), *Gossypium* (algodón), *Artemisia*, *Salvia* y *Mentha*. Preferentemente, la planta pertenece a la especie de *Nicotiana tabacum*.

25 En una realización más preferida el organismo huésped no humano usado para llevar a cabo el procedimiento de la invención *in vivo* es un microorganismo. Puede usarse cualquier microorganismo pero de acuerdo con una realización aún más preferida dicho microorganismo es una bacteria o levadura. Más preferentemente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

30 Algunos de estos organismos no producen FPP de forma natural. Para que sea adecuado llevar a cabo el procedimiento de la invención, estos organismos tienen que transformarse para producir dicho precursor. Pueden transformarse antes de la modificación con el ácido nucleico descrito de acuerdo con cualquiera de las realizaciones anteriores o simultáneamente, como se ha explicado anteriormente.

35 También pueden usarse células eucariotas superiores aisladas, en lugar de organismos completos, como huéspedes para llevar a cabo el procedimiento de la invención *in vivo*. Las células eucariotas adecuadas pueden ser cualquier célula no humana, pero son preferentemente células vegetales o fúngicas.

40 De acuerdo con una realización preferida, el al menos un polipéptido que tiene una actividad α -santaleno sintasa usado en cualquiera de las realizaciones anteriormente descritas o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas comprende una secuencia de aminoácidos al menos 55%, preferentemente al menos 60%, preferentemente al menos 65%, preferentemente al menos 70%, preferentemente al menos 75%, preferentemente al menos 80%, preferentemente al menos 85%, preferentemente al menos 90%, más preferentemente al menos 95% y aún más preferentemente al menos 98% idéntica a SEC ID N°: 1. De acuerdo con una realización más preferida, dicho polipéptido comprende la secuencia de aminoácidos SEC ID N°: 1. En una realización aún más preferida, dicho polipéptido consiste en SEC ID N°: 1.

45 De acuerdo con otra realización preferida, el al menos un polipéptido que tiene una actividad α -santaleno sintasa usado en cualquiera de las realizaciones anteriormente descritas o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas comprende una secuencia de aminoácidos que es una variante de SEC ID N°: 1 obtenida por ingeniería genética. En otros términos, dicho polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2. De acuerdo con una realización más preferida, el al menos un polipéptido que tiene una actividad α -santaleno sintasa usado en cualquiera de las realizaciones anteriormente descritas o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas consiste en una secuencia de aminoácidos que es una variante de SEC ID N°: 1 obtenida por ingeniería genética, es decir una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2.

Como se usa en el presente documento, se pretende que el polipéptido sea un fragmento polipeptídico o peptídico que abarque las secuencias de aminoácidos identificadas en el presente documento, así como polipéptidos truncados o variantes, siempre que mantengan su actividad como se ha definido anteriormente y que compartan al menos el porcentaje definido de identidad con el fragmento correspondiente de SEC ID N°: 1.

- 5 Son ejemplos de polipéptidos variantes, proteínas de origen natural que resultan de acontecimientos de corte y empalme de ARNm alternativos o de escisión proteolítica de los polipéptidos descritos en el presente documento. Las variaciones atribuibles a la proteólisis incluyen, por ejemplo diferencias en los extremos N o C tras la expresión en diferentes tipos de células huésped, debido a la retirada proteolítica de uno o más aminoácidos terminales de los polipéptidos de la invención. Los polipéptidos codificados por un ácido nucleico obtenido por mutación natural o
10 artificial de un ácido nucleico de la invención, como se describen posteriormente, también están abarcados por la invención.

También pueden usarse en los procedimientos de la invención variantes polipeptídicas que resulten de una fusión de secuencias peptídicas adicionales en los extremos amino y carboxilo terminal. En particular dicha fusión puede potenciar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente o sistema de expresión deseado. Dichas secuencias peptídicas adicionales pueden ser péptidos señal, por ejemplo. En consecuencia, la presente invención abarca procedimientos que usan polipéptidos variantes, tales como los obtenidos por fusión con otros oligo o polipéptidos y/o los que están unidos a péptidos señal. También pueden usarse provechosamente en los procedimientos de la invención polipéptidos que resulten de una fusión con otra proteína funcional, tal como otra proteína de la ruta de biosíntesis de terpenos.

- 20 De acuerdo con otra realización, el al menos un polipéptido que tiene una actividad α -santaleno sintasa usado en cualquiera de las realizaciones anteriormente descritas o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas se aísla de *Clausena lansium*.

Una herramienta importante para llevar a cabo el procedimiento de la invención es el polipéptido en sí mismo. Un polipéptido que tenga una actividad α -santaleno sintasa y que comprenda una secuencia de aminoácidos al menos
25 50% idéntica a SEC ID N°: 1 es por lo tanto otro objeto de la presente invención.

De acuerdo con una realización preferida, el polipéptido es capaz de producir α -santaleno como un producto principal. De acuerdo con una realización aún más preferida, es capaz de producir una mezcla de sesquiterpenos en la que α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% de los sesquiterpenos producidos.

- 30 De acuerdo con una realización más preferida, el polipéptido tiene una actividad (+)- α -santaleno sintasa.

De acuerdo con una realización aún más preferida, el polipéptido es capaz de producir (+)- α -santaleno como un producto importante. De acuerdo con una realización aún más preferida, es capaz de producir una mezcla de sesquiterpenos en la que (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% de los sesquiterpenos producidos.

- 35 De acuerdo con una realización preferida, el polipéptido comprende una secuencia de aminoácidos al menos 55%, preferentemente al menos 60%, preferentemente al menos 65%, preferentemente al menos 70%, preferentemente al menos 75%, preferentemente al menos 80%, preferentemente al menos 85%, preferentemente al menos 90%, más preferentemente al menos 95% y aún más preferentemente al menos 98% idéntica a SEC ID N°: 1. De acuerdo con una realización más preferida, el polipéptido comprende la secuencia de aminoácidos SEC ID N°: 1. De acuerdo con una
40 realización aún más preferida, el polipéptido consiste en SEC ID N°: 1.

De acuerdo con otra realización preferida, el al menos un polipéptido comprende una secuencia de aminoácidos que es una variante de SEC ID N°: obtenida por ingeniería genética. En otros términos, dicho polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2.

- 45 De acuerdo con una realización más preferida, el al menos un polipéptido que tiene una actividad α -santaleno sintasa consiste en una secuencia de aminoácidos que es una variante de SEC ID N°: 1 obtenida por ingeniería genética, es decir una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2.

De acuerdo con otra realización, el polipéptido se aísla de *Clausena lansium*.

- 50 Como se usa en el presente documento, se pretende que el polipéptido sea un fragmento polipeptídico o peptídico que abarque las secuencias de aminoácidos identificadas en el presente documento, así como polipéptidos truncados o variantes, siempre que mantengan su actividad como se ha definido anteriormente y que compartan al menos el porcentaje de identidad definido con el fragmento correspondiente de SEC ID N°: 1.

Son ejemplos de polipéptidos variantes proteínas de origen natural que resultan de acontecimientos de corte y empalme de ARNm alternativos o de escisión proteolítica de los polipéptidos descritos en el presente documento. Las variaciones atribuibles a la proteólisis incluyen, por ejemplo, diferencias en los extremos N o C tras la expresión en diferentes tipos de células huésped, debido a la retirada proteolítica de uno o más aminoácidos terminales de los
55

polipéptidos de la invención. Los polipéptidos codificados por un ácido nucleico obtenido por mutación natural o artificial de un ácido nucleico de la invención, como se describe posteriormente, también están abarcados por la invención.

5 Las variantes polipeptídicas resultantes de una fusión de secuencias peptídicas adicionales en los extremos amino y carboxilo terminales también están abarcadas por los polipéptidos de la invención. En particular dicha fusión puede potenciar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente o sistema de expresión deseado. Dichas secuencias peptídicas adicionales pueden ser péptidos señal, por ejemplo. En consecuencia, la presente invención abarca variantes de los polipéptidos de la invención, tales como los obtenidos mediante fusión con otros oligo o polipéptidos y/o los que están unidos a péptidos señal. Los polipéptidos resultantes de una fusión con otra proteína funcional, tal como otra proteína de la ruta de biosíntesis de terpeno también están abarcados por los polipéptidos de la invención.

Como se ha mencionado anteriormente, el ácido nucleico que codifica el polipéptido de la invención es una herramienta útil para modificar organismos o células huésped no humanos que se pretenden usar cuando el procedimiento se lleva a cabo *in vivo*.

15 Un ácido nucleico que codifica un polipéptido de acuerdo con cualquiera de las realizaciones anteriormente descritas es por lo tanto también un objeto de la presente invención.

De acuerdo con una realización preferida, el ácido nucleico comprende una secuencia de nucleótidos al menos 50%, preferentemente al menos 55%, preferentemente al menos 60%, preferentemente al menos 65%, preferentemente al menos 70%, preferentemente al menos 75%, preferentemente al menos 80%, preferentemente al menos 85%, preferentemente al menos 90%, más preferentemente al menos 95% y aún más preferentemente al menos 98% idéntica a SEC ID N°: 2 o el complemento de la misma. De acuerdo con una realización más preferida, el ácido nucleico comprende la secuencia de nucleótidos SEC ID N°: 2 o el complemento de la misma. De acuerdo con una realización aún más preferida, el ácido nucleico consiste en SEC ID N°: 2 o el complemento de la misma.

De acuerdo con otra realización, el ácido nucleico se aísla de *Clausena lansium*.

25 Puede definirse que el ácido nucleico de la invención incluye polímeros desoxirribonucleotídicos o ribonucleotídicos en forma mono o bicatenaria (ADN y/o ARN). También debería entenderse que la expresión "secuencia de nucleótidos" comprende una molécula polinucleotídica o una molécula oligonucleotídica en forma de un fragmento separado o como un componente de un ácido nucleico mayor. Los ácidos nucleicos de la invención también abarcan ciertas secuencias de nucleótidos aisladas incluyendo las que están sustancialmente sin material endógeno contaminante. El ácido nucleico de la invención puede estar truncado, siempre que codifique un polipéptido abarcado por la presente invención, como se ha descrito anteriormente.

De acuerdo con una realización más preferida, el al menos un ácido nucleico de acuerdo con cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2. Preferentemente dicho ácido nucleico consiste en una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2.

Los ácidos nucleicos que comprenden una secuencia obtenida por mutación de SEC ID N°: 2 o el complemento de la misma están abarcados por la invención, siempre que las secuencias que comprenden compartan al menos el porcentaje de identidad definido con los fragmentos correspondientes de SEC ID N°: 2 o con el complemento de los mismos y siempre que codifiquen un polipéptido que tenga una actividad α -santaleno sintasa, como se ha definido en cualquiera de las realizaciones anteriores. Las mutaciones pueden ser cualquier tipo de mutaciones de estos ácidos nucleicos, tales como mutaciones puntuales, mutaciones de delección, mutaciones de inserción y/o mutaciones de desplazamiento de fase. Puede prepararse un ácido nucleico variante para adaptar su secuencia de nucleótidos a un sistema de expresión específico. Por ejemplo, se sabe que los sistemas de expresión bacterianos expresan más eficazmente polipéptidos si los aminoácidos están codificados por un codón preferido. Debido a la degeneración del código genético, en la que más de un codón puede codificar el mismo aminoácido, múltiples secuencias de ADN pueden codificar el mismo polipéptido, abarcándose todas estas secuencias de ADN por la invención.

Otra herramienta importante para transformar organismos o células huésped adecuados para llevar a cabo el procedimiento de la invención *in vivo* es un vector de expresión que comprende un ácido nucleico de acuerdo con cualquier realización de la invención. Dicho vector es por lo tanto también un objeto de la presente invención.

50 Un "vector de expresión" como se usa en el presente documento incluye cualquier vector recombinante lineal o circular incluyendo pero sin limitación vectores virales, bacteriófagos y plásmidos. El experto en la materia es capaz de seleccionar un vector adecuado de acuerdo con el sistema de expresión. En una realización, el vector de expresión incluye el ácido nucleico de la invención unido operativamente a al menos una secuencia reguladora, que controla la transcripción, traducción, inicio y terminación, tal como un promotor de la transcripción, operador o potenciador, o un sitio de unión ribosómico ARNm y, opcionalmente, incluyendo al menos un marcador de selección. Las secuencias de nucleótidos están "unidas operativamente" cuando la secuencia reguladora está relacionada funcionalmente con el ácido nucleico de la invención.

Los vectores de expresión de la presente invención pueden usarse en los procedimientos para preparar un organismo y/o célula huésped transformado genéticamente, en organismos y/o células huésped que albergan los ácidos nucleicos de la invención y en los procedimientos para producir o preparar polipéptidos que tengan una actividad α -santaleno sintasa, como se desvela adicionalmente posteriormente.

5 Los organismos y células huésped no humanos recombinantes transformados para albergar al menos un ácido nucleico de la invención de modo que exprese o sobreexpresa de forma heteróloga al menos un polipéptido de la invención también son herramientas muy útiles para llevar a cabo el procedimiento de la invención. Dichos organismos y células huésped no humanos son por lo tanto otro objeto de la presente invención.

10 Un ácido nucleico de acuerdo con cualquiera de las realizaciones anteriormente descritas puede usarse para transformar los organismos y células huésped no humanos y el polipéptido expresado puede ser cualquiera de los polipéptidos anteriormente descritos.

15 Los organismos huésped no humanos de la invención pueden ser cualquier organismo multicelular o unicelular no humano. En una realización preferida, el organismo huésped no humano es una planta, un procarionta o un hongo. Cualquier planta, procarionta u hongo es adecuado para transformarse de acuerdo con la presente invención. Son plantas particularmente útiles las que producen de forma natural altas cantidades de terpenos. En una realización más preferida, la planta se selecciona de la familia de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Malvaceae*, *Asteraceae* o *Lamiaceae*. Por ejemplo, la planta se selecciona de los géneros *Nicotiana*, *Solanum*, *Sorghum*, *Arabidopsis*, *Brassica* (colza), *Medicago* (alfalfa), *Gossypium* (algodón), *Artemisia*, *Salvia* y *Mentha*. Preferentemente, la planta pertenece a la especie de *Nicotiana tabacum*.

20 En una realización más preferida el organismo huésped no humano es un microorganismo. Cualquier microorganismo es adecuado para la presente invención, pero de acuerdo con una realización aún más preferida dicho microorganismo es una bacteria o levadura. Más preferentemente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

25 También pueden transformarse células eucariotas superiores aisladas, en lugar de organismos completos. Como células eucariotas superiores los inventores entienden en el presente documento cualquier célula eucariota no humana excepto células de levadura. Las células eucariotas superiores preferidas son células vegetales o células fúngicas.

30 El término "transformado" se refiere al hecho de que el huésped se sometió a ingeniería genética para comprender una, dos o más copias de cada uno de los ácidos nucleicos requeridos en cualquiera de las realizaciones anteriormente descritas. Preferentemente el término "transformado" se refiere a huéspedes que expresan de forma heteróloga los polipéptidos codificados por el ácido nucleico con el que se transforman, así como que sobreexpresan dichos polipéptidos. En consecuencia, en una realización, la presente invención proporciona un organismo transformado, en el que los polipéptidos se expresan en mayor cantidad que en el mismo organismo no transformado de este modo.

35 Existen varios procedimientos conocidos en la técnica para la creación de organismos o células huésped transgénicos tales como plantas, hongos, procariontas o cultivos de células eucariotas superiores. Se describen vectores de clonación y expresión apropiados para su uso con huéspedes celulares bacterianos, fúngicos, de levadura, vegetales y mamíferos, por ejemplo, en Pouwels y col., *Cloning Vectors: A Laboratory Manual*, 1985, Elsevier, Nueva York y Sambrook y col., *Molecular Cloning: A Laboratory Manual*, 2ª edición, 1989, Cold Spring Harbor Laboratory Press. Están disponibles para los expertos en la materia vectores de clonación y expresión para plantas superiores y/o células vegetales en particular. Véase por ejemplo Schardl y col. *Gene* 61: 1-11, 1987.

40 Los expertos en la materia están familiarizados con procedimientos para transformar organismos o células huésped para que alberguen ácidos nucleicos transgénicos. Para la creación de plantas transgénicas, por ejemplo, los procedimientos actuales incluyen: electroporación de protoplastos vegetales, transformación mediada por liposomas, transformación mediada por agrobacterium, transformación mediada por polietilenglicol, bombardeo de partículas, microinyección de células vegetales y transformación usando virus.

45 En una realización, el ADN transformado se integra en un cromosoma de un organismo y/o célula huésped no humano de modo que resulte un sistema recombinante estable. En la práctica de la invención puede usarse cualquier procedimiento de integración cromosómica conocido en la técnica, incluyendo pero sin limitación intercambio de casetes mediado por recombinasa (RMCE), inserción cromosómica específica de sitio viral, adenovirus e inyección pronuclear.

50 Para llevar a cabo el procedimiento de producción de α -santaleno *in vitro*, como se ha expuesto anteriormente en el presente documento, es muy ventajoso proporcionar un procedimiento para preparar al menos un polipéptido que tenga una actividad α -santaleno sintasa como se ha descrito en cualquier realización de la invención. Por lo tanto, la invención proporciona un procedimiento para producir al menos un polipéptido de acuerdo con cualquiera realización de la invención que comprende:

55 a) cultivar un organismo o célula huésped no humano transformado con el vector de expresión de la invención, de modo que albergue un ácido nucleico de acuerdo con la invención y exprese o sobreexpresa un polipéptido de la invención;

b) aislar el polipéptido del organismo o célula huésped no humano cultivado en la etapa a).

De acuerdo con una realización preferida, dicho procedimiento comprende además, antes de la etapa a), transformar un organismo o célula huésped no humano con el vector de expresión de la invención, de modo que albergue un ácido nucleico de acuerdo con la invención y exprese o sobreexpresa el polipéptido de la invención.

5 Puede usarse un ácido nucleico de acuerdo con cualquiera de las realizaciones anteriormente descritas.

Puede llevarse a cabo transformación y cultivo del organismo o célula huésped no humano como se ha descrito anteriormente para el procedimiento de producir α -santaleno *in vivo*. La etapa b) puede realizarse usando cualquier técnica bien conocida en este campo para aislar un polipéptido particular de un organismo o célula.

10 Una "variante polipeptídica" como se menciona en el presente documento significa un polipéptido que tiene una actividad α -santaleno sintasa y que es sustancialmente homólogo del polipéptido de acuerdo con cualquiera de las realizaciones anteriores, pero que tiene una secuencia de aminoácidos diferente de la codificada por cualquiera de las secuencias de ácido nucleico de la invención debido a una o más deleciones, inserciones o sustituciones.

15 Las variantes pueden comprender secuencias sustituidas de forma conservativa, lo que significa que un resto de aminoácido dado se reemplaza por un resto que tenga características fisicoquímicas similares. Los ejemplos de sustituciones conservativas incluyen sustitución de un resto alifático por otro, tal como Ile, Val, Leu o Ala entre sí, o sustituciones de un resto polar por otro, tal como entre Lys y Arg; Glu y Asp; o Gln y Asn. Véase Zubay, *Biochemistry*, 1983, Addison-Wesley Pub. Co. Los efectos de dichas sustituciones pueden calcularse usando matrices de puntuación de sustitución tales como PAM-120, PAM-200 y PAM-250 como se analiza en Altschul, *J. Mol. Biol.*, 1991, 219, 555-565. Se conocen bien otras de dichas sustituciones conservativas, por ejemplo sustituciones de regiones completas que tienen características de hidrofobicidad similares.

20 También están abarcadas por la invención variantes peptídicas de origen natural. Son ejemplos de dichas variantes proteínas que resultan de acontecimientos de corte y empalme de ARNm alternativos o de escisión proteolítica de los polipéptidos descritos en el presente documento. Las variaciones atribuibles a proteólisis incluyen, por ejemplo diferencias de los extremos N o C tras expresión en diferentes tipos de células huésped, debido a la retirada proteolítica de uno o más aminoácidos terminales de los polipéptidos codificados por las secuencias de la invención.

25 Pueden usarse variantes de los polipéptidos de la invención para obtener por ejemplo actividad enzimática potenciada o reducida deseada, regioquímica o estereoquímica modificada, o utilización de sustrato o distribución de producto alterada, afinidad por el sustrato aumentada, especificidad mejorada para la producción de uno o más compuestos deseados, velocidad aumentada de la reacción enzimática, mayor actividad o estabilidad en un ambiente específico (pH, temperatura, disolvente, etc.) o nivel de expresión mejorado en un sistema de expresión deseado. Puede realizarse un mutante dirigido o variante por cualquier procedimiento conocido en la técnica. Pueden obtenerse variantes y derivados de polipéptidos nativos aislando variantes de origen natural, o la secuencia de nucleótidos de variantes, de otra o la misma línea o especie vegetal, o mediante mutaciones de programación artificial de secuencias de nucleótidos que codifiquen los polipéptidos de la invención. Pueden conseguirse alteraciones de la secuencia de aminoácidos por cualquiera de varios procedimientos convencionales.

30 Pueden usarse variantes polipeptídicas resultantes de una fusión de secuencias peptídicas adicionales en los extremos amino y carboxilo terminal de los polipéptidos de la invención para potenciar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente o sistema de expresión deseado. Dichas secuencias peptídicas adicionales pueden ser péptidos señal, por ejemplo. En consecuencia, la presente invención abarca variantes de los polipéptidos de la invención, tales como los obtenidos por fusión con otros oligo o polipéptidos y/o los que están unidos a péptidos señal. El polipéptido de fusión abarcado por la invención también comprende polipéptidos de fusión resultantes de una fusión u otras proteínas funcionales, tales como otras proteínas de la ruta de biosíntesis de terpeno.

35 Por lo tanto, en una realización, la presente invención proporciona un procedimiento para preparar un polipéptido variante que tenga una actividad α -santaleno sintasa, como se ha descrito en cualquiera de las realizaciones anteriores, y comprende las etapas de:

- (a) seleccionar un ácido nucleico de acuerdo con cualquiera de las realizaciones expuestas anteriormente;
- (b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
- (c) transformar células huésped y organismos unicelulares con la secuencia de ácido nucleico mutante para expresar un polipéptido codificado por la secuencia de ácido nucleico mutante;
- (d) explorar el polipéptido con respecto a al menos una propiedad modificada; y
- (e) opcionalmente, si el polipéptido no tiene actividad α -santaleno sintasa variante deseada, repetir las etapas de procedimiento (a) a (d) hasta que se obtenga un polipéptido con una actividad α -santaleno sintasa variante deseada;
- (f) opcionalmente, si se identificó un polipéptido que tenía una actividad α -santaleno sintasa variante deseada en la etapa (d), aislar el ácido nucleico mutante correspondiente obtenido de la etapa (c).

De acuerdo con una realización preferida, el polipéptido variante preparado es capaz de producir α -santaleno como un producto principal. De acuerdo con una realización aún más preferida, es capaz de producir una mezcla de sesquiterpenos en la que α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% de los sesquiterpenos producidos.

5 De acuerdo con una realización más preferida, el polipéptido variante preparado tiene una actividad (+)- α -santaleno sintasa.

De acuerdo con una realización aún más preferida, el polipéptido variante preparado es capaz de producir (+)- α -santaleno como un producto principal. De acuerdo con una realización aún más preferida, es capaz de producir una mezcla de sesquiterpenos en la que (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90%, preferentemente al menos el 92% de los sesquiterpenos producidos.

10 En la etapa (b), puede crearse un gran número de secuencias de ácido nucleico mutantes, por ejemplo por mutagénesis aleatoria, mutagénesis específica de sitio o combinación de ADN. Los procedimientos detallados de combinación génica se encuentran en Stemmer, DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Natl Acad Sci U S A., 1994, 91(22): 10747-1075. En resumen, la combinación de ADN se refiere a un procedimiento de recombinación aleatoria de secuencias conocidas *in vitro* que implica al menos dos ácidos nucleicos seleccionados para recombinación. Por ejemplo pueden introducirse mutaciones en loci particulares sintetizando oligonucleótidos que contienen una secuencia mutante, flanqueados por sitios de restricción que permiten la ligación con fragmentos de la secuencia nativa. Después de la ligación, la secuencia reconstruida resultante codifica un análogo que tiene la inserción, sustitución o delección de aminoácidos deseada. Como alternativa, pueden emplearse procedimientos de mutagénesis específicos de sitio dirigidos por oligonucleótidos para proporcionar un gen alterado en el que los codones predeterminados pueden alterarse por sustitución, delección o inserción.

15 En consecuencia, el polipéptido que comprende SEC ID N°: 1 puede recombinarse con cualquier otro ácido nucleico que codifique sesquiterpeno sintasa, por ejemplo aislados de un organismo distinto de *Clausena lansium*. Por lo tanto, pueden obtenerse y separarse ácidos nucleicos mutantes, que pueden usarse para transformar una célula huésped de acuerdo con procedimientos convencionales, por ejemplo tales como se desvelan en los presentes ejemplos.

20 En la etapa (d), el polipéptido obtenido en la etapa (c) se explora con respecto a al menos una propiedad modificada, por ejemplo una actividad enzimática modificada deseada. Los ejemplos de actividades enzimáticas deseadas, para las que puede explorarse un polipéptido expresado, incluyen actividad enzimática potenciada o reducida, como se mide por el valor K_M o $V_{m\acute{a}x}$, regioquímica o estereoquímica modificada y utilización de sustratos o distribución de productos alterada. La exploración de la actividad enzimática puede realizarse de acuerdo con procedimientos familiares para los expertos en la materia y los desvelados en los presentes ejemplos.

25 La etapa (e) proporciona la repetición de las etapas de procedimiento (a)-(d), que pueden realizarse preferentemente en paralelo. En consecuencia, creando un número significativo de ácidos nucleicos mutantes, muchas células huésped pueden transformarse con diferentes ácidos nucleicos mutantes al mismo tiempo, permitiendo la exploración posterior de un número elevado de polipéptidos. Las posibilidades de obtener un polipéptido variante deseado pueden por lo tanto aumentar a la discreción del experto en la materia.

Todas las publicaciones mencionadas en la presente solicitud se incorporan por referencia para desvelar y describir los procedimientos y/o materiales en relación con los que se citan en las publicaciones.

40 **Descripción de los dibujos**

Figura 1: Secuencias de aminoácidos deducidas de los fragmentos de sesquiterpeno sintasas obtenidos de la secuenciación de la biblioteca de *C. lansium* y alineadas con la secuencia de aminoácidos de sesquiterpeno sintasa con el N° de referencia de NCBI AAK54279.

45 Figura 2: Comparación de los perfiles de producto obtenidos de E,E-FPP con las proteínas recombinantes Cont2-1, Cont2B_22, Cont2B_26 y Cont2B 29. El análisis se realizó mediante CG-EM y se muestran los cromatogramas iónicos totales.

Figura 3: Identificación de α -santaleno por comparación del espectro de masas del pico en el tiempo de retención de 12,63 minutos y el espectro de masas de un patrón auténtico de α -santaleno.

Realizaciones específicas de la invención o Ejemplos

50 La invención se describirá ahora en más detalle mediante los siguientes Ejemplos.

Ejemplo 1

Material vegetal y construcción de biblioteca de ADNc

Se obtuvieron semillas de *Clausena lansium* (wampee) de granjeros localizados en la provincia de Hainan en China y particularmente en la ciudad de FuShan (Condado de ChengMai) y la ciudad de Yongxing (Ciudad de Haikou). Las

semillas se hicieron germinar y las plantas se cultivaron en un invernadero.

Se recogieron hojas jóvenes (de 1 a 2 cm de longitud) y se usaron para la construcción de una biblioteca de ADNc. Se extrajo ARN total de las hojas usando el Reactivo de ARN Vegetal Concert™ de Invitrogen (Carlsbad, CA) y el ARNm se purificó mediante cromatografía de afinidad de oligodT-celulosa usando el Kit de aislamiento de ARNm FastTrack® 2.0 (Invitrogen, Carlsbad, CA) de acuerdo con las instrucciones del fabricante. Se construyó una biblioteca de ADNc de este ARNm y usando el Kit de amplificación de ADNc Marathon™ (Clontech, Mountain View, CA).

Ejemplo 2

Secuenciación paralela masiva de la biblioteca de ADNc de hoja de *C. lansium*

Los inventores usaron la tecnología de secuenciación paralela masiva de fragmentos de ADN pequeños desarrollada por Illumina (San Diego, California) para obtener información de secuencia de la biblioteca de ADNc completa preparada a partir de hojas pequeñas de wampee. Esta técnica de secuenciación usa una química de secuenciación basada en terminador reversible y los aparatos Estación de Grupos y Secuenciador Genómico desarrollados por Solexa e Illumina (www.illumina.com).

La biblioteca de ADNc (1 µg) se cargó en primer lugar en un gel de agarosa y se escindieron las bandas correspondientes a un tamaño entre 1,5 y 3 Kb, se eluyeron y se usaron para la secuenciación. Este enriquecimiento de tamaño evita la dilución de la biblioteca por algunos ADNc que codifican proteínas implicadas en el metabolismo primario (tales como por ejemplo la ribulosa-1,5-bisfosfato carboxilasa) que están presentes con frecuencia en alta proporción en una biblioteca realizada a partir de tejidos vegetales y especialmente tejidos verdes. Los ADNc diana, que codifican sesquiterpeno sintasas, típicamente tienen un tamaño entre 1,8 y 2,5 Kb y están incluidos por lo tanto en la biblioteca de tamaño enriquecido.

La tecnología y equipamiento Illumina se preparó en FASTERIS SA (Ginebra, Suiza) y se realizó la preparación de la muestra de ADN y la secuenciación por FASTERIS SA. La biblioteca de ADNc se trató usando el Kit de Preparación de Muestras Genómicas (Illumina). Brevemente, el ADN se fragmenta por nebulización, los extremos se reparan para generar extremos romos, se ligan adaptadores a los extremos de los fragmentos de ADN y los fragmentos de ADN modificados por adaptador se amplifican por PCR. Después de controlar la calidad de la biblioteca por electroforesis en gel, se realiza la generación de los grupos de ADN en la celda de flujo y la reacción de secuenciación en los equipamientos de Estación de Grupos y Secuenciador Genómico. Usando esta tecnología, se obtuvieron 1,9 millones de secuencias cortas (lecturas) de al menos 35 bases.

El software Edena (Dr David Hernandez, Genomic Research Laboratory, Hospitales de la Universidad de Ginebra, Ginebra, Suiza, resultado no publicado) se usó para volver a ensamblar las secuencias contiguas. Las cinco últimas bases se retiraron en primer lugar de cada lectura debido a posibles incorporaciones erróneas debido a la menor fidelidad en los últimos ciclos del procedimiento de secuenciación. Se generaron varios conjuntos de contig (secuencias contiguas). Para cada conjunto, se conservaron los contig de longitud mínima de 50 bases. En primer lugar se ajustaron los parámetros del programa informático para permitir el ensamblaje con un solapamiento mínimo de 25 bases e identidad estricta (100%) o no estricta (emparejamiento erróneo de 2 bases). Se generaron por lo tanto dos conjuntos de 3634 y 3756 contig respectivamente. Se generó otro conjunto de 4540 contig permitiendo el ensamblaje con un mínimo de 18 bases y solapamiento no estricto. Las secuencias de los contig se usaron para buscar homología con terpeno sintasas en bases de datos de proteínas públicamente disponibles usando el algoritmo Blastx (Altschul y col, J. Mol. Biol. 215, 403-410, 1990; <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). De los tres conjuntos de contig, se seleccionaron 14, 15 y 14 contig. A lo largo del análisis de las secuencias obtenidas de la biblioteca de ADNc de *Clausena lansium*, se observó homología de secuencia fuerte con secuencias de especies de cítricos, una observación coherente con la relación filogenética de *Clausena lansium* y especies de Cítricos (pertenecientes ambas a la familia Rutaceae). Por lo tanto, se usó el software Eland (Illumina) para buscar en lecturas no ensambladas identidad de secuencia de ADN con sesquiterpeno sintasas de cítrico (Nº de Referencia de NCBI CQ813507, CQ813505, CQ813508, CQ813506). A partir de este análisis, se seleccionaron 117 lecturas.

Los contig y las lecturas seleccionados se procesaron después usando el programa CAP (Huang, Genomics 14(1), 18-25, 1992) y se generaron nuevos contig. Después de la confirmación de homología de secuencia con sesquiterpeno sintasas, se conservaron 17 contig de longitud de 30 a 436 bases (véase SEC ID Nº: 3 a 19). Las secuencias deducidas se alinearon con una secuencia de sesquiterpeno sintasa de cítrico (la beta-farneseno sintasa de *C. junos*, Nº de referencia de NCBI AAK54279) para mapear su posición relativa a lo largo de una secuencia de sesquiterpeno sintasa de longitud completa y evaluar el número de ADNc de sesquiterpeno diferentes presentes (Figura 1). Se diseñó un conjunto de oligonucleótidos específicos de 6 de los 19 contig que surgen supuestamente de ADNc de sesquiterpeno sintasas diferentes.

Ejemplo 3

Amplificación de ADNc de sesquiterpeno sintasas de longitud completa

Los cebadores específicos de sesquiterpeno sintasas deducidos a partir de la secuenciación paralela masiva (Ejemplo 2) se usaron en combinación con cebadores adaptadores de ADNc en amplificaciones de PCR de tipo

3'/5'RACE. Las amplificaciones se realizaron usando la biblioteca de ADNc de *C. lansium*, preparada como se ha descrito anteriormente en el Ejemplo 1, y la Mezcla de Polimerasa Advantage[®] 2 (Clontech) después del protocolo del Kit de Amplificación de ADNc Marathon[™] (Clontech, Mountain View, CA). Las condiciones de termociclación fueron las siguientes: 1 minuto a 94 °C, 32 ciclos de 1 minuto a 94 °C y 3 minutos a 68 °C, y 3 minutos a 68 °C.

- 5 Usando el cebador FS2 cont2 F1 (SEC ID N°: 20), se obtuvo una secuencia de ADN de 1049 pb. El análisis de las secuencias de varios clones obtenidos a partir de esta amplificación mostró que estaban presentes dos variantes de secuencia (Cont2_RACE_F1 (SEC ID N°: 23) y Cont2_RACE_F2 (SEC ID N°: 25)) con 96% de identidad de secuencia. Cada una de las dos secuencias correspondía al extremo 3' de un ADNc de sesquiterpeno sintasa y contenía una región codificante de 735 pb. Las dos secuencias de aminoácidos deducidas (SEC ID N°: 24 y 26) tuvieron 92% de identidad de secuencia entre sí. Con el cebador FS2_cont2_R1 (SEC ID N°: 21), se amplificó un fragmento de 1101 pb (Cont2_RACE_R, SEC ID N°: 27)) que contenía el codón de partida y que codificaba los 349 aminoácidos N terminales del sesquiterpeno correspondiente al contig2. El alineamiento de las dos secuencias del 3'RACE (Cont2_RACE_F1 y Cont2_RACE_F2, SEC ID N°: 23 y 25) con la secuencia del 5'RACE (cont2_RACE_R, SEC ID N°: 27) mostró un solapamiento de 132 bases. En esta región solapante, las secuencias Cont2_RACE_F2 y Cont2_RACE_R (SEC ID N°: 25 y 27) eran casi idénticas (una diferencia de una única base) mientras que se observaron diferencias de 9 bases entre las secuencias de Cont2_RACE_F1 y Cont2_RACE_R (SEC ID N°: 23 y 27). Por lo tanto se usaron las secuencias Cont2_RACE_F2 (SEC ID N°: 25) y Cont2_RACE_R (SEC ID N°: 27) para reconstituir una secuencia de ADNc de longitud completa (Cont2_RACE_1, SEC ID N°: 28) que codificaba una proteína de 551 aminoácidos (SEC ID N°: 29).
- 20 Con el cebador FS2_Cont10_F (SEC ID N°: 22) se obtuvieron dos secuencias de 1342 pb (Cont10_RACE_Fa y Cont10_RACE_Fb, SEC ID N°: 30 y 31) que mostraban diferencias significativas (67 pb, que representa el 95% de la identidad de secuencia de ADN) y que sugerían la presencia de dos ADNc de sesquiterpeno sintasas estrechamente relacionados. Las dos secuencias contenían una región codificante de 1135 pb. Resulta interesante que la secuencia de Cont10_RACE_Fa (SEC ID N°: 30) era 99,9% idéntica a la secuencia de Cont2_RACE_F2 (SEC ID N°: 25, una diferencia de solamente 1 base en el alineamiento de 1 Kb) y la secuencia de Cont10_RACE_Fb (SEC ID N°: 31) era 99% idéntica a la secuencia de Cont2_RACE_F1 (SEC ID N°: 23, diferencia de solamente 8 bases en el alineamiento de 1 Kb), lo que sugiere por lo tanto que los fragmentos de ADN amplificados con los cebadores Cont2 y Cont10 permitían amplificaciones a partir de dos secuencias relacionadas sin verdadera diferenciación. Se diseñaron dos cebadores (Cont2_inicio (SEC ID N°: 32) y Cont2_parada (SEC ID N°: 33)), que son específicos de las regiones de los codones de inicio y de parada de las secuencias del 5'RACE y el 3'RACE de los fragmentos cont2 y cont10, para amplificar simultáneamente los dos o más ADNc de longitud completa correspondientes. El cebador Cont2_inicio (SEC ID N°: 32) se extendió con la secuencia CACC para permitir la inserción directa en el plásmido pET101/D-TOPO (Invitrogen). La amplificación se realizó en primer lugar usando la Mezcla de Polimerasa Advantage[®] 2 (Clontech). Cada mezcla de PCR contenía, en un volumen total de 50 µl, 5 µl de Tampón de PCR Advantage[®] 2, dNTP 200 µM, 200 nM de cada cebador oligonucleotídico, 5 µl de ADNc diluido 100 veces y 1 µl de Mezcla de Polimerasa Advantage[®] 2. Las condiciones de termociclación fueron las siguientes: 2 minutos a 95 °C; 35 ciclos de 30 segundos a 95 °C, 30 segundos a 60 °C y 4 minutos a 72 °C; y 10 minutos a 72 °C. Se realizó después un segundo ciclo de amplificación usando 5 µl del producto de PCR purificado del primer ciclo de amplificación y usando la ADN polimerasa *Pfu* (Promega), en un volumen final de 50 µl que contenía 5 µl de tampón 10X de ADN polimerasa *Pfu*, 200 µM de cada dNTP, 0,4 µM de cada cebador directo e inverso, 2,9 unidades de ADN polimerasa *Pfu*. Las condiciones de termociclación fueron idénticas a las condiciones usadas en el primer ciclo. Los productos de PCR purificados se ligaron en el vector pET1001/D-TOPO siguiendo las instrucciones del fabricante (Invitrogen). Se seleccionaron varios clones y después de secuenciar el inserto, se observaron algunas variaciones en las secuencias. Se seleccionaron los siguientes clones: Cont2-1 (SEC ID N°: 2), Cont2B_22 (SEC ID N°: 38), Cont2B_26 (SEC ID N°: 39) y Cont2B_29 (SEC ID N°: 40). Las secuencias de las proteínas codificadas por estos clones se proporcionan en SEC ID N°: 1 y 41 a 43, respectivamente.

Ejemplo 4

Expresión heteróloga y actividades enzimáticas de las sesquiterpeno sintasas recombinantes.

- 50 Los plásmidos pET101 con Cont2_1 (SEC ID N°: 2), Cont2B_22 (SEC ID N°: 38), Cont2B_26 (SEC ID N°: 39) y Cont2B_29 (SEC ID N°: 40) preparados como se ha descrito en el Ejemplo 3 se usaron para transformar células *E. coli* BI21(DE3). Se usaron colonias individuales de células transformadas para inocular 5 ml de medio LB. Después de 5 a 6 horas de incubación a 37 °C, el cultivo se transfirió a un incubador a 20 °C y se dejó durante 1 hora para equilibrado. Después se indujo expresión de la proteína mediante la adición de IPTG 1 mM y el cultivo se incubó durante una noche a 20 °C. Al día siguiente, las células se recogieron por centrifugación, se resuspendieron en 0,1 volúmenes de MOPSO 50 mM pH 7, glicerol al 10%, DTT 1 mM y se lisaron por sonicación. El extracto se aclaró por centrifugación (30 minutos a 20.000 g) y el sobrenadante que contenía la proteína soluble se usó para experimentos adicionales.

- 60 El extracto de proteína en bruto se usó para evaluar la actividad enzimática. El ensayo enzimático se realizó en un tubo de vidrio sellado con Teflón usando de 50 a 100 µl de extracto proteico en un volumen final de 1 ml de MOPSO 50 mM pH 7, glicerol al 10% complementado con DTT 1 mM, MgCl₂ 20 mM y E,E-farnesil difosfato (FPP) purificado de 50 a 200

- 5 μM (preparado como se describe en Keller y Thompson, J. Chromatogr 645(1), 161-167, 1993). El tubo se incubó de 18 a 24 horas a 30 °C y los productos enzimáticos se extrajeron dos veces con un volumen de pentano. Después de concentración en un flujo de nitrógeno, el extracto se analizó por CG y se confirmó la identidad de los productos mediante CG-EM basándose en la concordancia de los índices de retención y espectros de masas de patrones auténticos. El análisis de CG-EM se realizó en un sistema de CG de Hewlett-Packard Serie 6890 equipado con un detector de ionización de llama usando una columna capilar SPB-1 de diámetro interno de 0,25 mm por 30 m (Supelco, Bellefonte, PA). El gas portador fue He a un flujo constante de 1,5 ml/minuto. La temperatura inicial del horno fue de 80 °C seguido de un gradiente de 10 °C/minuto hasta 280 °C. Los espectros se registraron a 70 eV con una tensión multiplicadora de electrones de 2200V.
- 10 El ensayo reveló la formación de (+)- α -santaleno como un producto principal (92,7% de los sesquiterpenos totales producidos) y cantidades traza de cinco sesquiterpenos adicionales que representaban del 4,8 al 0,95% de los productos enzimáticos. (+)- α -santaleno se identificó con análisis CG-EM por coincidencia del espectro de masas y del índice de retención con los valores publicados (Joulain, D., y König, W. A. The Atlas of Spectral Data of Sesquiterpene Hydrocarbons, EB Verlag, Hamburgo, 1998). La identificación de (+)- α -santaleno se confirmó
- 15 adicionalmente por ^1H RMN, ^{13}C RMN y mediante medición de la rotación óptica. Para producir suficientes cantidades para estas mediciones, el ensayo enzimático descrito anteriormente se aumentó de escala hasta 1 l. Los productos enzimáticos se extrajeron con un volumen igual de pentano, se concentraron y la fracción de hidrocarburos de sesquiterpeno (5,5 mg) se purificó mediante filtración en una columna de sílice corta. Se proporcionan en la Figura 2 datos espectrales obtenidos con Cont2_1.
- 20 El espectro de RMN se registró en un espectrómetro Bruker-Avance-500. Los datos de RMN son los siguientes:
- ^1H RMN (500,13 MHz, CDCl_3): δ 0,82 (s, 2H), 0,83 (s, 3H), 0,99 (s, 3H), 1,00-1,08 (m, 2H), 1,08-1,26 (m, 2H), 1,57-1,63 (m, 6H), 1,68 (s, 3H), 5,12 (t 3 q, $J = 7,2, 1,4$ Hz, 1H)
- ^{13}C RMN (125,76 MHz, CDCl_3): δ 10,7 (q), 17,5 (q), 19,6 (d), 23,3 (t), 25,7 (q), 27,4 (s), 31,0 (t), 31,5 (t), 34,6 (t), 38,2 (d), 45,9 (s), 125,5 (d), 130,8 (s);
- 25 El hecho de que se produjera el estereoisómero (+)- α -santaleno se ha demostrado midiendo la rotación óptica (como se mide en un polarímetro Perkin-elmer 241): $[\alpha]_D^{20} = +12,0$ ($C = 0,3$, CHCl_3).

Ejemplo 5

Producción *in vivo* de (+)- α -santaleno en *E. coli*

- 30 El uso de la santaleno sintasa de *C. lansium* para la producción *in vivo* de sesquiterpenos en células de *E. coli* se evaluó co-expresando la sesquiterpeno sintasa con una FPP sintasa y las enzimas de una ruta biosintética de cuatro etapas que permitía la conversión de mevalonato a FPP. Los genes de la ruta de mevalonato se organizaron en un único operón y codificaban una mevalonato quinasa (mvaK1), una fosfomevalonato quinasa (mvaK2), una mevalonato difosfato descarboxilasa (MvaD) y una isopentenil difosfato isomerasa (idi), convirtiendo todas las enzimas mevalonato exógeno en isopentenil difosfato (IPP) y dimetilalil difosfato (DMAPP), los dos sustratos de la
- 35 FPP sintasa. La co-expresión de esta ruta de mevalonato parcial se usó para aumentar la cantidad de FPP intracelular disponible para la sesquiterpeno sintasa y de este modo las cantidades de sesquiterpeno producido.

- El gen de FPP sintasa de levadura (número de Referencia J05091) se amplificó a partir de ADN genómico de *S. cerevisiae* usando los cebadores FPPy NcoI (SEC ID N°: 34) y FPPy-Eco (SEC ID N°: 35). El ADN genómico se aisló de *S. cerevisiae* usando el Kit Maxi de ARN/ADN de Qiagen (Qiagen AG, Basilea, Suiza). La PCR se realizó con la
- 40 ADN polimerasa *Pfu* (Promega AG, Dubendorf, Suiza) en un volumen final de 50 μl que contenía 0,4 μl de cada cebador, dNTP 200 μM , 0,5 μl de ADN polimerasa, 5 μl de ADN genómico de *S. cerevisiae*. Las condiciones de ciclación de PCR fueron las siguientes: 90 segundos a 95 °C; 28 ciclos de 45 segundos a 95 °C, 30 segundos a 54 °C y 4 minutos a 72 °C; 10 minutos a 72 °C. El ADN amplificado se ligó como fragmento NdeI-EcoRI en el primer sitio de clonación múltiple (MCS1) del plásmido pACYCDuet-1 (Novagen, Madison, WI) que proporciona el plásmido
- 45 pACYCDuet-FPP que alberga el gen de FPP bajo el control de un promotor T7.

- Se amplificó un operón que contenía los genes que codifican mvaK1, mvaK2, MvaD e idi a partir de ADN genómico de *Streptococcus pneumoniae* (ATCC BAA-334, LGC Standards, Molsheim, Francia) con los cebadores MVA-up1-inicio (SEC ID N°: 36) y MVA-up2-parada (SEC ID N°: 37). La PCR se realizó usando la ADN polimerasa *PfuUltra*™ II Fusion HS (Stratagene, Agilent Technologies Inc., Santa Clara, CA, Estados Unidos). La composición de la mezcla
- 50 de PCR fue de acuerdo con las instrucciones del fabricante. Las condiciones de termociclación fueron 2 minutos a 95 °C; 30 ciclos de 20 segundos a 95 °C, 20 segundos a 58 °C y 90 segundos a 72 °C; y 3 minutos a 72 °C. El fragmento de 3,8 Kb se purificó en un gel de agarosa y se ligó usando el Kit de Clonación de PCR In-Fusion™ Dry-Down (Clontech Laboratories) en el segundo MCS del plásmido pACYCDuet-FPPs digerido con *NdeI* y *XhoI* que proporciona el plásmido pACYCDuet-4506. Las secuencias de los dos insertos se secuenciaron completamente para
- 55 excluir cualquier mutación.

5 Se transformaron células *E. coli* BL21 Star™ (DE3) (Invitrogen, Carlsbad, CA) con los plásmidos pET101-cont2 1 (SEC ID N°: 2) preparados como se ha descrito en el Ejemplo 3 y con el plásmido pACYCDuet-4506. Las células transformadas se seleccionaron en placas de LB-agarosa de carbenicilina (50 µg/ml) y cloranfenicol (34 µg/ml). Se usaron colonias individuales para inocular 5 ml de medio LB líquido complementado con los mismos antibióticos. El cultivo se incubó durante una noche a 37 °C. Al día siguiente se inocularon 2 ml de medio TB complementado con los mismos antibióticos con 0,2 ml del cultivo de una noche. Después de 6 horas de incubación a 37 °C, el cultivo se enfrió a 28 °C y se añadió IPTG 1 mM, mevalonato 2 mg/ml (preparado disolviendo mevalonolactona (Sigma) en NaOH 0,5 N a una concentración de 1 g/ml e incubando la solución durante 30 minutos a 37 °C) y 0,2 ml de decano a cada tubo. Los cultivos se inocularon durante 48 horas a 28 °C. Los cultivos se extrajeron después dos veces con dos volúmenes de etil-acetato, la fase orgánica se concentró a 500 µl y se analizó mediante CG-EM como se ha descrito anteriormente en el Ejemplo 4. En estas condiciones las células produjeron (+)-α-santaleno a 250 mg/l de cultivo en 48 horas.

15 Este ejemplo muestra que una célula *E. coli* transformada con una α-santaleno sintasa, como se ha definido en la presente invención, es capaz de producir α-santaleno. Las otras enzimas con las que se transforma la célula *E. coli* no son esenciales para la producción de α-santaleno. De hecho α-santaleno también se produce cuando se transforma una célula *E. coli* solamente con la α-santaleno sintasa, pero en cantidades menores. Las otras enzimas con las que se transforma la célula *E. coli* se añaden con el único fin de aumentar la cantidad de precursor disponible para la α-santaleno sintasa.

LISTADO DE SECUENCIAS

- 20 <110> Firmenich SA
- <120> Procedimiento de producción de (+)-alfa-santaleno
- 25 <130> 7280
- <160> 43
- <170> PatentIn versión 3.5
- 30 <210> 1
- <211> 551
- <212> PRT
- <213> Clausena lansium
- 35 <400> 1

ES 2 410 906 T3

Met Ser Thr Gln Gln Val Ser Ser Glu Asn Ile Val Arg Asn Ala Ala
 1 5 10 15

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

Gln Thr Ile Asp Ser Trp Thr Gln Gln His His Lys Glu Leu Lys Glu
 35 40 45

Glu Val Arg Lys Met Met Val Ser Asp Ala Asn Lys Pro Ala Gln Arg
 50 55 60

Leu Arg Leu Ile Asp Thr Val Gln Arg Leu Gly Val Ala Tyr His Phe
 65 70 75 80

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

Asp Asp Lys Asp Asp Leu Tyr Ile Val Ser Leu Cys Phe Arg Leu Leu
 100 105 110

Arg Gln His Gly Ile Lys Ile Ser Cys Asp Val Phe Glu Lys Phe Lys
 115 120 125

Asp Asp Asp Gly Lys Phe Lys Ala Ser Leu Met Asn Asp Val Gln Gly
 130 135 140

Met Leu Ser Leu Tyr Glu Ala Ala His Leu Ala Ile His Gly Glu Asp
 145 150 155 160

Ile Leu Asp Glu Ala Ile Val Phe Thr Thr Thr His Leu Lys Ser Thr
 165 170 175

ES 2 410 906 T3

Val Ser Asn Ser Pro Val Asn Ser Thr Phe Ala Glu Gln Ile Arg His
 180 185 190

Ser Leu Arg Val Pro Leu Arg Lys Ala Val Pro Arg Leu Glu Ser Arg
 195 200 205

Tyr Phe Leu Asp Ile Tyr Ser Arg Asp Asp Leu His Asp Lys Thr Leu
 210 215 220

Leu Asn Phe Ala Lys Leu Asp Phe Asn Ile Leu Gln Ala Met His Gln
 225 230 235 240

Lys Glu Ala Ser Glu Met Thr Arg Trp Trp Arg Asp Phe Asp Phe Leu
 245 250 255

Lys Lys Leu Pro Tyr Ile Arg Asp Arg Val Val Glu Leu Tyr Phe Trp
 260 265 270

Ile Leu Val Gly Val Ser Tyr Gln Pro Lys Phe Ser Thr Gly Arg Ile
 275 280 285

Phe Leu Ser Lys Ile Ile Cys Leu Glu Thr Leu Val Asp Asp Thr Phe
 290 295 300

Asp Ala Tyr Gly Thr Phe Asp Glu Leu Ala Ile Phe Thr Glu Ala Val
 305 310 315 320

Thr Arg Trp Asp Leu Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys
 325 330 335

Phe Ile Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu
 340 345 350

Leu Ala Lys Glu Gly Arg Ser Tyr Ser Ile His Tyr Ala Ile Arg Ser
 355 360 365

Phe Gln Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn
 370 375 380

Lys Gly Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg
 385 390 395 400

Ser Ile Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly
 405 410 415

ES 2 410 906 T3

Asp Ile Ala Thr Lys Glu Val Phe Glu Trp Glu Met Asn Asn Pro Lys
 420 425 430

Ile Ile Ile Ala Ala Glu Thr Ile Phe Arg Phe Leu Asp Asp Ile Ala
 435 440 445

Gly His Arg Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu
 450 455 460

Cys Tyr Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala
 465 470 475 480

Leu Ser Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu
 485 490 495

Leu Leu Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu
 500 505 510

Asp Leu Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg
 515 520 525

Phe Thr His Ser Thr Met Met Lys Asp Gln Val Asp Leu Val Leu Lys
 530 535 540

Asp Pro Val Lys Leu Asp Asp
 545 550

<210> 2
 <211> 1656
 <212> ADN
 <213> Clausena lansium

<400> 2

atgtcaactc aacaagtttc atcagagaac attgttcgta acgctgcgaa tttccatcct 60
 aatatatggg gaaaccattt cctcacatgt ccttctcaga cgattgatag ttggactcaa 120
 cagcaccaca aagaactgaa agaagagggtg aggaaaatga tgggtgtctga tgcaaataaa 180
 cctgcccaga gattgcgctt gattgatact gtccaaaggc taggtgtggc ttaccacttt 240
 gaaaaggaga ttgatgatgc attggagaaa atagggtcatg acccttttga tgataaagat 300
 gatctctaca ttgtctctct ttgttttcga ttgctgaggc agcatggaat taagatatca 360
 tgtgatgtgt ttgagaagtt taaagatgac gatggaaaat tcaaggcatc attgatgaat 420
 gatgttcaag gcatgctaag tttatatgag gcagcacacc tagccattca cggagaagat 480
 attttagatg aagcaattgt tttcacgacc actcacctta agtcaacggg atctaattct 540

5

10

ES 2 410 906 T3

cctgtaaact ctacttttgc tgaacaaata cgtcattctc tcagagttcc tctccgtaa 600
gctgtaccta ggtagagtc gaggtatttc ttggatatct attcaagaga tgatttgac 660
gataaaactt tgctcaattt cgcaaagtta gactttaata tactacaagc aatgcaccag 720
aaggaagcaa gtgagatgac caggtggtgg agagattttg acttccttaa aaagctgcct 780
tatataagag acagagtcgt ggagctatat ttttggattc tgggtggagt gtcttatcag 840
cccaaattca gcaactggtag aatTTTTTTG tccaaaataa tatgccttga gaccctgta 900
gatgatacat ttgacgccta cggactttt gacgagctcg caatctttac tgaagcagtt 960
acaagatggg accttggcca cagagatgca ctaccagaat acatgaaatt cattttcaag 1020
acactcattg atgtctacag tgaagctgag caagaactgg caaaggaagg gagatcatac 1080
agcatacact atgcaatacg atcgttccaa gaactagtta tgaagtactt ctgcgaagcc 1140
aagtggttaa ataaaggtta tgttccgagc ctggacgatt ataaatcagt ttcattaaga 1200
agtatcggtt tttaccgat agcggtagct tccttcgttt tcatgggtga tattgcaact 1260
aaggaggtct ttgaatggga aatgaataac cctaagatca taatagccgc agaaacgatt 1320
ttcagattcc tggatgacat agcaggccat aggtttgagc aaaagagaga acatagtcca 1380
tcagctattg aatgctacaa gaatcaacat ggagtgtctg aggaagaggc agttaaagcg 1440
ttgtcgttag aagttgctaa tagttggaaa gatataaatg aggagctgct tctcaacca 1500
atggctatlc ctttacctct gottcaggtg attcttgatc tctcacgttc ggccgatttt 1560
atgtacggta atgctcaaga tcgcttcacg cattcaacga tgatgaaaga ccaagttgat 1620
ttggtgctga aggaccccg taaagcttgac gattaa 1656

<210> 3
<211> 211
<212> ADN
<213> Clausena lansium

5

<400> 3

gccgcagaaa cgattttcag attcctggat gatgtagcag gccataagtt tgagcaaaag 60
agagaacatt gtccatcagc tattgaaatgc tacaagaatc aacatggagt gtctgaggaa 120
gaggcagtta aagcgttgtc gttagaagtt gctaatagatt ggaaagatat aatgaggag 180
ctgcttctca acccaatggc tattccttta c 211

10

<210> 4
<211> 436
<212> ADN
<213> Clausena lansium

15

<400> 4

ES 2 410 906 T3

	gtccaaaata atatgccttg agaccctcgt agatgataca tttgacgcct acggactttt	60
	tgaagagctc acaatcttta ctgaagcagt tacaagatgg gacattggcc acacagatgc	120
	actaccagat tacatgaaat tccttttcaa gacactcatt gatgtctata gtgaagctga	180
	ggaagaactg gcaaaggaag gaagatcata cagcatacaa tatgcaatac gatcgtttca	240
	agaactagct atgaaatact tctgcgaagc caagtggta aataaaggtt atgttccgag	300
	cctggacgat tataaatcag tttcattaag aagtatcggg tttttaccga tagcggtagc	360
	ttccttcggt ttcattgggtg atattgcaac taaggaggtc tttgaatggg aatgaataa	420
	ccctaagatc ataata	436
5	<210> 5 <211> 71 <212> ADN <213> Clausena lansium	
	<400> 5	
10	cacgttcggc cgattttatg tacggcaatg gtcaagatcg ctacacgcat tcaacgatga	60
	tgaaagacca a	71
15	<210> 6 <211> 133 <212> ADN <213> Clausena lansium	
	<400> 6	
20	tgaccctttt gatgataaag atgatctcta cattgtctct ctttgttttc gattgctgag	60
	gcagcatgga attaagatat catgtgatgt gtttgagaag tttaaagatg acgatggaaa	120
	attcaaggca tcc	133
25	<210> 7 <211> 172 <212> ADN <213> Clausena lansium	
	<400> 7	
30	agagtcgagg tatttcttgg atatctattc aagagatgat ttgcacgata aaactttgct	60
	caatttcgca aagttagact ttaatatact acaagcaatg caccagaagg aagcaagtga	120
	gataaccagg tgggtggagag attttgatt ccttgaaaag ctgccttatg ta	172
35	<210> 8 <211> 73 <212> ADN <213> Clausena lansium	
	<400> 8	

ES 2 410 906 T3

actaatggat gacatagtgt cacacaagtt tgaacaaagc agagggcacg ttgcctcgag 60

cgttgagtgt tac 73

5 <210> 9
<211> 46
<212> ADN
<213> Clausena lansium

10 <400> 9
actgctttc ctggtgcttt aattgagaga cctttcaata tcgcac 46

15 <210> 10
<211> 47
<212> ADN
<213> Clausena lansium

20 <400> 10
cactcttat gcgaattctc aatcttacgc gcggtataga tgtcatt 47

25 <210> 11
<211> 44
<212> ADN
<213> Clausena lansium

30 <400> 11
ccagaaggaa ctcggtgaca ttcaagggtg gtggaaagaa ttag 44

35 <210> 12
<211> 42
<212> ADN
<213> Clausena lansium

40 <400> 12
gcatggaatg agttccggaa acaagttca aatgcctgga ag 42

45 <210> 13
<211> 39
<212> ADN
<213> Clausena lansium

50 <400> 13
taatcgtgct gaacaaatta atcatgctct cgactgtcc 39

55 <210> 14
<211> 38
<212> ADN
<213> Clausena lansium

60 <400> 14
aagtatgaag atggctacac tcattctgca gttgtgct 38

65 <210> 15
<211> 38
<212> ADN
<213> Clausena lansium

70 <400> 15
ttttagatga agcaattgtt ttcacgacca ctcacctt 38

75 <210> 16
<211> 37
<212> ADN
<213> Clausena lansium

ES 2 410 906 T3

<400> 16
 cctggaagga tataaacgag gagtgcttac gcccaac 37

5
 <210> 17
 <211> 30
 <212> ADN
 <213> Clausena lansium

10
 <400> 17
 acacctttg ttggcatggg agacattgta 30

15
 <210> 18
 <211> 30
 <212> ADN
 <213> Clausena lansium

20
 <400> 18
 gagggcacgt tgcctcaagc gtagagtgtt 30

25
 <210> 19
 <211> 30
 <212> ADN
 <213> Clausena lansium

30
 <400> 19
 cgactgctta ggcagcaagg atttaaggtt 30

35
 <210> 20
 <211> 31
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador

40
 <400> 20
 cttgagaccc tcgtagatga tacattgac g 31

45
 <210> 21
 <211> 31
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador

50
 <400> 21
 gtatgatctt ccttccttg ccagttcttc c 31

55
 <210> 22
 <211> 31
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador

60
 <400> 22
 gatgaagcaa ttgtttcac gaccactcac c 31

65
 <210> 23
 <211> 1049
 <212> ADN
 <213> Clausena lansium

<400> 23

ES 2 410 906 T3

```

cctatggtac tcttgaagag ctacacagtct ttactgaagc aattacaaga tgggacgttg      60
gccacacaga tgcactacca gattacatga aattcctttt caagacactc attgatgtct      120
atagtgaagc tgaggaagaa ctggcaaagg aaggaagatc atacagcata caatatgcaa      180
taogatcggt tcaagaacta gctatgaaat acttctgcga agcggagtgg ttaaataaag      240
gttatggtcc gagcctggac gagtataaat cagtttcagt aagaagtgtc ggtttttttc      300
cgatagcggg agcttctctc gttttcatgg gtgatattgc aactaaggag gtcttttgaat      360
gggaaatgaa taaccctaag atcataatag cgcgagaaac gattttcaga ttcttgatg      420
atgcagcagg ccataagttt gagcaaaaga gagaacattg tccatcagct attgaatgct      480
acaagaatca acatggagtg tctgaggaag aggcagttaa agcgttgctg ttagaagttg      540
ctaatagttg gaaagatata aatgaggagc tgcttctcaa cccaatggct attcctttac      600
ctctacttca ggtgattctt gatctctcac gttcggccga ttttatgtac ggcaatggtc      660
aagatcgcta cacgcattca acgatgatga aagaccaagt tgacttggtg ctgaaggacc      720
ccgttaagct tgaogattaa agttatggtg ctgatttctc atcgtatatt tgagaagttg      780
gtaataaatt aagttgggtc ttgctagtta ttagctagc tagtcatgcg tagctaaggg      840
atggttcaat tgattaggcc tatattctag taaaaataaa aggagtaaga acgaatctcc      900
ctcacaccaa cttcgcaata atgtaattta tttcatctat gtctgttaca aaaattttga      960
gattaaaata aacagcaatc ctacttgcac ggaataaaca ataataccat ttttactaat     1020
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa                                         1049

```

5 <210> 24
 <211> 245
 <212> PRT
 <213> Clausena lansium

10 <400> 24

ES 2 410 906 T3

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

ES 2 410 906 T3

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

Val Lys Leu Asp Asp
 245

5 <210> 25
 <211> 906
 <212> ADN
 <213> Clausena lansium

<400> 25

cctacggtac ttttgacgag ctcaaatct ttactgaagc agttacaaga tgggacattg 60
 gccacagaga tgcactacca gaatacatga aattcatttt caagacactc attgatgtct 120
 acagtgaagc tgagcaagaa ctggcaaagg aagggagatc atacagcata caatatgcaa 180
 tacgatcgtt ccaagaacta gttatgaagt acttctgcga agccaagtgg ttaaataaag 240
 gttatgttcc gagcctggac gattataaat cagtttcatt aagaagtatc ggttttttac 300
 cgatagcggc agcttccttc gttttcatgg gtgatattgc aactaaggag gtctttgaat 360
 gggaaatgaa taaccctaag atcataatag ccgcagaaac gattttcaga ttctctggatg 420
 acatagcagg ccataagttt gagcaaaaga gagaacatag tccatcagct attgaatgct 480
 acaagaatca acatggagtg tctgaggaag aggcagttaa agcgttgtcg ttagaagttg 540
 ctaatagttg gaaagatata aatgaggagc tgcttctcaa cccaatggct attcctttac 600
 ctctgcttca ggtgattctt gatctctcac gttcggccga ttttatgtac ggtaatgctc 660
 aagatcgctt cacgcattca acgatgatga aagaccaagt tgatttgggtg ctgaaggacc 720
 ccgttaagct tgacgattaa agttatgttg ctgatttcct atcgtatatt tgagaagttg 780
 gtaataaatt aagttgggtg ttgctagtta tttagctagc tagtcatgcg tagctagggg 840
 atggttcaat tgattaggcc tatattctag taaaaataaa cgatgtaaga acaaatctcc 900
 ctcgca 906

10

15 <210> 26
 <211> 245
 <212> PRT
 <213> Clausena lansium

<400> 26

Tyr Gly Thr Phe Asp Glu Leu Thr Ile Phe Thr Glu Ala Val Thr Arg
 1 5 10 15

Trp Asp Ile Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys Phe Ile
 20 25 30

ES 2 410 906 T3

Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu Leu Ala
 35 40 45

Lys Glu Gly Arg Ser Tyr Ser Ile Gln Tyr Ala Ile Arg Ser Phe Gln
 50 55 60

Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn Lys Gly
 65 70 75 80

Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg Ser Ile
 85 90 95

Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly Asp Ile
 100 105 110

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

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

Lys Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu Cys Tyr
 145 150 155 160

Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala Leu Ser
 165 170 175

Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu Leu Leu
 180 185 190

Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu Asp Leu
 195 200 205

Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg Phe Thr
 210 215 220

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

Val Lys Leu Asp Asp
 245

<210> 27
 <211> 1101
 <212> ADN
 <213> Clausena lansium

5

<400> 27

ES 2 410 906 T3

ctagtagcta aaaaaattat taagttcaat cttttttgtc tgtctagaga aagatgtcaa 60
ctcaacaagt ttcacagag aacattgttc gtaatgctgc agatttccat cctaataat 120
gggaaacca tttcctcaca tgtctttctc aaacgattga tagttggact caacagcacc 180
aaaagaact gaaagaagag gtgaggaaaa tgatggtgtc tgatgcaaat aaacctgccc 240
agagattgag cttgattgat actgtccaaa ggtaggtgt ggcttaccac tttgaaaagg 300
agattgatga tgcattggag aaaataggtc atgacccttt tgatgataaa gatgatctct 360
acattgtctc tctttgtttt cgattgctga ggcagcatgg aattaagata tcatgtgatg 420
tgtttgagaa gtttaagat gacgatggaa aattcaaggc atcattgatg aatgatgttc 480
aaggcatgct aagtttatat gaggcagcac acctagccat tcacggagaa gatattttag 540
atgaagcaat tgttttcacg accactcacc ttaagtcaac ggtatcta tctcctgtaa 600
actctacttt tgctgaacaa atacgtcatt ctctcagagt tcctctcctg aaagctgtac 660
ctaggttaga gtcgaggtat ttcttgata tctattcaag agatgatttg cacgataaaa 720
ctttgctcaa tttcgcaaag tcagacttta atatactaca agcaatgcac cagaaggaag 780
caagtgagat gaccaggtgg tggagagatt ttgacttctt taaaagctg ccttatataa 840
gagacagagt cgtggagcta ttttttggg ttctggtggg agtgtcttat cagcccaaat 900
tcagcactgg tagaattttt ttgtccaaaa taatagcct tgagaccctc gtagatgata 960
catttgacgc ctacggtact tttgacgagc tcacaatctt tactgaagca gttacaagat 1020
gggacattgg ccacagagat gcaactaccag aatacatgaa attcattttc aagacactca 1080
ttgatgtcta tagtgaagct g 1101

<210> 28
<211> 1875
<212> ADN
<213> Clausena lansium

5

<400> 28

ctagtagcta aaaaaattat taagttcaat cttttttgtc tgtctagaga aagatgtcaa 60
ctcaacaagt ttcacagag aacattgttc gtaatgctgc agatttccat cctaataat 120
gggaaacca tttcctcaca tgtctttctc aaacgattga tagttggact caacagcacc 180
aaaagaact gaaagaagag gtgaggaaaa tgatggtgtc tgatgcaaat aaacctgccc 240
agagattgag cttgattgat actgtccaaa ggtaggtgt ggcttaccac tttgaaaagg 300
agattgatga tgcattggag aaaataggtc atgacccttt tgatgataaa gatgatctct 360
acattgtctc tctttgtttt cgattgctga ggcagcatgg aattaagata tcatgtgatg 420

10

ES 2 410 906 T3

tgtttgagaa gtttaaagat gacgatggaa aattcaaggc atcattgatg aatgatgttc 480
 aaggcatgct aagtttatat gaggcagcac acctagccat tcacggagaa gatattttag 540
 atgaagcaat tgttttcacg accactcacc ttaagtcaac ggtatctaat tctcctgtaa 600
 actctacttt tgctgaacaa atacgtcatt ctctcagagt tcctctccgt aaagctgtac 660
 ctaggttaga gtcgaggtat ttcttgata tctattcaag agatgatttg cacgataaaa 720
 ctttgctcaa tttcgcaaag tcagacttta atatactaca agcaatgcac cagaaggaag 780
 caagtgagat gaccaggtgg tggagagatt ttgacttcct taaaaagctg ccttatataa 840
 gagacagagt cgtggagcta tatttttggga ttctgggtggg agtgtcttat cagcccaaat 900
 tcagcactgg tagaattttt ttgtccaaaa taatatgcct tgagaccctc gtagatgata 960
 catttgacgc ctacggtact tttgacgagc tcacaatctt tactgaagca gttacaagat 1020
 gggacattgg ccacagagat gcactaccag aatacatgaa attcattttc aagacactca 1080
 ttgatgtcta cagtgaagct gagcaagaac tggcaaagga agggagatca tacagcatac 1140
 aatatgcaat acgatcgttc caagaactag ttatgaagta cttctgcgaa gccaaagtgg 1200
 taaataaagg ttatgttccg agcctggacg attataaatc agtttcatta agaagtatcg 1260
 gttttttacc gatagcggta gcttccctcg ttttcatggg tgatattgca actaaggagg 1320
 tctttgaatg ggaaatgaat aaccctaaga tcataatagc cgcagaaacg attttcagat 1380
 tcttgatga catagcaggc cataagtttg agcaaaagag agaacatagt ccatcagcta 1440
 ttgaatgcta caagaatcaa catggagtgt ctgaggaaga ggcagttaaa gcggttgcgt 1500
 tagaagttgc taatagttgg aaagatataa atgaggagct gcttctcaac ccaatggcta 1560
 ttcttttacc tctgcttcag gtgattcttg atctctcacg ttcggcogat tttatgtacg 1620
 gtaatgctca agatcgcttc acgcattcaa cgatgatgaa agaccaagtt gatttggtgc 1680
 tgaaggaccc cgttaagctt gacgattaaa gttatgttgc tgatttctta tcgtatattt 1740
 gagaagttgg taataaatta agttgggtgct tgctagttat ttagctagct agtcatgctg 1800
 agctagggga tggttcaatt gattaggcct atattctagt aaaaataaac gatgtaagaa 1860
 caaatctccc tcgca 1875

<210> 29
 <211> 551
 <212> PRT
 <213> Clausena lansium

<400> 29

Met Ser Thr Gln Gln Val Ser Ser Glu Asn Ile Val Arg Asn Ala Ala
 1 5 10 15

5

10

ES 2 410 906 T3

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

Gln Thr Ile Asp Ser Trp Thr Gln Gln His His Lys Glu Leu Lys Glu
 35 40 45

Glu Val Arg Lys Met Met Val Ser Asp Ala Asn Lys Pro Ala Gln Arg
 50 55 60

Leu Arg Leu Ile Asp Thr Val Gln Arg Leu Gly Val Ala Tyr His Phe
 65 70 75 80

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

Asp Asp Lys Asp Asp Leu Tyr Ile Val Ser Leu Cys Phe Arg Leu Leu
 100 105 110

Arg Gln His Gly Ile Lys Ile Ser Cys Asp Val Phe Glu Lys Phe Lys
 115 120 125

Asp Asp Asp Gly Lys Phe Lys Ala Ser Leu Met Asn Asp Val Gln Gly
 130 135 140

Met Leu Ser Leu Tyr Glu Ala Ala His Leu Ala Ile His Gly Glu Asp
 145 150 155 160

Ile Leu Asp Glu Ala Ile Val Phe Thr Thr Thr His Leu Lys Ser Thr
 165 170 175

Val Ser Asn Ser Pro Val Asn Ser Thr Phe Ala Glu Gln Ile Arg His
 180 185 190

Ser Leu Arg Val Pro Leu Arg Lys Ala Val Pro Arg Leu Glu Ser Arg
 195 200 205

Tyr Phe Leu Asp Ile Tyr Ser Arg Asp Asp Leu His Asp Lys Thr Leu
 210 215 220

Leu Asn Phe Ala Lys Ser Asp Phe Asn Ile Leu Gln Ala Met His Gln
 225 230 235 240

Lys Glu Ala Ser Glu Met Thr Arg Trp Trp Arg Asp Phe Asp Phe Leu
 245 250 255

ES 2 410 906 T3

Lys Lys Leu Pro Tyr Ile Arg Asp Arg Val Val Glu Leu Tyr Phe Trp
260 265 270

Ile Leu Val Gly Val Ser Tyr Gln Pro Lys Phe Ser Thr Gly Arg Ile
275 280 285

Phe Leu Ser Lys Ile Ile Cys Leu Glu Thr Leu Val Asp Asp Thr Phe
290 295 300

Asp Ala Tyr Gly Thr Phe Asp Glu Leu Thr Ile Phe Thr Glu Ala Val
305 310 315 320

Thr Arg Trp Asp Ile Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys
325 330 335

Phe Ile Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu
340 345 350

Leu Ala Lys Glu Gly Arg Ser Tyr Ser Ile Gln Tyr Ala Ile Arg Ser
355 360 365

Phe Gln Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn
370 375 380

Lys Gly Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg
385 390 395 400

Ser Ile Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly
405 410 415

Asp Ile Ala Thr Lys Glu Val Phe Glu Trp Glu Met Asn Asn Pro Lys
420 425 430

Ile Ile Ile Ala Ala Glu Thr Ile Phe Arg Phe Leu Asp Asp Ile Ala
435 440 445

Gly His Lys Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu
450 455 460

Cys Tyr Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala
465 470 475 480

Leu Ser Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu
485 490 495

Leu Leu Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu

ES 2 410 906 T3

500

505

510

Asp Leu Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg
515 520 525

Phe Thr His Ser Thr Met Met Lys Asp Gln Val Asp Leu Val Leu Lys
530 535 540

Asp Pro Val Lys Leu Asp Asp
545 550

5 <210> 30
<211> 1368
<212> ADN
<213> Clausena lansium

10 <220>
<221> misc_feature
<222> (1346)..(1346)
<223> n e s a, c, g o t

15 <220>
<221> misc_feature
<222> (1357)..(1357)
<223> n e s a, c, g o t

20 <220>
<221> misc feature
<222> (1362)..(1362)
<223> n e s a, c, g o t

<400> 30

ES 2 410 906 T3

ttaagtcaat ggtatctaact tctcttgtaa actctacttt tgctgaacaa atacgtcatt 60
ctctcagagt tctctccat aaagccttac ctaggttaga atcgaggtat ttcttgata 120
tctattcaag agacgatttg cacgataaaa ctttgctcaa ttctcggaag ttagacttta 180
atatactaca agtaatgcac cagaaggaag caagtggat gaccaggtgg tggagagatt 240
ttgacttctt taaaaagctg cttatataa gagacagagt cgtggagcta ttttttgga 300
ttctggtggg agtgtcttat cagcccaaat tcagcactgg tagaattttt ttgtccaaaa 360
taatatgcct tgagaccctc gtagatgata catttgacgc ctacggtact tttgacgagc 420
tcacaatctt tactgaagca gttacaagat gggacattgg ccacagagat gcactaccag 480
aatacatgaa attcattttc aagacactca ttgatgtcta cagtgaagct gagcaagaac 540
tggcaaagga agggagatca tacagcatac aatatgcaat acggtcgttc caagaactag 600
ttatgaagta cttctgcgaa gccaaagtgg taaataaagg ttatgttccg agcctggagc 660
attataaatc agtttcatta agaagtatcg gttttttacc gatagcggta gcttccttcg 720

ttttcatggg tgatattgca actaaggagg tctttgaatg ggaaatgaat aaccctaaga 780
tcataatagc cgcagaaacg attttcagat tcttgatga catagcaggc cataagtttg 840
agcaaaagag agaacatagt ccatcagcta ttgaatgcta caagaatcaa catggagtgt 900
ctgaggaaga ggcagttaaa gcgttgcgt tagaagttgc taatagttgg aaagatataa 960
atgaggagct gcttctcaac ccaatggcta ttctttacc tctgcttcag gtgattcttg 1020
atctctcagc ttcgccgat tttatgtacg gtaatgctca agatcgcttc acgcattcaa 1080
cgatgatgaa agaccaagtt gatttggc tgaaggacc cgttaagctt gacgattaaa 1140
gttatgttgc tgatttcta tcgtatattt gagaagttgg taataaatta agttggtgct 1200
tgctagtat ttagctagct agtcatgct agctagggga tggttcaatt gattagcct 1260
atattctagt aaaaataaac gatgtaagaa caaatctccc tcgcaccaac ttcgcaataa 1320
tgtaatttat ttcattatg tctatngcag gggtcanaac cnaaaaaa 1368

5 <210> 31
<211> 1372
<212> ADN
<213> Clausena lansium

10 <220>
<221> misc_feature
<222> (1356)..(1356)
<223> n es a, c g o t

15 <220>
<221> misc_feature
<222> (1366)..(1366)
<223> n es a, c g o t

<400> 31

ES 2 410 906 T3

ttaagtcaac ggtatcctaat tctcctgtaa actctacttt tgccgaacaa atacgtcatt 60
 ctctcagagt tcctctccgt aaagctgtac ctagggtaga gtcgaggtat ttcttgata 120
 tctattcaag agatgatttg cacgataaaa ctttgctcaa ttctgcaaag ttagacttta 180
 atatactaca agcaatgcac cagaaggaag caagtgagat aaccaggtgg tggagagatt 240
 ttggattcct tgaaaagctg ccttatgtaa gagacagaat cgtggagata tatttttgga 300
 tattggtggg atggtcctat gagccaaaat tcagcactgg tagaatcatt ttgtccaaaa 360
 tattatgcct cgtgtccctt gtagatgata ctttgacgc ctatggtact cttgaagagc 420
 tcacagtctt tactgaagca attacaagat gggacattgg ccacacagat gcaactaccag 480
 attacatgaa attccttttc aagacactca ttgatgtcta tagtgaagct gaggaagaac 540
 tggcaaaggg aggaagatca tacagcatac aatatgcaat acgatcgttt caagaactag 600
 ctatgaaata cttctgcgaa gcggagtggg taaataaagg ttatgttccg agcctggacg 660

 agtataaatc agtttcagta agaagtgtcg gttttttcc gatagcggta gcttccttcg 720
 ttttcatggg tgatattaca actaaggagg tctttgaatg ggaaatgaat aaccctaaga 780
 tcataatagc cgcagaaaacg attttcagat tcctggatga tgtggcaggc cataagtttg 840
 agcaaaagag agaacattgt ccatcagcta ttgaatgcta caagaatcaa catggagtgt 900
 ctgaggaaga ggcagttaaa gcgttgtcgt tagaagttgc taatagttgg aaagatataa 960
 atgaggagct gcttctcaac ccaatggcta ttcctttacc tctacttcag gtgattcttg 1020
 atctctcagc ttcggcogat tttatgtacg gcaatggcca agatcgctac acgcattcaa 1080
 cgatgatgaa agaccaagtt gacttgggtc tgaaggaccc cgtaagctt gacgattaaa 1140
 gttatgttcg tgatttccta ttgtatattt gagaagttgg taataaatta agttgggtcct 1200
 tgctagtatt ttagctagct agtcatcgt agctaaggga tggttcaatt gattaggcct 1260
 atattctagt aaaaataaaa ggtgtaagaa cgaatctccc tcacaccaac ttcgcaataa 1320
 tgtaatttat ttcattatg tctgttacia aaatngaga taaaanaaca gc 1372

5 <210> 32
 <211> 31
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Cebador

<400> 32
 caccatgtca actcaacaag ttcatcaga g 31

15 <210> 33
 <211> 27
 <212> ADN
 <213> Secuencia artificial

ES 2 410 906 T3

<220>
 <223> Cebador

 5 <400> 33
 actttaatcg tcaagcttaa cggggtc 27

 <210> 34
 <211> 30
 <212> ADN
 10 <213> Artificial

 <220>
 <223> Cebador

 15 <400> 34
 ctgacctgg ctcagaaaa agaaattagg 30

 <210> 35
 <211> 40
 20 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador

 25 <400> 35
 ccggaattcc tattgcttc tctgtaaac ttgttcaag 40

 <210> 36
 <211> 42
 30 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador

 35 <400> 36
 aaggagatat acatatgaca aaaaaagttg gtgtcggta gg 42

 40 <210> 37
 <211> 43
 <212> ADN
 <213> Artificial

 45 <220>
 <223> Cebador

 <400> 37
 50 cttaccaga ctgagttac gccttttca tctgatcctt tgc 43

 <210> 38
 <211> 1656
 <212> ADN
 <213> Clausena lansium

 55 <400> 38

ES 2 410 906 T3

atgtcaactc aacaagtttc atcagagaac attgttcgta acgctgcgaa tttccatcct 60
aatatatggg gaaaccatth cctcacatgt ccttctcaga cgattgatag ttggactcaa 120
cagcaccaca aagaactgaa agaagaggtg aggaaaatga tgggtgtctga tgcaaataaa 180
cctgcccaga gattgcgctt gattgatact gtccaaaggt taggtgtggc ttaccacttt 240
gaaaaggaga ttgatgatgc attggagaaa ataggatcatg acccttttga tgataaagat 300
gatctctaca ttgtctctct ttgttttcga ttgctgaggg agcatggaat taagatatca 360
tgtgatgtgt ttgagaagtt taaagatgac gatggaaaat tcaaggcatc attgatgaat 420
gatgttcaag gcatgctaag tttatatgag gcagcacacc tagccattca cggagaagat 480
atthtagatg aagcaattgt tttcacgacc actcacctta agtcaacggg atctaattct 540
cctgtaaact ctacttttgc tgaacaaata cgtcattctc tcagagttcc tctccgtaaa 600
gctgtacctt ggttagagtc gaggtatttc ttggatatct attcaagaga tgatttgac 660
gataaaactt tgctcaatth cgcaaagtta gactttaata tactacaagc aatgcaccag 720
aaggaagcaa gtgagatgac caggtggtgg agagatthtg acttccttaa aaagctgcct 780
tatataagag acagagtcgt ggagctatat thttggattc tgggtgggagt gtcttatcag 840
cccaaattca gcaactggtag aathththtg tccaaaataa tatgccttga gaccctcgta 900
gatgatacat ttgacgccta cggtaactth gacgagctca caatctthac tgaagcagtt 960
acaagatggg acattggcca cagagatgca ctaccagaat acatgaaatt caththcaag 1020
acactcattg atgtctacag tgaagctgag caagaactgg caaaggaagg gagatcatac 1080
agcatacaat atgcaatacg atcgttccaa gaactagtta tgaagtactt ctgcgaagcc 1140
aagtggthaa ataaaggtta tgttccgagc ctggacgatt ataaatcagt thcattaaga 1200
agtatcggth thttaccgat agcggtagct tcttcgtht tcatgggtga tattgcaact 1260
aaggaggtct ttgaatggga aatgaataac cctaagatca taatagccgc agaaacgatt 1320
thcagattcc tggatgacat agcagccat aagthtgagc aaaagagaga acatagtcca 1380
tcagctattg aatgctacaa gaatcaacat ggagtgtctg aggaagaggc agthaaagcg 1440
thgtcgthtag aagthgctaa tagthgaaa gatataaatg aggagctgct tctcaaccca 1500
atggctattc thttacctct gcttcagtg attcttgatc tctcacgthc ggccgathth 1560
atgtacggta atgctcaaga tgcctcagc cattcaacga tgatgaaaga ccaagthgat 1620
thgggtgctga aggaccccgth taagcttgac gattag 1656

<210> 39
<211> 1656
<212> ADN
<213> *Clausena lansium*
<400> 39

5

10

ES 2 410 906 T3

atgtcaactc aacaagtttc atcagagaac attgttcgta acgctgcgaa ttccatcct 60
aatatatggg gaaaccattt cctcacatgt ccttctcaga cgattgatag ttggactcaa 120
cagcaccaca aagaactgaa agaagaggtg aggaaaatga tgggtgtctga tgcaaataaa 180
cctgcccaga gattgogctt gattgatact gtccaaaggt taggtgtggc ttaccacttt 240
gaaaaggaga ttgatgatgc attggagaaa ataggtcatg acccttttga tgataaagat 300
gatctctaca ttgtctctct ttgttttcga ttgctgaggc agcatggaat taagatatca 360
tgtgatgtgt ttgagaagtt taaagatgac gatggaaaat tcaaggcatc attgatgaat 420
gatgttcaag gcatgctaag tttatatgag gcagcacacc tagccattca cggagaagat 480
atthtagatg aagcaattgt tttcaogacc actcacctta agtcaacggt atctaattct 540
cctgtaaact ctacttttgc tgaacaaata cgtcattctc tcagagttcc tctccgtaaa 600
gctgtacctg ggtagagtc gaggtatttc ttggatatct attcaagaga tgatttgcac 660
gataaaactt tgctcaattt cgcaaagtta gactttaata tactacaagc aatgcaccag 720
aaggaagcaa gtgagatgac caggtggtgg agagattttg acttccttaa aaagctgcct 780
tatataagag acagagtcgt ggagctatat ttttggatc tgggtgggagt gtcttatcac 840
cccaaattca gcaactgtag aatttttttg tccaaaataa tatgccttga gaccctcgta 900
gatgatacat ttgacgccta cggfactttt gacgagctca caatctttac tgaagcagtt 960
acaagatggg acattggcca cagagatgca ctaccagaat acatgaaatt cattttcaag 1020
acactcattg atgtctacag tgaagctgag caagaactgg caaaggaagg gagatcatic 1080
agcatacaat atgcaatacg atcgttccaa gaactagtta tgaagtactt ctgogaagcc 1140
aagtgggtaa ataaaggta tgttccgagc ctggacgatt ataaatcagt ttcattaaga 1200
agtatcgggt ttttaccgat agcggtagct tccttcgttt tcatgggtga tattgcaact 1260
aaggaggtct ttgaatggga aatgaataac cctaagatca taatagccgc agaaacgatt 1320
ttcagattcc tggatgacat agcaggccat aagtttgagc aaaagagaga acatagtcca 1380
tcagctattg aatgctacaa gaatcaacat ggagtgctct aggaagaggc agttaaagcg 1440
ttgtcgttag aagttgctaa tagttggaaa gatataaatg aggagctgct tctcaacca 1500
atggctatc ctttacctct gcttcagggt attcttgatc tctcacgttc ggccgatttt 1560
atgtacggta atgctcaaga tcgcctcacg cattcaacga tgatgaaaga ccaagttgat 1620
ttgggtgctga aggaccccgt taagcttgac gattag 1656

<210> 40
<211> 1656
<212> ADN
<213> Clausena lansium

<400> 40

ES 2 410 906 T3

atgtcaactc aacaagtttc atcagagaac attgttcgta acgctgcgaa tttccatcct 60
 aatatatggg gaaaccattt cctcacatgt ccttctcaga cgattgatag ttggactcaa 120
 cagcaccaca aagaactgaa agaagaggtg aggaaaatga tgggtgtctga tgcaaataaa 180
 cctgcccaga gattgcgctt gattgatact gtccaaaggc taggtgtggc ttaccacttt 240
 gaaaaggaga ttgatgatgc attggagaaa ataggtcatg acccttttga tgataaagat 300
 gatctctaca ttgtctctct ttgttttcga ttgctgaggc agcatggaat taagatatca 360
 tgtgatgtgt ttgagaagtt taaagatgac gatggaaaat tcaaggcatc attgatgaat 420
 gatgttcaag gcatgctaag tttatatgag gcagcacacc tagccattca cggagaagat 480
 attttagatg aagcaattgt tttcacgacc actcacctta agtcaacggg atetaattct 540
 cctgtaaact ctacttttgc tgaacaaata cgtcattctc tcagagttcc tctccgtaaa 600
 gctgtacctt ggtagagtc gaggtatttc ttggatatct attcaagaga tgatttgcac 660
 gataaaactt tgctcaattt cgcaaagtta gactttaata tactacaagc aatgcaccag 720
 aaggaagcaa gtgagatgac cagggtgtgg agagattttg acttocttaa aaagctgcct 780
 tatataagag acagagtcgt ggagctatat ttttggattc tgggtgggagt gtcttatcag 840
 cccaaattca gcaactggtg aattttttgc tccaaaataa tatgccttga gaccctcgta 900
 gatgatacat ttgacgccta cggtaacttt gacgagctca caatctttac tgaagcagtt 960
 acaagatggg acattggcca cagagatgca ctaccagaat acatgaaatt cattttcaag 1020
 aactcattg atgtctacag tgaagctgag caagaactgg caaaggaagg gagatcatac 1080
 agcatacaat atgcaatacg atcgttccaa gaactagtta tgaagtactt ctgcgaagcc 1140
 aagtggttaa ataaaggtta tgttccgagc ctggacgatt ataaatcagt ttcattaaga 1200
 agtatcgggt ttttaccgat agcggtagct tccttcggtt tcatgggtga tattgcaact 1260
 aaggaggtct ttgaatggga aatgaataac cctaagatca taatagccgc agaaacgatt 1320
 ttcagattcc tggatgacat agcaggccat aagtttgagc aaaagagaga acatagtcca 1380
 tcagctattg aatgctacaa gaatcaacat ggagtgtctg aggaagaggc agttaaagcg 1440
 ttgtcgttag aagttgctaa tagttggaaa gatataaatg aggagctgct tctcaacca 1500
 atggctatcc ctttacctct gcttcaggty attcttgatc tctcacgttc ggccgatttt 1560
 atgtacggta atgctcaaga tcgcttcacg cattcaacga tgatgaaaga ccaagttgat 1620
 ttgggtgctga aggaccccg taaagcttgac gattaa 1656

<210> 41
 <211> 551
 <212> PRT
 <213> Clausena lansium

ES 2 410 906 T3

<400> 41

Met Ser Thr Gln Gln Val Ser Ser Glu Asn Ile Val Arg Asn Ala Ala
1 5 10 15

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

Gln Thr Ile Asp Ser Trp Thr Gln Gln His His Lys Glu Leu Lys Glu
35 40 45

ES 2 410 906 T3

Glu Val Arg Lys Met Met Val Ser Asp Ala Asn Lys Pro Ala Gln Arg
 50 55 60

Leu Arg Leu Ile Asp Thr Val Gln Arg Leu Gly Val Ala Tyr His Phe
 65 70 75 80

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

Asp Asp Lys Asp Asp Leu Tyr Ile Val Ser Leu Cys Phe Arg Leu Leu
 100 105 110

Arg Gln His Gly Ile Lys Ile Ser Cys Asp Val Phe Glu Lys Phe Lys
 115 120 125

Asp Asp Asp Gly Lys Phe Lys Ala Ser Leu Met Asn Asp Val Gln Gly
 130 135 140

Met Leu Ser Leu Tyr Glu Ala Ala His Leu Ala Ile His Gly Glu Asp
 145 150 155 160

Ile Leu Asp Glu Ala Ile Val Phe Thr Thr Thr His Leu Lys Ser Thr
 165 170

Val Ser Asn Ser Pro Val Asn Ser Thr Phe Ala Glu Gln Ile Arg His
 180 185 190

Ser Leu Arg Val Pro Leu Arg Lys Ala Val Pro Arg Leu Glu Ser Arg
 195 200 205

Tyr Phe Leu Asp Ile Tyr Ser Arg Asp Asp Leu His Asp Lys Thr Leu
 210 215 220

Leu Asn Phe Ala Lys Leu Asp Phe Asn Ile Leu Gln Ala Met His Gln
 225 230 235 240

Lys Glu Ala Ser Glu Met Thr Arg Trp Trp Arg Asp Phe Asp Phe Leu
 245 250 255

Lys Lys Leu Pro Tyr Ile Arg Asp Arg Val Val Glu Leu Tyr Phe Trp
 260 265 270

Ile Leu Val Gly Val Ser Tyr Gln Pro Lys Phe Ser Thr Gly Arg Ile
 275 280 285

ES 2 410 906 T3

Phe Leu Ser Lys Ile Ile Cys Leu Glu Thr Leu Val Asp Asp Thr Phe
 290 295 300

Asp Ala Tyr Gly Thr Phe Asp Glu Leu Thr Ile Phe Thr Glu Ala Val
 305 310 315 320

Thr Arg Trp Asp Ile Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys
 325 330 335

Phe Ile Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu
 340 345 350

Leu Ala Lys Glu Gly Arg Ser Tyr Ser Ile Gln Tyr Ala Ile Arg Ser
 355 360 365

Phe Gln Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn
 370 375 380

Lys Gly Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg
 385 390 395 400

Ser Ile Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly
 405 410 415

Asp Ile Ala Thr Lys Glu Val Phe Glu Trp Glu Met Asn Asn Pro Lys
 420 425 430

Ile Ile Ile Ala Ala Glu Thr Ile Phe Arg Phe Leu Asp Asp Ile Ala
 435 440 445

Gly His Lys Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu
 450 455 460

Cys Tyr Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala
 465 470 475 480

Leu Ser Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu
 485 490 495

Leu Leu Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu
 500 505 510

Asp Leu Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg
 515 520 525

Leu Thr His Ser Thr Met Met Lys Asp Gln Val Asp Leu Val Leu Lys

ES 2 410 906 T3

530

535

540

Asp Pro Val Lys Leu Asp Asp
545 550

5

<210> 42
<211> 551
<212> PRT
<213> Clausena lansium

<400> 42

Met Ser Thr Gln Gln Val Ser Ser Glu Asn Ile Val Arg Asn Ala Ala
1 5 10 15

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

Gln Thr Ile Asp Ser Trp Thr Gln Gln His His Lys Glu Leu Lys Glu
35 40 45

Glu Val Arg Lys Met Met Val Ser Asp Ala Asn Lys Pro Ala Gln Arg
50 55 60

Leu Arg Leu Ile Asp Thr Val Gln Arg Leu Gly Val Ala Tyr His Phe
65 70 75 80

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

Asp Asp Lys Asp Asp Leu Tyr Ile Val Ser Leu Cys Phe Arg Leu Leu
100 105 110

Arg Gln His Gly Ile Lys Ile Ser Cys Asp Val Phe Glu Lys Phe Lys
115 120 125

Asp Asp Asp Gly Lys Phe Lys Ala Ser Leu Met Asn Asp Val Gln Gly
130 135 140

Met Leu Ser Leu Tyr Glu Ala Ala His Leu Ala Ile His Gly Glu Asp
145 150 155 160

Ile Leu Asp Glu Ala Ile Val Phe Thr Thr Thr His Leu Lys Ser Thr
165 170 175

Val Ser Asn Ser Pro Val Asn Ser Thr Phe Ala Glu Gln Ile Arg His
180 185 190

10

ES 2 410 906 T3

Ser Leu Arg Val Pro Leu Arg Lys Ala Val Pro Arg Leu Glu Ser Arg
 195 200 205

Tyr Phe Leu Asp Ile Tyr Ser Arg Asp Asp Leu His Asp Lys Thr Leu
 210 215 220

Leu Asn Phe Ala Lys Leu Asp Phe Asn Ile Leu Gln Ala Met His Gln
 225 230 235 240

Lys Glu Ala Ser Glu Met Thr Arg Trp Trp Arg Asp Phe Asp Phe Leu
 245 250 255

Lys Lys Leu Pro Tyr Ile Arg Asp Arg Val Val Glu Leu Tyr Phe Trp
 260 265 270

Ile Leu Val Gly Val Ser Tyr His Pro Lys Phe Ser Thr Gly Arg Ile
 275 280 285

Phe Leu Ser Lys Ile Ile Cys Leu Glu Thr Leu Val Asp Asp Thr Phe
 290 295 300

Asp Ala Tyr Gly Thr Phe Asp Glu Leu Thr Ile Phe Thr Glu Ala Val
 305 310 315 320

Thr Arg Trp Asp Ile Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys
 325 330 335

Phe Ile Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu
 340 345 350

Leu Ala Lys Glu Gly Arg Ser Tyr Ser Ile Gln Tyr Ala Ile Arg Ser
 355 360 365

Phe Gln Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn
 370 375 380

Lys Gly Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg
 385 390 395 400

Ser Ile Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly
 405 410 415

Asp Ile Ala Thr Lys Glu Val Phe Glu Trp Glu Met Asn Asn Pro Lys
 420 425 430

ES 2 410 906 T3

Ile Ile Ile Ala Ala Glu Thr Ile Phe Arg Phe Leu Asp Asp Ile Ala
435 440 445

Gly His Lys Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu
450 455 460

Cys Tyr Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala
465 470 475 480

Leu Ser Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu
485 490 495

Leu Leu Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu
500 505 510

Asp Leu Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg
515 520 525

Leu Thr His Ser Thr Met Met Lys Asp Gln Val Asp Leu Val Leu Lys
530 535 540

Asp Pro Val Lys Leu Asp Asp
545 550

- <210> 43
- <211> 551
- <212> PRT
- <213> Clausena lansium
- <400> 43

5

ES 2 410 906 T3

Met Ser Thr Gln Gln Val Ser Ser Glu Asn Ile Val Arg Asn Ala Ala
1 5 10 15

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

Gln Thr Ile Asp Ser Trp Thr Gln Gln His His Lys Glu Leu Lys Glu
35 40 45

Glu Val Arg Lys Met Met Val Ser Asp Ala Asn Lys Pro Ala Gln Arg
50 55 60

Leu Arg Leu Ile Asp Thr Val Gln Arg Leu Gly Val Ala Tyr His Phe
65 70 75 80

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

ES 2 410 906 T3

Asp Asp Lys Asp Asp Leu Tyr Ile Val Ser Leu Cys Phe Arg Leu Leu
 100 105 110
 Arg Gln His Gly Ile Lys Ile Ser Cys Asp Val Phe Glu Lys Phe Lys
 115 120 125
 Asp Asp Asp Gly Lys Phe Lys Ala Ser Leu Met Asn Asp Val Gln Gly
 130 135 140
 Met Leu Ser Leu Tyr Glu Ala Ala His Leu Ala Ile His Gly Glu Asp
 145 150 155 160
 Ile Leu Asp Glu Ala Ile Val Phe Thr Thr Thr His Leu Lys Ser Thr
 165 170 175
 Val Ser Asn Ser Pro Val Asn Ser Thr Phe Ala Glu Gln Ile Arg His
 180 185 190
 Ser Leu Arg Val Pro Leu Arg Lys Ala Val Pro Arg Leu Glu Ser Arg
 195 200 205
 Tyr Phe Leu Asp Ile Tyr Ser Arg Asp Asp Leu His Asp Lys Thr Leu
 210 215 220
 Leu Asn Phe Ala Lys Leu Asp Phe Asn Ile Leu Gln Ala Met His Gln
 225 230 235 240
 Lys Glu Ala Ser Glu Met Thr Arg Trp Trp Arg Asp Phe Asp Phe Leu
 245 250 255
 Lys Lys Leu Pro Tyr Ile Arg Asp Arg Val Val Glu Leu Tyr Phe Trp
 260 265 270
 Ile Leu Val Gly Val Ser Tyr Gln Pro Lys Phe Ser Thr Gly Arg Ile
 275 280 285
 Phe Leu Ser Lys Ile Ile Cys Leu Glu Thr Leu Val Asp Asp Thr Phe
 290 295 300
 Asp Ala Tyr Gly Thr Phe Asp Glu Leu Thr Ile Phe Thr Glu Ala Val
 305 310 315 320
 Thr Arg Trp Asp Ile Gly His Arg Asp Ala Leu Pro Glu Tyr Met Lys
 325 330 335

ES 2 410 906 T3

Phe Ile Phe Lys Thr Leu Ile Asp Val Tyr Ser Glu Ala Glu Gln Glu
 340 345 350

Leu Ala Lys Glu Gly Arg Ser Tyr Ser Ile Gln Tyr Ala Ile Arg Ser
 355 360 365

Phe Gln Glu Leu Val Met Lys Tyr Phe Cys Glu Ala Lys Trp Leu Asn
 370 375 380

Lys Gly Tyr Val Pro Ser Leu Asp Asp Tyr Lys Ser Val Ser Leu Arg
 385 390 395 400

Ser Ile Gly Phe Leu Pro Ile Ala Val Ala Ser Phe Val Phe Met Gly
 405 410 415

Asp Ile Ala Thr Lys Glu Val Phe Glu Trp Glu Met Asn Asn Pro Lys
 420 425 430

Ile Ile Ile Ala Ala Glu Thr Ile Phe Arg Phe Leu Asp Asp Ile Ala
 435 440 445

Gly His Lys Phe Glu Gln Lys Arg Glu His Ser Pro Ser Ala Ile Glu
 450 455 460

Cys Tyr Lys Asn Gln His Gly Val Ser Glu Glu Glu Ala Val Lys Ala
 465 470 475 480

Leu Ser Leu Glu Val Ala Asn Ser Trp Lys Asp Ile Asn Glu Glu Leu
 485 490 495

Leu Leu Asn Pro Met Ala Ile Pro Leu Pro Leu Leu Gln Val Ile Leu
 500 505 510

Asp Leu Ser Arg Ser Ala Asp Phe Met Tyr Gly Asn Ala Gln Asp Arg
 515 520 525

Phe Thr His Ser Thr Met Met Lys Asp Gln Val Asp Leu Val Leu Lys
 530 535 540

Asp Pro Val Lys Leu Asp Asp
 545 550

REIVINDICACIONES

1. Un procedimiento de producción de α -santaleno que comprende
 - a) poner en contacto FPP (farnesil pirofosfato) con al menos un polipéptido que tiene una actividad α -santaleno sintasa y que comprende una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1;
 - b) opcionalmente, aislar el α -santaleno producido en la etapa a).
2. El procedimiento de la reivindicación 1, en el que la etapa a) comprende cultivar un organismo o célula huésped no humano capaz de producir FPP y transformado para expresar al menos un polipéptido que comprende una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1 y que tiene una actividad α -santaleno sintasa, en condiciones que conducen a la producción de α -santaleno.
3. El procedimiento de la reivindicación 2, en el que el procedimiento comprende además, antes de la etapa a), transformar un organismo o célula huésped no humano capaz de producir FPP con al menos un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1 y que tiene una actividad α -santaleno sintasa, de modo que dicho organismo exprese dicho polipéptido.
4. El procedimiento de la reivindicación 3, en el que el al menos un ácido nucleico que codifica la α -santaleno sintasa comprende una secuencia de nucleótidos al menos 50%, preferentemente al menos 70%, preferentemente al menos 90% idéntica a SEC ID N°: 2 o el complemento de la misma.
5. El procedimiento de la reivindicación 4, en el que el ácido nucleico comprende la secuencia de nucleótidos SEC ID N°: 2 o el complemento de la misma.
6. El procedimiento de la reivindicación 5, en el que el ácido nucleico consiste en SEC ID N°: 2 o el complemento de la misma.
7. El procedimiento de la reivindicación 3 o 4, en el que el al menos un ácido nucleico usado en cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que ha sido obtenida modificando SEC ID N°: 2 o el complemento de la misma.
8. El procedimiento de cualquiera de las reivindicaciones 2 a 7, en el que el organismo huésped no humano es una planta, un procarionta o un hongo.
9. El procedimiento de cualquiera de las reivindicaciones 2 a 7, en el que el organismo huésped no humano es un microorganismo, preferentemente una bacteria o una levadura.
10. El procedimiento de la reivindicación 9, en el que dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.
11. El procedimiento de cualquiera de las reivindicaciones 2 a 7, en el que la célula huésped no humana es una célula vegetal o una fúngica.
12. El procedimiento de cualquiera de las reivindicaciones precedentes, de producción de α -santaleno como un producto principal.
13. El procedimiento de la reivindicación 12, en el que el α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos obtenidos.
14. El procedimiento de cualquiera de las reivindicaciones precedentes, de producción de (+)- α -santaleno y en el que el polipéptido que tiene una actividad α -santaleno sintasa tiene una actividad (+)- α -santaleno sintasa.
15. El procedimiento de la reivindicación 14, en el que se obtiene (+)- α -santaleno como un producto principal.
16. El procedimiento de la reivindicación 15, en el que (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos obtenidos.
17. El procedimiento de cualquiera de las reivindicaciones precedentes, en el que el al menos un polipéptido comprende una secuencia de aminoácidos al menos 70%, preferentemente al menos 90% idéntica a SEC ID N°: 1.
18. El procedimiento de la reivindicación 17, en el que el al menos un polipéptido comprende la secuencia de aminoácidos SEC ID N°: 1.
19. El procedimiento de la reivindicación 18, en el que el al menos un polipéptido consiste en SEC ID N°: 1.
20. El procedimiento de cualquiera de las reivindicaciones 1 a 4 o 7 a 17, en el que el al menos un polipéptido comprende una secuencia de aminoácidos que ha sido obtenida modificando SEC ID N°: 1.

21. Un polipéptido que tiene una actividad α -santaleno sintasa y que comprende una secuencia de aminoácidos al menos 50% idéntica a SEC ID N°: 1.
22. El polipéptido de la reivindicación 21, en el que dicho polipéptido es capaz de producir α -santaleno como un producto principal.
- 5 23. El polipéptido de la reivindicación 22, en el que dicho polipéptido es capaz de producir una mezcla de sesquiterpenos en la que α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos producidos.
24. El polipéptido de la reivindicación 21, en el que dicho polipéptido tiene una actividad (+)- α -santaleno sintasa.
- 10 25. El polipéptido de la reivindicación 24, en el que dicho polipéptido es capaz de producir (+)- α -santaleno como un producto principal.
26. El polipéptido de la reivindicación 25, en el que dicho polipéptido es capaz de producir una mezcla de sesquiterpenos en la que (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos producidos.
- 15 27. El polipéptido de cualquiera de las reivindicaciones 21 a 26, en el que dicho polipéptido comprende una secuencia de aminoácidos al menos 70%, preferentemente al menos 90% idéntica a SEC ID N°: 1.
28. El polipéptido de la reivindicación 27, en el que dicho polipéptido comprende la secuencia de aminoácidos SEC ID N°: 1.
29. El polipéptido de la reivindicación 28, en el que dicho polipéptido consiste en SEC ID N°: 1.
- 20 30. El polipéptido de cualquiera de las reivindicaciones 21 a 27, en el que dicho polipéptido comprende una secuencia de aminoácidos que ha sido obtenida modificando SEC ID N°: 1.
31. Un ácido nucleico que codifica un polipéptido de acuerdo con cualquiera de las reivindicaciones 21 a 30.
32. El ácido nucleico de la reivindicación 31, que comprende una secuencia de nucleótidos al menos 50%, preferentemente al menos 70%, preferentemente al menos 90% idéntica a SEC ID N°: 2 o el complemento de la misma.
- 25 33. El ácido nucleico de la reivindicación 32, que comprende la secuencia de nucleótidos SEC ID N°: 2 o el complemento de la misma.
34. El ácido nucleico de la reivindicación 33, que consiste en SEC ID N°: 2 o el complemento de la misma.
35. El ácido nucleico de la reivindicación 31 o 32, que comprende una secuencia de nucleótidos que se ha obtenido modificando SEC ID N°: 2 o el complemento de la misma.
- 30 36. Un vector de expresión que comprende el ácido nucleico de cualquiera de las reivindicaciones 31 a 35.
37. El vector de expresión de la reivindicación 36, en forma de un vector viral, un bacteriófago o un plásmido.
38. El vector de expresión de la reivindicación 36 o 37, que incluye el ácido nucleico de la invención unido operativamente al menos a una secuencia reguladora que controla la transcripción, inicio o terminación de la traducción, tal como un promotor transcripcional, operador o potenciador o un sitio de unión ribosómico de ARNm y, opcionalmente, que incluye al menos un marcador de selección.
- 35 39. Un organismo o célula huésped no humano transformado para albergar al menos un ácido nucleico de acuerdo con cualquiera de las reivindicaciones 31 a 35, de modo que exprese de forma heteróloga o sobreexpresa al menos un polipéptido de acuerdo con cualquiera de las reivindicaciones 21 a 30.
- 40 40. El organismo huésped no humano de la reivindicación 39, en el que dicho organismo huésped no humano es una planta, un procarionta o un hongo.
41. El organismo huésped no humano de la reivindicación 39, en el que dicho organismo huésped no humano es un microorganismo, preferentemente una bacteria o levadura.
42. El organismo huésped no humano de la reivindicación 41, en el que dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.
- 45 43. La célula eucariota superior de la reivindicación 39, en la que dicha célula eucariota superior es una célula vegetal o una célula fúngica.
44. Un procedimiento de producción de al menos un polipéptido de acuerdo con cualquiera de las reivindicaciones

21 a 30 que comprende

- 5 a) cultivar un organismo o célula huésped no humano transformado con el vector de expresión de cualquiera de las reivindicaciones 36 a 38, de modo que albergue un ácido nucleico de acuerdo con cualquiera de las reivindicaciones 31 a 35 y exprese o sobreexpresa un polipéptido de acuerdo con cualquiera de las reivindicaciones 21 a 30;
- b) aislar el polipéptido del organismo o célula huésped no humano cultivado en la etapa a).
- 10 45. El procedimiento de la reivindicación 44, que comprende además, antes de la etapa a), transformar un organismo o célula huésped no humano con el vector de expresión de cualquiera de las reivindicaciones 36 a 38, de modo que albergue un ácido nucleico de acuerdo con cualquiera de las reivindicaciones 31 a 35 y exprese o sobreexpresa el polipéptido de cualquiera de las reivindicaciones 21 a 30.
46. Un procedimiento de preparación de un polipéptido variante que tiene actividad α -santaleno sintasa que comprende las etapas de:
- 15 (a) seleccionar un ácido nucleico de acuerdo con cualquiera de las reivindicaciones 31 a 35;
- (b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
- (c) transformar células u organismos unicelulares huésped con la secuencia de ácido nucleico mutante para expresar un polipéptido codificado por la secuencia de ácido nucleico mutante;
- (d) explorar el polipéptido con respecto a al menos una propiedad modificada; y
- (e) opcionalmente, si el polipéptido no tiene actividad α -santaleno sintasa variante deseada, repetir las etapas de procedimiento (a) a (d) hasta que se obtenga un polipéptido con una actividad α -santaleno sintasa variante deseada;
- 20 (f) opcionalmente, si se identificó un polipéptido que tenía una actividad α -santaleno sintasa variante deseada en la etapa (d), aislar el ácido nucleico mutante correspondiente obtenido en la etapa (c).
47. Un procedimiento de acuerdo con la reivindicación 46, en el que el polipéptido variante preparado es capaz de producir α -santaleno como un producto principal.
- 25 48. Un procedimiento de acuerdo con la reivindicación 47, en el que el polipéptido variante preparado es capaz de producir una mezcla de sesquiterpenos en la que α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos producidos.
49. Un procedimiento de acuerdo con la reivindicación 46, en el que el polipéptido variante preparado tiene una actividad (+)- α -santaleno sintasa.
- 30 50. Un procedimiento de acuerdo con la reivindicación 49, en el que el polipéptido variante preparado es capaz de producir (+)- α -santaleno como un producto principal.
51. Un procedimiento de acuerdo con la reivindicación 50, en el que el polipéptido variante preparado es capaz de producir una mezcla de sesquiterpenos en la que (+)- α -santaleno representa al menos el 60%, preferentemente al menos el 80%, preferentemente al menos el 90% de los sesquiterpenos producidos.

Figura 2

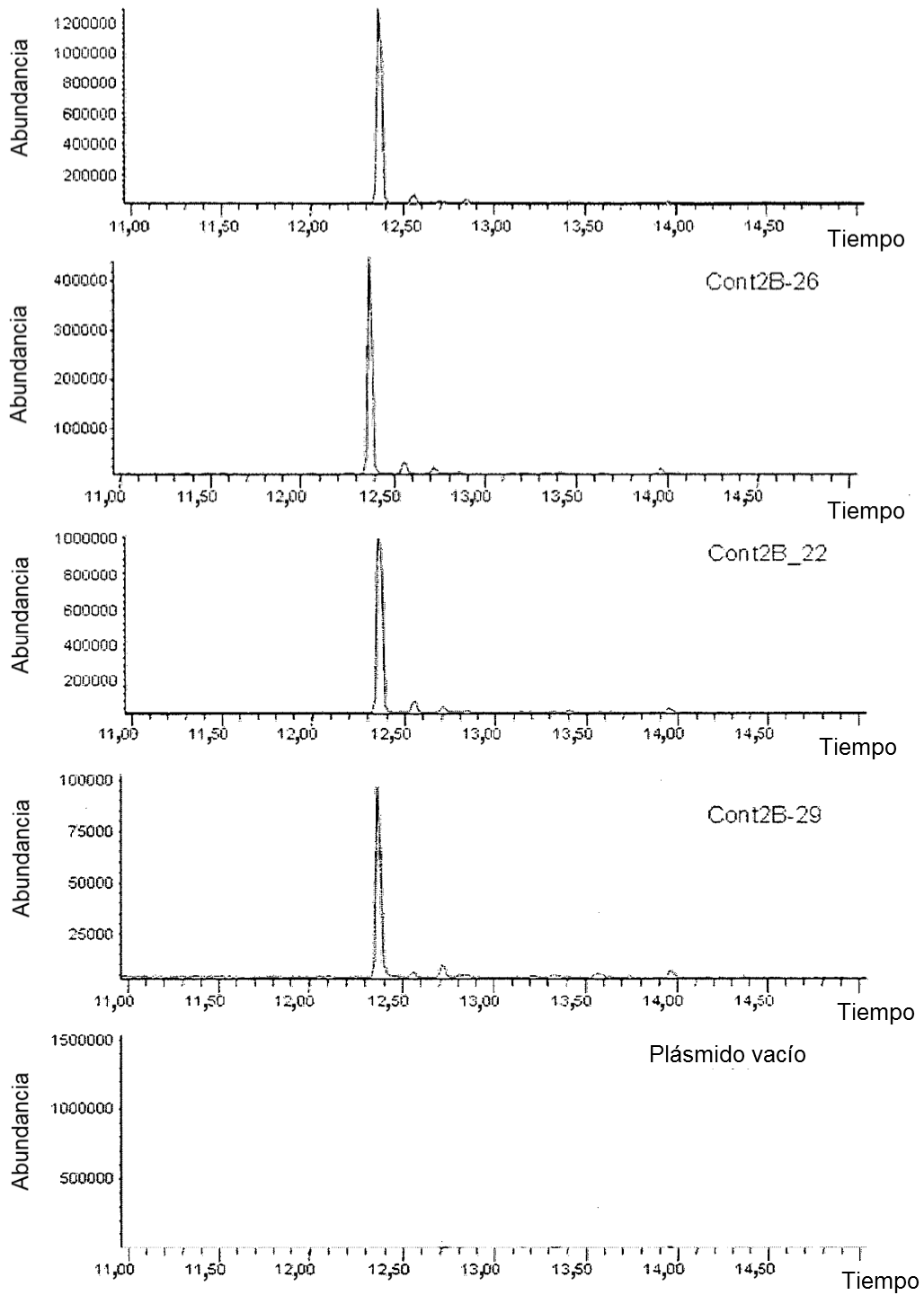


Figura 3

