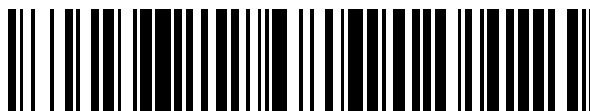


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 792 041**

51 Int. Cl.:

C12N 9/88

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.05.2011** **E 17199822 (2)**

97 Fecha y número de publicación de la concesión europea: **19.02.2020** **EP 3312276**

54 Título: **Procedimiento de producción de sesquiterpenos**

30 Prioridad:

10.05.2010 US 333002 P
10.05.2010 EP 10162433

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
06.11.2020

73 Titular/es:

FIRMENICH SA (100.0%)
7, Rue de la Bergère
1242 Satigny, CH

72 Inventor/es:

SCHALK, MICHEL y
DEGUERRY, FABIENNE

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 792 041 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 sesquiterpenos

Campo técnico

5 Se describe en el presente documento un procedimiento de producción de pachulol y 7-epi-a-selineno, comprendiendo dicho procedimiento poner en contacto al menos un polipéptido con pirofosfato de farnesilo (FPP). En particular, dicho procedimiento puede llevarse a cabo *in vitro* o *in vivo* para producir pachulol y 7-epi-a-selineno, compuestos que pueden ser útiles en el campo de la perfumería. La presente invención proporciona la secuencia de aminoácidos de un polipéptido útil en dicho procedimiento. 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 usar en el procedimiento de producción de pachulol y 7-epi-a-selineno también es un objeto de la presente invención.

Estado de la técnica

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

25 La producción biosintética de los terpenos implica a enzimas llamadas las terpeno sintasas. Existe prácticamente una infinidad de sesquiterpeno sintasas presentes en el reino de las plantas, todas usando el mismo sustrato (pirofosfato de farnesilo, FPP), pero que tienen diferentes perfiles de producto. Han sido clonados genes y ADNc que codifican sesquiterpeno sintasas y caracterizado las enzimas recombinantes correspondientes. La biosíntesis de los terpenos en plantas y otros organismos ha sido ampliamente estudiada y no se detalla más aquí.

30 Generalmente, el precio y la disponibilidad de los extractos naturales de planta dependen de la abundancia, rendimiento del aceite y origen geográfico de las plantas. Además, la disponibilidad y calidad de los extractos naturales es muy dependiente del clima y otras condiciones locales que conducen a la variabilidad de año a año, haciendo que el uso de tales componentes en la perfumería de alta calidad sea muy difícil o incluso imposible algunos años.

El aceite de valeriana es uno de estos extractos naturales. Es un aceite aromático, del que algunos componentes pueden ventajosamente usarse en el campo de la perfumería y los aromas. Sin embargo, la purificación de estos constituyentes individuales del aceite no es factible a gran escala.

35 Un procedimiento independiente de la planta de producción de los constituyentes del aceite de valeriana sería, por tanto, muy deseable, pero hasta la fecha no está disponible una síntesis química rentable de tales compuestos.

El pachulol y 7-epi-a-selineno son moléculas de sesquiterpeno que existen de forma natural, que son útiles en el campo de la perfumería y los aromas. Una vía bioquímica que conduce a la síntesis de pachulol y 7-epi-a-selineno sería, por tanto, de gran interés.

40 El análisis de la composición del aceite de *Valeriana jatamansi* mostró que el pachulol era el principal constituyente de este aceite. Véase, por ejemplo, Mathela C.S. y col. (2005), Chem. Biodivers. 2(9), 1174-1182 y Bos R. y col. Flav. Fragr. J. 12(2), 123-131. Sin embargo, este documento no proporciona ni incluso sugiere una secuencia de aminoácidos o de nucleótidos que conduzca a la producción de pachulol y 7-epi-a-selineno.

45 Se han descrito sesquiterpeno sintasas y procedimientos para su producción y uso en el documento WO 2004/031376. Una sesquiterpeno sintasa capaz de sintetizar pachulol ya ha sido descrita en el documento WO 2005/052163. Se ha informado una sesquiterpeno sintasa capaz de sintetizar 7-epi-a-selineno sintasa en Lucker y col. (2004), Phytochemistry 65, 2649-2659. Sin embargo, una sesquiterpeno sintasa capaz de sintetizar tanto pachulol como 7-epi-a-selineno nunca ha sido desvelada en el estado de la técnica.

50 El porcentaje de identidad entre las sesquiterpeno sintasas conocidas y el polipéptido de la invención es muy bajo. En particular, la pachulol sintasa descrita en el documento WO 2005/052163 es solo el 39 % idéntica a la sesquiterpeno sintasa de la invención. La secuencia de proteínas más próxima a la sesquiterpeno sintasa de la invención son una vetispiradieno sintasa de *Solanum tuberosus* (acceso de NCBI N.º AAD02269) y una proteína predicha de *Populus trichocarpa* (acceso de NCBI N.º XP_002321642 que tienen ambas el 49 % de identidad con la presente sesquiterpeno sintasa.

A pesar de los amplios estudios de ciclación de terpenos, el aislamiento y la caracterización de las terpeno sintasas es todavía difícil, particularmente en plantas, debido a su baja abundancia, sus patrones de expresión frecuentemente transitorios, y la complejidad de purificarlos de las mezclas de resinas y compuestos fenólicos en tejidos donde se expresan.

- 5 Es un objetivo de la presente invención proporcionar enzimas capaces de sintetizar pachulol y 7-epi- α -selineno que son útiles como ingredientes de perfumería y/o aroma.

Abreviaturas usadas

	pb	par de bases
	BSA	albúmina de suero bovino
10	ADNc	ADN complementario
	ADN	ácido desoxirribonucleico
	dNTP	desoxinucleótido trifosfato
	DTT	ditiotreitól
	FPP	pirofosfato de farnesilo
15	CG	cromatografía de gases
	IPTG	isopropil-D-tiogalacto-piranósido
	LB	caldo de lisogenia
	MOPSO	ácido 3-(N-morfolino)-2-hidroxiopropanosulfónico
	EM	espectrómetro de masas
20	PCR	reacción en cadena de la polimerasa
	RMCE	intercambio de casete mediado por recombinasa
	3'-/5' -RACE	amplificación rápida de los extremos de ADNc 3' y 5'
	ARN	ácido ribonucleico
	ARNm	ácido ribonucleico mensajero

Descripción de la invención

La presente invención proporciona polipéptidos que tienen actividad pachulol y 7-epi- α -selineno sintasa.

Una "sesquiterpeno sintasa" o un "polipéptido que tiene una actividad de sesquiterpeno sintasa" está previsto aquí 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.

- 30 Como "pachulol y 7-epi- α -selineno sintasa" o como "polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa", los presentes inventores indican aquí un polipéptido capaz de catalizar la síntesis de tanto el pachulol como el 7-epi- α -selineno a partir de FPP. El pachulol y el 7-epi- α -selineno pueden ser los únicos productos o pueden ser parte de una mezcla con otros sesquiterpenos.

- 35 La capacidad de un polipéptido para catalizar la síntesis de sesquiterpenos particulares (por ejemplo, pachulol y 7-epi- α -selineno) puede ser simplemente confirmada realizando el ensayo enzimático que se detalla en el Ejemplo 2.

Según la presente invención, también se indica que los polipéptidos incluyen polipéptidos truncados, a condición de que mantengan su actividad de sesquiterpeno sintasa como se define en cualquiera de las realizaciones de la presente invención y que compartan al menos el porcentaje de identidad definido con el fragmento correspondiente de SEQ ID NO: 1.

- 40 Como se pretende en el presente documento más adelante, "una secuencia de nucleótidos obtenida modificando SEQ ID NO: 2 o el complemento de la misma" engloba cualquier secuencia que ha sido obtenida cambiando la secuencia de SEQ ID NO: 2 o del complemento de la misma usando cualquier procedimiento conocido en la técnica, por ejemplo introduciendo cualquier tipo de mutaciones tales como mutaciones de delección, inserción o sustitución. Ejemplos de tales procedimientos se citan en la parte de la descripción con respecto a los polipéptidos de variante y los procedimientos para prepararlos.

- 45 El porcentaje de identidad entre dos secuencias peptídicas o nucleotídicas es una 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. Restos idénticas 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
- 50 alineamiento óptimo tomando el número de restos idénticos entre dos secuencias dividiéndolo entre 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 más alto posible. Pueden introducirse huecos en una o ambas secuencias en una o más posiciones del alineamiento para obtener el alineamiento óptimo. Estos huecos se tienen entonces en cuenta como restos no idénticos para el cálculo del porcentaje de identidad de secuencia.

- 55 El alineamiento con el fin de determinar el porcentaje de identidad de secuencia de aminoácidos o de ácidos nucleicos puede lograrse de diversas formas usando programas informáticos y, por ejemplo, programas informáticos

públicamente disponibles disponible de la web mundial. Preferentemente, puede usarse el programa BLAST (Tatiana y col, FEMS Microbiol Lett., 1999, 174:247-250, 1999) establecido a los parámetros por defecto, disponible del Centro Nacional para Información Biotecnológica (NCBI) en <http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi>, para obtener un alineamiento óptimo de secuencias peptídicas o nucleotídicas y para calcular el porcentaje de identidad de secuencia.

Se describe en el presente documento un procedimiento de producción de pachulol y 7-epi- α -selineno que comprende

- a) poner en contacto FPP con al menos un polipéptido que tiene una actividad de pachulol y 7-epi-a-selineno sintasa y que comprende una secuencia de aminoácidos al menos el 70 % idéntica a SEQ ID NO: 1;
- b) opcionalmente, aislar el pachulol y el 7-epi- α -selineno producidos en la etapa a).

El procedimiento es preferentemente un procedimiento para producir pachulol y 7-epi- α -selineno como productos principales. Incluso más preferentemente, el pachulol y el 7-epi-a-selineno representan en total al menos el 50 %, preferentemente al menos el 60 %, preferentemente al menos el 70 %, preferentemente al menos el 80 %, preferentemente al menos el 90 % del producto producido por el procedimiento descrito en el presente documento.

En otro aspecto, el 7-epi- α -selineno es el producto principal y el pachulol es el segundo producto principal. Por ejemplo, el 7-epi- α -selineno puede representar al menos el 20 %, preferentemente al menos el 30 %, preferentemente al menos el 34 % del producto obtenido por el procedimiento descrito en el presente documento. Como otro ejemplo, el pachulol puede representar al menos el 10 %, preferentemente al menos el 15 %, preferentemente al menos el 16 % del producto obtenido por el procedimiento descrito en el presente documento.

El producto del procedimiento descrito en el presente documento comprende además al menos un compuesto seleccionado del grupo que consiste en germacreno A, α -guaiano, α -humuleno, α -selineno, seicheleno, α -pachuleno, α -bulneseno y (E)-nerolidol, definiéndose estos compuestos por la forma de su estructura como se representa en la Figura 1. Más preferentemente, el producto del procedimiento descrito en el presente documento comprende todos estos compuestos.

El procedimiento puede llevarse a cabo *in vitro*, además de *in vivo*, como se explicará en detalle más adelante.

El polipéptido que va a ponerse en contacto con FPP *in vitro* puede obtenerse por extracción de cualquier organismo que lo expresa, usando tecnologías de extracción de proteínas o enzimas convencionales. Si el organismo huésped es un organismo unicelular o célula que libera el polipéptido de la invención en el medio de cultivo, el polipéptido puede simplemente ser recogido del medio de cultivo, por ejemplo por 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 adicionalmente extracción del polipéptido del lisado celular.

El polipéptido que tiene una actividad de pachulol y 7-epi-a-selineno sintasa, ya sea en una forma aislada o junto con otras proteínas, por ejemplo en un extracto de proteína en bruto obtenido de células cultivadas o microorganismos, puede entonces suspenderse en una solución de tampón a pH óptimo. Si es adecuado, pueden añadirse sales, BSA y otros tipos de co-factores enzimáticos con el fin de optimizar la actividad enzimática. Condiciones apropiadas se describen en más detalle en los ejemplos más adelante.

El precursor FPP puede entonces añadirse a la suspensión o solución, que entonces 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, pueden aislarse el pachulol y 7-epi- α -selineno, y opcionalmente los otros subproductos de sesquiterpeno producidos, de la solución incubada por procedimientos de aislamiento convencionales, tales como extracción con disolvente y destilación, opcionalmente después de eliminar los polipéptidos de la solución.

Según otro aspecto preferido, el procedimiento descrito en el presente documento se lleva a cabo *in vivo*. En este caso, la etapa a) comprende cultivar un organismo huésped no humano o célula capaz de producir FPP y transformarlo para expresar al menos un polipéptido que comprende una secuencia de aminoácidos al menos el 70 % idéntica a SEQ ID NO: 1 y que tiene una actividad de pachulol y 7-epi-a-selineno sintasa, en condiciones propicias para la producción de pachulol y 7-epi- α -selineno.

Según un aspecto más preferido, el procedimiento comprende además, antes de la etapa a), transformar un organismo no humano o célula capaz de producir FPP con al menos un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos al menos el 70 % idéntica a SEQ ID NO: 1 y que tiene una actividad de pachulol y 7-epi-a-selineno sintasa, de manera que dicho organismo exprese dicho polipéptido.

Estos aspectos del procedimiento descrito en el presente documento son particularmente ventajosos, ya 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 transformado para expresar dicho polipéptido.

Según una realización particular de la invención, el al menos un ácido nucleico que codifica el pachulol y la 7-epi- α -selineno sintasa comprende una secuencia de nucleótidos al menos el 70 %, preferentemente al menos el 75 %, de la secuencia de SEQ ID NO: 1.

preferentemente al menos el 80 %, preferentemente al menos el 85 %, preferentemente al menos el 90 %, más preferentemente al menos el 95 % e incluso más preferentemente al menos el 98 % idéntica a SEQ ID NO: 2 o el complemento de la misma. Según una realización incluso más preferida, dicho ácido nucleico comprende la secuencia de nucleótidos SEQ ID NO: 2 o el complemento de la misma. En una realización más preferida, dicho ácido nucleico consiste en SEQ ID NO: 2 o el complemento de la misma.

Según una realización preferida, el al menos un ácido nucleico usado en cualquiera de los procedimientos anteriores comprende una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma. Según una realización incluso más preferida, dicho al menos un ácido nucleico consiste en una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma.

Según otra realización, el al menos un ácido nucleico se aísla de *Valeriana jatamansi*.

Se indica que el organismo o célula "expresa" un polipéptido, a condición de que el organismo o célula se transforme para alojar un ácido nucleico que codifica dicho polipéptido, este ácido nucleico se transcribe en ARNm y el polipéptido se encuentra en el organismo huésped o célula. El término "expresar" engloba "expresar heterológamente" y "expresar en exceso", el último con referencia a niveles de ARNm, polipéptido y/o actividad enzimática, además de lo que se mide en un organismo o célula no transformado. Una descripción más detallada de procedimientos adecuados para transformar un organismo huésped no humano o célula se describirá después en el parte de la memoria descriptiva que se dedica a tales organismos huéspedes no humanos transformados o células como objetos específicos de la presente invención y en los ejemplos.

Se indica que un organismo o célula particular es "capaz de producir FPP" cuando produce FPP naturalmente o cuando no producen FPP naturalmente, pero se transforma para producir FPP, ya sea antes de la transformación con un ácido nucleico como se describe en el presente documento o junto con dicho ácido nucleico. Organismos o células transformados para producir una cantidad más alta de FPP que el organismo que existe de forma natural o célula también están englobados por los "organismos o células capaces de producir FPP". Procedimientos para transformar organismos, por ejemplo microorganismos, de manera que produzcan FPP ya son conocidos en la técnica. Tales procedimientos pueden encontrarse, por ejemplo, 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).

Para llevar a cabo los procedimientos descritos en el presente documento *in vivo*, el organismo huésped o célula se cultiva en condiciones propicias para la producción de pachulol y 7-epi- α -selineno. Por consiguiente, si el huésped es una planta transgénica, se proporcionan condiciones de crecimiento óptimas, tales como condiciones óptimas de luz, agua y nutrientes, por ejemplo. Si el huésped es un organismo unicelular, condiciones propicias para la producción de pachulol y 7-epi- α -selineno pueden comprender la adición de cofactores adecuados al medio de cultivo del huésped. Además, puede seleccionarse un medio de cultivo para maximizar la síntesis de pachulol y 7-epi- α -selineno. Condiciones de cultivo óptimas se describen de un modo más detallado en los siguientes ejemplos.

Organismos huéspedes no humanos adecuados para llevar a cabo el procedimiento descrito en el presente documento *in vivo* pueden ser cualquier organismo multicelular o unicelular no humano. En un aspecto preferido, el organismo huésped no humano usado para llevar a cabo el procedimiento descrito en el presente documento *in vivo* es una planta, un procariota o un hongo. Puede usarse cualquier planta, procariota u hongo. Plantas particularmente útiles son aquellas que producen naturalmente altas cantidades de terpenos. Más preferentemente, la planta está seleccionada de la familia de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Malvaceae*, *Asteraceae* o *Lamiaceae*. Por ejemplo, la planta está seleccionada 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*.

Más preferentemente, el organismo huésped no humano usado para llevar a cabo el procedimiento descrito en el presente documento *in vivo* es un microorganismo. Puede usarse cualquier microorganismo, pero incluso más preferentemente, dicho microorganismo es una bacteria o levadura. Lo más preferentemente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

Algunos de estos organismos no producen FPP naturalmente. Para ser adecuados para llevar a cabo el procedimiento descrito en el presente documento, estos organismos tienen que transformarse para producir dicho precursor. Así, pueden transformarse ya sea antes de la modificación con el ácido nucleico descrito según cualquiera de las realizaciones anteriores o simultáneamente, como se ha explicado anteriormente.

También puede usarse células eucariotas superiores aisladas, en lugar de organismos completos, como huéspedes para llevar a cabo el procedimiento descrito en el presente documento *in vivo*. Células eucariotas adecuadas pueden ser cualquier célula no humana, pero son preferentemente células de planta o fúngicas.

Según un aspecto preferido, el al menos un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa

usado en cualquiera de los procedimientos anteriormente descritos o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas comprende una secuencia de aminoácidos al menos el 70 %, preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, preferentemente al menos el 90 %, más preferentemente al menos el 95 % e incluso más preferentemente al menos el 98 % idéntica a SEQ ID NO: 1. Según un aspecto más preferido, dicho polipéptido comprende la secuencia de aminoácidos SEQ ID NO: 1. Incluso más preferentemente, dicho polipéptido consiste en SEQ ID NO: 1.

Según otro aspecto preferido, el al menos un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa usado en cualquiera de los procedimientos anteriormente descritos o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas comprende una secuencia de aminoácidos que es una variante de SEQ ID NO: 1 obtenida por ingeniería genética. En otros términos, dicho polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma. Según un aspecto más preferido, el al menos un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa usado en cualquiera de los procedimientos anteriormente descritos 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 SEQ ID NO: 1 obtenida por ingeniería genética, es decir, una secuencia de aminoácidos codificada por una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma.

Como se usa en el presente documento, el polipéptido está previsto como un polipéptido o fragmento de péptido que engloba las secuencias de aminoácidos identificadas en el presente documento, además de polipéptidos truncados o de variante, a condición de que mantengan su actividad como se ha definido anteriormente y que compartan al menos el porcentaje de identidad definido con el fragmento correspondiente de SEQ ID NO: 1.

Ejemplos de polipéptidos de variante son proteínas que existen de forma natural que resultan de eventos de corte y empalme de ARNm alternativos o de escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a la proteólisis incluyen, por ejemplo, diferencias en los extremos N o C tras la expresión en diferentes tipos de células hospedadoras, debido a eliminación 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 artificial de un ácido nucleico de la invención, como se describe más adelante, también están englobados por la invención.

También pueden usarse variantes de polipéptido resultantes de una fusión de secuencias de péptidos adicionales en los extremos amino y carboxilo en los procedimientos descritos en el presente documento. En particular, una fusión tal puede potenciar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un entorno o sistema de expresión deseado. Tales secuencias de péptidos adicionales pueden ser péptidos señal, por ejemplo. Por consiguiente, la presente divulgación engloba procedimientos usando polipéptidos de variante, tales como aquellos obtenidos por fusión con otros oligo- o polipéptidos y/o aquellos que están unidos a péptidos señal. Los polipéptidos resultantes de una fusión con otra proteína funcional, tal como otra proteína de la vía de biosíntesis de los terpenos, también pueden ser usados ventajosamente en los procedimientos descritos en el presente documento.

Según otro aspecto, el al menos un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa usado en cualquiera de los procedimientos anteriormente descritos o codificado por el ácido nucleico usado en cualquiera de las realizaciones anteriormente descritas se aísla de *Valeriana jatamansi*.

Una herramienta importante para llevar a cabo el procedimiento descrito en el presente documento es el polipéptido en sí. Un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa y que comprende una secuencia de aminoácidos al menos el 50 % idéntica a SEQ ID NO: 1 es un objeto de la presente invención.

Según una realización preferida, el polipéptido es capaz de producir pachulol y 7-epi- α -selineno como los principales productos. Según una realización incluso más preferida, el pachulol y el 7-epi- α -selineno representan en total al menos el 50 %, preferentemente al menos el 60 %, preferentemente al menos el 70 %, preferentemente al menos el 80 %, preferentemente al menos el 90 % del producto producido por el polipéptido de la invención.

Según otra realización de la invención, el 7-epi- α -selineno es el producto principal y el pachulol es el segundo producto principal. Por ejemplo, el 7-epi- α -selineno puede representar al menos 20 %, preferentemente al menos el 30 %, preferentemente al menos el 34 % del producto obtenido por el procedimiento de la invención. Como otro ejemplo, el pachulol puede representar al menos el 10 %, preferentemente al menos el 15 %, preferentemente al menos el 16 % del producto obtenido por el procedimiento de la invención.

Según otra realización, el producto del polipéptido de la presente invención comprende además al menos un compuesto seleccionado del grupo que consiste en germacreno A, α -guaiano, α -humuleno, α -selineno, seicheleno, α -pachuleno, α -bulneseno y (E)-nerolidol, definiéndose estos compuestos por la forma de su estructura como se representa en la Figura 1. En una realización más preferida, el producto de la presente invención comprende todos estos compuestos.

Según una realización preferida, el polipéptido comprende una secuencia de aminoácidos al menos el 70 %, preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, preferentemente al menos el 90 %, más preferentemente al menos el 95 % e incluso más preferentemente al menos el 98 % idéntica a SEQ ID NO: 1.

preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, preferentemente al menos el 90 %, más preferentemente al menos el 95 % e incluso más preferentemente al menos el 98 % idéntica a SEQ ID NO: 1. Según una realización más preferida, el polipéptido comprende la secuencia de aminoácidos SEQ ID NO: 1. Según una realización incluso más preferida, el polipéptido consiste en SEQ ID NO: 1.

- 5 Según otra realización preferida, el polipéptido comprende una secuencia de aminoácidos que es una variante de SEQ ID NO: 1 obtenida por ingeniería genética. En otros términos, dicho polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma. Según una realización más preferida, el polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa consiste en una secuencia de aminoácidos que es una variante de SEQ ID NO: 1 obtenida por ingeniería genética, es decir, una secuencia de aminoácidos codificada por una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma.

Según otra realización, el polipéptido se aísla de *Valeriana jatamansi*.

- 15 Como se usa en el presente documento, el polipéptido está previsto como un polipéptido o fragmento de péptido que engloba las secuencias de aminoácidos identificadas en el presente documento, además de polipéptidos truncados o de variante, a condición de que mantengan su actividad como se ha definido anteriormente y que compartan al menos el porcentaje de identidad definido con el fragmento correspondiente de SEQ ID NO: 1.

- 20 Ejemplos de polipéptidos de variante son proteínas que existen de forma natural que resultan de eventos de corte y empalme de ARNm alternativos o de escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a la proteólisis incluyen, por ejemplo, diferencias en los extremos N o C tras la expresión en diferentes tipos de células hospedadoras, debido a eliminación 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 artificial de un ácido nucleico de la invención, como se describen más adelante, también están englobados por la invención.

- 25 También están englobados por los polipéptidos de la invención variantes de polipéptido resultantes de una fusión de secuencias de péptidos adicionales en los extremos amino y carboxilo. En particular, una fusión tal puede potenciar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un entorno o sistema de expresión deseado. Tales secuencias de péptidos adicionales pueden ser un péptido señal, por ejemplo. Por consiguiente, la presente invención engloba variantes de los polipéptidos de la invención, tales como aquellas obtenidas por fusión con otros oligo- o polipéptidos y/o aquellas están unidas a péptidos señal.
- 30 Polipéptidos resultantes de una fusión con otra proteína funcional, tal como otra proteína de la vía de biosíntesis de los terpenos, también están englobados 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 hospedadores no humanos o células previstos para ser usados cuando el procedimiento se lleve a cabo *in vivo*.

- 35 Un ácido nucleico que codifica un polipéptido según cualquiera de las realizaciones descritas anteriormente es, por lo tanto, también un objeto de la presente invención.

- 40 Según una realización preferida, el ácido nucleico comprende una secuencia de nucleótidos al menos el 70 %, preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, preferentemente al menos el 90 %, más preferentemente al menos el 95 % e incluso más preferentemente al menos el 98 % idéntica a SEQ ID NO: 2 o el complemento de la misma. Según una realización más preferida, el ácido nucleico comprende la secuencia de nucleótidos SEQ ID NO: 2 o el complemento de la misma. Según una realización incluso más preferida, el ácido nucleico consiste en SEQ ID NO: 2 o el complemento de la misma.

Según otra realización, el ácido nucleico se aísla de *Valeriana jatamansi*.

- 45 El ácido nucleico de la invención puede definirse como que incluye polímeros de desoxirribonucleótido o ribonucleótido en ya sea forma mono- o bicatenaria (ADN y/o ARN). El término "secuencia de nucleótidos" también debe entenderse como que comprende una molécula de polinucleótido o una molécula de oligonucleótido en forma de un fragmento separado o como un componente de un ácido nucleico mayor. Los ácidos nucleicos de la invención también engloban ciertas secuencias de nucleótidos aisladas que incluyen aquellas que están sustancialmente libres de material endógeno contaminante. El ácido nucleico de la invención puede ser truncado, a condición de que codifique un polipéptido englobado por la presente invención, como se ha descrito anteriormente.

Según una realización más preferida, el al menos un ácido nucleico según cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma. Preferentemente, dicho ácido nucleico consiste en una secuencia de nucleótidos que ha sido obtenida modificando SEQ ID NO: 2 o el complemento de la misma.

- 55 Los ácidos nucleicos que comprenden una secuencia obtenida por mutación de SEQ ID NO: 2 o el complemento de la misma están englobados por la invención, a condición de que las secuencias comprendan compartan al menos el

porcentaje de identidad definido con los fragmentos correspondientes de SEQ ID NO: 2 o el complemento de la misma y a condición de que codifiquen un polipéptido que tiene una actividad de pachulol y 7-epi-a-selineno sintasa, como se define en cualquiera de las realizaciones anteriores. Las mutaciones pueden ser cualquier tipo de mutaciones de estos ácidos nucleicos, tales como mutaciones puntuales, mutaciones por delección, mutaciones por inserción y/o mutaciones por desplazamiento del marco. Puede prepararse un ácido nucleico de variante con el fin de adaptar su secuencia de nucleótidos a un sistema de expresión específico. Por ejemplo, se sabe que los sistemas de expresión bacterianos expresan más eficientemente polipéptidos si los aminoácidos están codificados por un codón preferido. Debido a la degeneración del código genético, en el que más de un codón puede codificar el mismo aminoácido, múltiples secuencias de ADN pueden codificar el mismo polipéptido, estando todas estas secuencias de ADN englobadas por la invención.

Otra herramienta importante para transformar organismos hospedadores o células adecuados para llevar a cabo el procedimiento descrito en el presente documento *in vivo* es un vector de expresión que comprende un ácido nucleico según cualquier realización de la invención. Dicho vector también es un objeto de la presente invención.

Un "vector de expresión", como se usa en el presente documento, incluye cualquier vector recombinante lineal o circular que incluye, pero no se limitan a, vectores virales, bacteriófagos y plásmidos. El experto es capaz de seleccionar un vector adecuado según el sistema de expresión. En una realización, el vector de expresión incluye el ácido nucleico de la invención operativamente unido a al menos una secuencia reguladora, que controla la iniciación y/o terminación de la transcripción y/o traducción, tal como un promotor de la transcripción, operador o potenciador, o un sitio de unión al ribosoma de ARNm y, opcionalmente, que incluye al menos un marcador de selección. Las secuencias de nucleótidos están "operativamente unidas" cuando la secuencia reguladora se refiere funcionalmente al ácido nucleico de la invención.

Pueden usarse vectores de expresión de la presente invención en los procedimientos de preparación de un huésped organismo y/o célula genéticamente transformado, en organismos hospedadores y/o células que alojan los ácidos nucleicos de la invención y en los procedimientos para producir o preparar los polipéptidos que tienen una actividad de pachulol y 7-epi-a-selineno sintasa, como se desvela adicionalmente más adelante.

Organismos hospedadores no humano recombinantes y células transformados para alojar al menos un ácido nucleico de la invención de manera que exprese heterológamente o exprese por exceso al menos un polipéptido de la invención también son herramientas muy útiles para llevar a cabo el procedimiento descrito en el presente documento. Dichos organismos y células huéspedes no humanos son otro objeto de la presente invención.

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

Organismos hospedadores 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 procariota o un hongo. Cualquier planta, procariota u hongo es adecuado para ser transformado según la presente invención. Plantas particularmente útiles son aquellas que producen naturalmente altas cantidades de terpenos. En una realización más preferida, la planta está seleccionada de la familia de *Solanaceae*, *Poaceae*, *Brassicaceae*, *Fabaceae*, *Malvaceae*, *Asteraceae* o *Lamiaceae*. Por ejemplo, la planta está seleccionada 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*.

En una realización más preferida, el organismo huésped no humano es un microorganismo. Cualquier microorganismo es adecuado para la presente invención, pero según una realización incluso más preferida, dicho microorganismo es una bacteria o levadura. Lo más preferentemente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

También puede transformarse células eucariotas superiores aisladas, en lugar de organismos completos. Como células eucariotas superiores, los presentes inventores indican aquí cualquier célula eucariota no humana, excepto células de levadura. Células eucariotas superiores preferidas son células vegetales o células fúngicas.

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 hospedadores que expresan heterológamente los polipéptidos codificados por el ácido nucleico con el que se transforman, además de expresar en exceso dichos polipéptidos. Por consiguiente, en una realización, la presente invención proporciona un organismo transformado, en el que los polipéptidos se expresan en cantidad más alta que en el mismo organismo no transformado así.

Se conocen varios procedimientos en la técnica para la creación de organismos hospedadores transgénicos o células tales como plantas, hongos, procariotas, o cultivos de células eucariotas superiores. Vectores de clonación y expresión apropiados para su uso como hospedadores de células bacterianas, fúngicas, de levadura, planta y de mamífero se describen, por ejemplo, en Pouwels y col., Cloning Vectors: A Laboratory Manual, 1985, Elsevier, New York and Sambrook y col., Molecular Cloning: A Laboratory Manual, 2ª edición, 1989, Cold Spring Harbor Laboratory

Press. Vectores de clonación y de expresión para plantas superiores y/o células vegetales en particular están disponibles para el experto. Véase, por ejemplo, Schardl y col. *Gene* 61: 1-11, 1987.

Los procedimientos para transformar organismos hospedadores o células para alojar ácidos nucleicos transgénicos son familiares para el experto. Para la creación de plantas transgénicas, por ejemplo, procedimientos actuales incluyen: electroporación de protoplastos de planta, transformación mediada por liposomas, transformación mediada por *Agrobacterium*, transformación mediada por polietilenglicol, bombardeo de partículas, microinyección de células vegetales y transformación usando virus.

En una realización, se integra ADN transformado en un cromosoma de un organismo huésped no humano y/o célula de forma que resulte un sistema recombinante estable. Puede usarse cualquier procedimiento de integración cromosómica conocida en la técnica en la práctica de la invención, que incluye, pero no se limita a, intercambio de casete mediado por recombinasa (RMCE), inserción cromosómica específica de sitio viral, inyección de adenovirus y pronuclear.

Con el fin de llevar a cabo el procedimiento para producir pachulol y 7-epi-a-selineno *in vitro*, como se expone en el presente documento anteriormente, es muy ventajoso proporcionar un procedimiento de preparación de al menos un polipéptido que tiene una actividad de pachulol y 7-epi-a-selineno sintasa como se describe en cualquier realización de la invención. Por lo tanto, la invención proporciona un procedimiento para producir al menos un polipéptido según cualquier realización de la invención que comprende

- a) cultivar un organismo huésped no humano o célula transformado con el vector de expresión de la invención, de manera que aloje un ácido nucleico según la invención y exprese o exprese en exceso un polipéptido de la invención;
- b) aislar el polipéptido del organismo huésped no humano o célula cultivado en la etapa a).

Según una realización preferida, dicho procedimiento comprende además, antes de la etapa a), transformar un organismo huésped no humano o célula con el vector de expresión de la invención de manera que aloje un ácido nucleico según la invención y exprese o expresa en exceso el polipéptido de la invención.

Puede usarse un ácido nucleico según cualquiera de las realizaciones anteriormente descritas.

La transformación y el cultivo del organismo huésped no humano o célula puede llevarse a cabo como se ha descrito anteriormente para el procedimiento de producción de pachulol y 7-epi- α -selineno *in vivo*. La etapa b) puede realizarse usando cualquier técnica muy conocida en la técnica para aislar un polipéptido particular de un organismo o célula.

Una "variante de polipéptido", como se hace referencia en el presente documento, significa un polipéptido que tiene una actividad de pachulol y 7-epi-a-selineno sintasa y que es sustancialmente homólogo al polipéptido según cualquiera de las realizaciones anteriores, pero que tiene una secuencia de aminoácidos diferente de la codificada por cualquiera de las secuencias de ácidos nucleicos de la invención debido a una o más delecciones, inserciones o sustituciones.

Las variantes pueden comprender secuencias conservativamente sustituidas, que significa que un resto de aminoácido dado está sustituido por un resto que tiene características fisicoquímicas similares. Ejemplos de sustituciones conservativas incluyen sustitución de un resto alifático por otro, tales como Ile, Val, Leu, o Ala por otro, o sustituciones de un resto polar por otro, tales como entre Lys y Arg; Glu y Asp; o Gln y Asn. Véase Zubay, *Biochemistry*, 1983, Addison-Wesley Pub. Co. Los efectos de tales sustituciones pueden calcularse usando matrices de puntuación de sustituciones tales como PAM-120, PAM-200 y PAM-250, como se trata en Altschul, *J. Mol. Biol.*, 1991, 219, 555-565.

Son muy conocidas otras sustituciones conservativas tales, por ejemplo sustituciones de regiones enteras que tienen características de hidrofobia similares.

También están englobadas por la invención variantes de péptido que existen de forma natural. Ejemplos de tales variantes son proteínas que resultan de eventos de corte y empalme de ARNm alternos o de la escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a la proteólisis incluyen, por ejemplo, diferencias en los extremos N o C tras la expresión en diferentes tipos de células hospedadoras, debido a eliminación proteolítica de uno o más aminoácidos terminales de los polipéptidos codificados por las secuencias de la invención.

Pueden usarse variantes de los polipéptidos de la invención para obtener, por ejemplo, actividad enzimática potenciada o reducida deseada, regioquímica o estereoquímica modificadas, o utilización de sustrato o distribución de productos alteradas, elevada afinidad por el sustrato, mejorada especificidad por la producción de uno o más de los compuestos deseados, elevada velocidad de la reacción enzimática, actividad o estabilidad más alta en un entorno específico (pH, temperatura, disolvente, etc.), o nivel de expresión mejorado en un sistema de expresión deseado. Puede prepararse una variante o mutante dirigido a sitio por cualquier procedimiento conocido en la técnica. Pueden obtenerse variantes y derivados de polipéptidos nativos aislando variantes que existen de forma natural, o la secuencia de nucleótidos de variantes, de otras líneas o especie de planta o de las mismas, o programando artificialmente mutaciones de secuencias de nucleótidos que codifican los polipéptidos de la invención. Pueden llevarse a cabo alteraciones de la secuencia de aminoácidos nativa por cualquiera de varios procedimientos convencionales.

Pueden usarse variantes de polipéptido resultantes de una fusión de secuencias de péptidos adicionales en los extremos amino y carboxilo 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 entorno o sistema de expresión deseado. Tales secuencias de péptidos adicionales pueden ser péptidos señal, por ejemplo. Por consiguiente, la presente invención engloba variantes de los polipéptidos de la invención, tales como aquellos obtenidos por fusión con otros oligo- o polipéptidos y/o aquellos que están unidos a péptidos señal. Polipéptidos de fusión englobados por la invención también comprenden polipéptidos de fusión resultantes de una fusión de otras proteínas funcionales, tales como otras proteínas de la vía de biosíntesis de los terpenos.

Por tanto, en una realización, la presente invención proporciona un procedimiento de preparación de un polipéptido de variante que tiene una actividad de pachulol y 7-epi- α -selineno sintasa, como se describe en cualquiera de las realizaciones anteriores, y que comprende las etapas de:

- (a) seleccionar un ácido nucleico según cualquiera de las realizaciones expuestas anteriormente;
- (b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
- (c) transformar células hospedadoras u organismos unicelulares con la secuencia de ácidos nucleicos mutante para expresar un polipéptido codificado por la secuencia de ácidos nucleicos mutante;
- (d) cribar el polipéptido según al menos una propiedad modificada; y,
- (e) opcionalmente, si el polipéptido no tiene actividad de pachulol y 7-epi- α -selineno sintasa de variante deseada, repetir las etapas de proceso (a) a (d) hasta que se obtenga un polipéptido con una actividad de pachulol y 7-epi- α -selineno sintasa de variante deseada;
- (f) opcionalmente, si se identificó un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa de variante deseada en la etapa d), aislar el ácido nucleico mutante correspondiente obtenido en la etapa (c).

Según una realización preferida, el polipéptido de variante preparado es capaz de producir pachulol y 7-epi- α -selineno como los productos principales. Según una realización incluso más preferida, el pachulol y el 7-epi- α -selineno representan en total al menos el 50 %, preferentemente al menos el 60 %, preferentemente al menos el 70 %, preferentemente al menos el 80 %, preferentemente al menos el 90 % del producto producido por el polipéptido de variante preparado.

Según otra realización de la invención, el 7-epi- α -selineno es el producto principal y el pachulol es el segundo producto principal. Por ejemplo, el 7-epi- α -selineno puede representar al menos el 20 %, preferentemente al menos el 30 %, preferentemente al menos el 34 % del producto obtenido con el polipéptido de variante preparado. Como otro ejemplo, el pachulol puede representar al menos el 10 %, preferentemente al menos el 15 %, preferentemente al menos el 16 % del producto obtenido con el polipéptido de variante preparado.

Según otra realización, el producto del polipéptido de variante preparado comprende además al menos un compuesto seleccionado del grupo que consiste en germacreno A, α -guaiano, α -humuleno, α -selineno, seicheleno, α -pachuleno, α -bulneseno y (E)-nerolidol, definiéndose estos compuestos por la forma de su estructura como se representa en la Figura 1. En una realización más preferida, el producto del polipéptido de variante preparado comprende todos estos compuestos.

En la etapa (b), puede crearse un gran número de secuencias de ácidos nucleicos mutantes, por ejemplo por mutagénesis al azar, mutagénesis específica de sitio o barajado de ADN. Los procedimientos detallados de barajado de genes 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, barajado de ADN se refiere a un proceso de recombinación al azar 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, flanqueada por sitios de restricción que permiten la unión a fragmentos de la secuencia nativa. Tras la unión, la secuencia reconstruida resultante codifica un análogo que tiene la inserción, sustitución o delección de aminoácido deseada. Alternativamente, pueden emplearse procedimientos de mutagénesis específica de sitio dirigida por oligonucleótidos para proporcionar un gen alterado en el que codones predeterminados pueden ser alterados por sustitución, delección o inserción.

Por consiguiente, el polipéptido que comprende SEQ ID NO: 1 puede ser recombinado con cualquier otro ácido nucleico de sesquiterpeno sintasa que codifica, por ejemplo aislado de un organismo distinto de *Valeriana jatamansi*. Así, pueden obtenerse ácidos nucleicos mutantes y separarse, que pueden usarse para transformar una célula hospedadora según procedimientos convencionales, por ejemplo tal como se desvela en los presentes ejemplos.

En la etapa (d), el polipéptido obtenido en la etapa (c) se criba según al menos una propiedad modificada, por ejemplo una actividad enzimática modificada deseada. Ejemplos de actividades enzimáticas deseadas, según las que puede cribarse un polipéptido expresado, incluyen actividad enzimática potenciada o reducida, por ejemplo como se mide por el valor K_M o $V_{m\acute{a}x}$, regioquímica o estereoquímica modificadas y utilización de sustrato o distribución de productos alteradas. El cribado de actividad enzimática puede realizarse según procedimientos familiares para el experto y los desvelados en los presentes ejemplos.

La etapa (e) proporciona la repetición de las etapas de proceso (a)-(d), que puede realizarse preferentemente en

paralelo. Por consiguiente, creando un número significativo de ácidos nucleicos mutantes, pueden transformarse muchas células hospedadoras con diferentes ácidos nucleicos mutantes al mismo tiempo, permitiendo el posterior cribado de un número elevado de polipéptidos. Así pueden aumentarse las probabilidades de obtener un polipéptido de variante deseado a criterio del experto.

- 5 Todas las publicaciones mencionadas en la presente solicitud desvelan y describen los procedimientos y/o materiales a propósito de los que se citan las publicaciones.

Descripción de los dibujos

Figura 1: Cromatogramas iónicos totales del análisis de CG-EM de los productos de sesquiterpeno generados por el pachulol y la 7-epi- α -selineno sintasa de la presente invención (SEQ ID NO: 1) (A) y perfil obtenido con un control negativo (B). Los picos marcados con números se identificaron como los sesquiterpenos: β -elemeno (producto de degradación de (+)-germacreno A) (1); α -guaiano (2); seicheleno (3); α -humuleno (4); α -pachuleno (5); α -selineno (6); α -bulneseno (7); 7-epi- α -selineno (8); (E)-nerolidol (9); alcohol de sesquiterpeno no identificado (10); pachulol (11).

Figura 2: Espectros de masas de los dos productos principales de pachulol y 7-epi- α -selineno sintasa de la invención (SEQ ID NO: 2) (pico 8 y pico 11 en la Figura 1 (A y B, respectivamente)) y comparación con los espectros de masas del 7-epi- α -selineno y pachulol auténticos (C y D, respectivamente).

Figura 3: Estructura de los sesquiterpenos producidos por el pachulol y la 7-epi- α -selineno sintasa de la invención (SEQ ID NO: 2).

Realizaciones específicas de la invención o ejemplos

- 20 La invención se describirá ahora en más detalle a modo de los siguientes ejemplos.

Ejemplo 1

Aislamiento de un ADNc de sesquiterpeno sintasa de raíces de *Valeriana jatamansi*

Se obtuvieron plantas de *Valeriana jatamansi* (sinónimo: *Valeriana walichii*) de B & T World Seeds (Paguignan, Aigues-Vives, Francia). Las plantas se cultivaron en un invernadero a temperatura mínima de 14 °C. Se recogieron por separado rizomas, raíces jóvenes y maduras y se congelaron inmediatamente en nitrógeno líquido. El material se machacó y se pulverizó a un polvo fino en nitrógeno líquido usando un mortero y pistilo. Se extrajo ARN total usando Concert™ Plant RNA Reagent de Invitrogen siguiendo las instrucciones del fabricante, excepto que la precipitación con isopropanol se sustituyó por una precipitación con LiCl 2 M. Se evaluó la calidad del ARN en un gel de agarosa verificando la integridad de las bandas de ARN ribosómico. El ARNm se purificó del ARN total por cromatografía de afinidad en oligodT-celulosa usando el kit de aislamiento de ARNm FastTrack® 2.0 (Invitrogen) siguiendo las instrucciones del fabricante.

Se diseñaron oligonucleótidos degenerados para motivos conservados observados en alineamientos de las secuencias de aminoácidos de sesquiterpeno sintasas de planta (Deguerry y col, 2006, Arch Biochem Biophys. 454(2), 123-36). La RT-PCR con estos oligonucleótidos específicos de sesquiterpeno sintasas se realizó usando el kit Qiagen OneStep RT-PCR y un ciclador de gradiente térmico Mastercycler de Eppendorf. Mezclas de reacción típicas contienen 10 μ l de 5X tampón Qiagen OneStep RT-PCR, 200 μ M de cada dNTP, 0,4 μ M de cada cebador, 2 μ l de Qiagen OneStep RT-PCR Enzyme Mix, 1 μ l de inhibidor de ribonucleasas RNasin® (Promega Co.) y 1 μ g de ARN total en un volumen final de 50 μ l. Las condiciones del ciclador térmico fueron: 30 min a 50 °C (transcripción inversa); 15 min a 95 °C (activación por ADN polimerasa); 35 ciclos de 30 s a 94 °C, 30 s a 41 a 51 °C y 1 min a 72 °C. Los tamaños de los productos de PCR se evaluaron en un gel de agarosa al 1,2 %. Las bandas correspondientes al tamaño esperado se eliminaron del gen, se purificaron usando el kit de extracción en gel QIAquick® (Qiagen) y se clonaron en el vector pCR®2.1-TOPO usando el kit de clonación TOPO TA (Invitrogen). Entonces se sometieron ADNc insertados a una secuenciación de ADN y la secuencia se comparó contra la base de datos de proteínas no redundantes de GenBank (NCBI) usando el algoritmo de BLASTX (Altschul y col. 1990, J. Mol. Biol. 215, 403-410). La combinación del cebador directo TpsCF2 (SEQ ID NO: 3) (5'-GGGA(A/T)(A/T)G(A/T)(A/T/G/C)(A/T)(C/T/G)GTTGAA(T/G)(T/G)TTATTTTGG-3') y el cebador inverso TpsCR3 (SEQ ID NO: 4) (5'-GT(A/T)(C/G)CGTG(A/T/G/C)G(A/C/T)GTCGTA(A/C/T)G(T/G)GTCATC-3') proporcionaron un fragmento de 81 pb que presentaba homología de secuencias con sesquiterpeno sintasas de planta conocidas.

Entonces se usó una combinación de amplificación rápida de los extremos de ADNc 3' y 5' (RACE) para obtener la secuencia de longitud completa del ADNc correspondiente a este fragmento. Para la 3'RACE, se dedujeron dos oligonucleótidos codificantes de la secuencia de 81 pb obtenida por RT-PCR: 20-3R1 (SEQ ID NO: 5) y 20-3R2 (SEQ ID NO: 6). Se preparó un ADNc bicatenario unido a adaptador usando el kit de amplificación de ADNc Marathon™ (Clontech) siguiendo el protocolo del fabricante. Este ADNc se preparó a partir de ARNm purificado de ARN total de rizoma de *V. jatamansi*. Mezclas de reacción de RACE típicas contienen, en un volumen final de 50 μ l, 5 μ l de 10X de tampón de reacción de PCR (Clontech), 200 μ M de cada dNTP, 1 μ l de mezcla de polimerasas Advantage® 2, 200 nM de cada cebador y 5 μ l de ADNc 250 veces diluido. La amplificación se realizó en un ciclador de gradiente térmico Mastercycler de Eppendorf. Las condiciones del ciclo térmico fueron las siguientes: 1 min a 94 °C, 5 ciclos de 5 s a 94 °C y 3 min a 72 °C, 5 ciclos de 5 s a 94 °C y 3 min a 70 °C, 20 ciclos de 5 s a 94 °C y 3 min

a 68 °C. Los productos de amplificación se evaluaron, se subclonaron, y la secuencia se analizó como se ha descrito anteriormente. El extremo 3' del ADNc se obtuvo después de una primera ronda de amplificación con el cebador 20-3R1 y el cebador API (Clontech) y una segunda ronda de amplificación con el cebador 20-3R2 y el cebador AP2 (Clontech).

- 5 Se dedujeron dos cebadores inversos de la secuencia obtenida por 3'RACE: 20-5R1 (SEQ ID NO: 7) y 20-5R2 (SEQ ID NO: 8). Se usó el kit de amplificación de ADNc SMART™ RACE (Clontech) para preparar un 5'RACE-ready cDNA de ARN total de raíz de *V. jatamansi*. Se realizó la 5'RACE en 50 µl de 1X tampón de PCR Advantage 2 que contenía 200 µM de dNTP, 5 µl de Universal Primer Mix (Clontech), 2 µM del primer cebador específico de gen (20-5R1), 2,5 µl de 5'RACE-ready cDNA y 1 µl de mezcla de polimerasas Advantage® 2. Se realizó una segunda ronda de amplificación en la misma condición con 1 µl de la primera amplificación, el cebador específico de gen anidado (20-5R2) y el cebador universal anidado (Clontech). La condición de ciclado de temperaturas y el análisis de fragmentos de ADN fueron como se ha descrito anteriormente para la 3'RACE.

- 15 La combinación de 3' y 5'-RACE permitió la reconstitución de la secuencia de longitud completa de un ADNc que se llamó ValR20 (SEQ ID NO: 9). La secuencia de aminoácidos deducida mostró homología con las sesquiterpeno sintasas de planta y contuvo motivos de aminoácidos de terpeno sintasas típicas tales como el motivo DDxxD. Las secuencias más próximas fueron una secuencia de aminoácidos de vetispiradieno sintasa de *S. tuberosum* (49 % de identidad) y una supuesta sesquiterpeno sintasa de *V. vinífera* (hasta el 46 % de identidad).

Ejemplo 2

Expresión heteróloga y caracterización de ValR20 como 7-epi-α-selineno y pachulol sintasa

- 20 Se diseñaron dos oligonucleótidos, Val-R20-topo-inicio (SEQ ID NO: 10) y Val-R20-parada (SEQ ID NO: 11), a partir de las regiones de inicio y parada de ValR20 y se usaron para amplificar la secuencia de longitud completa de este ADNc. Se diseñó Val-R20-topo-inicio según los kits de expresión Champion™ pET Direccional TOPO® (Invitrogen). La amplificación se realizó con la ADN polimerasa Pfu (Promega) del conjunto de 5'-RACE-Ready cDNA preparado con el kit de amplificación de ADNc Smart RACE (Clontech). Las condiciones de ciclado térmico fueron las siguientes:
- 25 2 min a 95 °C; 32 ciclos de 5 s a 94 °C, 20 s a 53 °C y 3 min a 72 °C. Los productos de PCR se purificaron sobre un gel de agarosa y se eluyeron usando el kit de extracción en gel QIAquick® (Qiagen, Valencia, CA). El producto de PCR se ligó en el plásmido pET101 siguiendo el protocolo del fabricante (Invitrogen). Las construcciones se verificaron por secuenciación de ADN. Se seleccionó un clon, 504-ValR20 (SEQ ID NO: 2), para la posterior expresión heteróloga y experimentos de ensayo enzimático. En comparación con las secuencias reconstituidas de RACE, la secuencia de
- 30 504-ValR20 (SEQ ID NO: 2) mostró cuatro diferencias de nucleótidos que condujeron a tres cambios de aminoácidos. La secuencia de aminoácidos codificada por 504-ValR20 se proporciona en SEQ ID NO: 1.

- El plásmido se transfirió a células de *E. coli* BL21 Star™ (DE3) (Invitrogen). Se usaron colonias individuales de células transformadas para inocular 5 ml de medio LB. Después de alcanzar una DO de 0,3, se usaron cultivos de 5 ml para
- 35 inocular 25 ml de medio LB. Los cultivos se incubaron a 37 °C hasta que alcanzó una DO de 0,5 y entonces se transfirieron a una estufa de incubación a 20 °C. Después de 1 hora de equilibrado, la expresión de las proteínas se indujo mediante la adición de IPTG 1 mM y el cultivo se incubó durante la noche a 20 °C. Al día siguiente, las células se recogieron por centrifugación, se resuspendieron en 1 ml de tampón de extracción (MOPSO 50 mM, pH 7,4, DTT 1 mM, 10 % de glicerol) y se rompieron por sonicación. Los residuos de células se sedimentaron por centrifugación de 30 min a 18.000 g y se recuperó el sobrenadante que contenía las proteínas solubles.

- 40 Los ensayos enzimáticos se realizaron en tubos de vidrio sellados con Teflon usando 250 µl de extracto de proteína en un volumen final de 5 ml de tampón de extracción complementado con MgCl₂ 10 mM y pirofosfato de farnesilo 100 a 250 µM (preparado como se describe por Keller y Thompson, J. Chromatogr 645(1), 161-167, 1993). Los ensayos se solaparon con 3 ml de pentano y los tubos se incubaron durante la noche a 30 °C. Se recuperó la fase de pentano y el medio se extrajo con un segundo volumen de pentano. Las fracciones de pentano combinadas se concentraron
- 45 bajo nitrógeno y se analizaron por CG-EM en un sistema detector selectivo de masas de cuadrupolo de CG Hewlett-Packard 6890N, equipado con una columna capilar de 0,25 mm de diámetro interno por 30 m de longitud DB-1EM (J&W Scientific). La temperatura del horno se programó de 50 °C (1 min de mantenimiento) a 280 °C a 10 °C/min. El gas portador fue He a un flujo constante de 1 ml/min. La identidad de los productos se confirmó basándose en la concordancia de los índices de retención y espectros de masas de patrones auténticos cuando estuvieron disponibles
- 50 o se basó en datos publicados (Joulain y Koenig, 1998). Se realizaron controles negativos con proteínas obtenidas de *E. coli* transformada con el plásmido pET101 sin inserción.

- Se encontró que la enzima recombinante (SEQ ID NO: 1) era una sesquiterpeno sintasa de múltiples productos que convierte el pirofosfato de farnesilo en al menos 11 productos. El principal producto fue 7-epi-α-selineno, que representó el 34,2 % de la mezcla de sesquiterpenos total. El pachulol fue el segundo producto más abundante de la enzima (16,4 %). Otros sesquiterpenos producidos fueron germacreno A (12,2 %), α-guaieno (2,4 %), seicheleno (4,6 %), α-humuleno (5,3 %), α-pachuleno (2,4 %), α-selineno (9,6 %), α-bulneseno (6,3 %), (E)-nerolidol (2,5 %) y un alcohol de sesquiterpeno no identificado (4,1 %).

LISTADO DE SECUENCIAS

<110> Firmenich SA

<120> PROCEDIMIENTO DE PRODUCCIÓN DE SESQUITERPENOS

<130> M/58103-EP-DIV1

5 <160> 11

<170> PatentIn versión 3.5

<210> 1

<211> 553

<212> PRT

10 <213> *Valeriana jatamansi*

<400> 1

```

Met Leu Ser Thr Glu Ser Gln Val Phe Arg Pro Leu Ala Asn Phe Glu
1      5      10      15

Lys Ser Leu Trp Gly Asn Leu Phe Thr Ser Phe Ser Val Asp Tyr Leu
      20      25      30

Thr Lys Lys Thr Asn Thr Glu Glu His Glu Gly Leu Leu Glu Lys Val
      35      40      45

Arg Leu Met Phe Leu Asp Ala Ser Lys Leu Lys Ile Pro Glu Lys Ile
      50      55      60

Asn Phe Ile Asn Thr Leu Glu Arg Leu Gly Val Ser Tyr His Met Glu
65      70      75      80

Arg Glu Ile Glu Asp Gln Leu His Gln Met Phe Asp Ala His Ser Lys
      85      90      95

Phe Gln Asp Asp Ile Gln Arg Phe Asp Leu Phe Thr Leu Gly Ile Tyr
      100      105      110

Phe Arg Ile Leu Arg Gln His Gly Tyr Asn Ile Ser Ser Asp Val Phe
      115      120      125

Lys Lys Leu Lys Asp Ser Asn Gly Lys Phe Lys Glu Glu Leu Lys Asp
      130      135      140

Asp Val Ile Gly Ile Leu Ser Leu Tyr Glu Ala Thr His Val Arg Thr
      145      150      155      160

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

```

Leu Glu Ser Met Ser Thr Ala Ser Leu Ser Pro Phe Leu Gly Met Gln
 180 185 190

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

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

Asp Leu Leu Leu Arg Phe Ser Thr Ile Asp Phe Asn Leu Leu Gln Met
 225 230 235 240

Leu His Lys Gln Glu Leu Cys Asp Ala Ser Arg Trp Trp Asn Glu Met
 245 250 255

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

Tyr Leu Trp Thr Leu Ser Ala Tyr Tyr Glu Pro Lys Tyr Ser Leu Ala
 275 280 285

Arg Arg Ile Leu Ile Lys Leu Met Ile Leu Val Ser Leu Thr Asp Asp
 290 295 300

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

Ala Val Glu Arg Leu Asp Glu Gly Ser Ile Asn Gln Leu Pro Asp Tyr
 325 330 335

Met Lys Ile Leu Tyr Lys Ala Leu Leu Asp Phe Phe Glu Glu Thr Glu
 340 345 350

Asp Ile Leu Cys Lys His Gly Ile Ile Asn Gly Ser His Arg Val Asn
 355 360 365

Tyr Gly Lys Tyr Val Tyr Lys Glu Ile Val Asn Cys Tyr Asn Thr Glu
 370 375 380

Tyr Lys Trp Phe Asn Lys Arg Tyr Val Pro Asp Phe Glu Glu Tyr Met
 385 390 395 400

Gln Lys Ala Val Val Thr Ser Gly Asn Asn Leu Leu Ile Thr Trp Ser
 405 410 415

Phe Gln Gly Met Asp Gln Val Ala Ser Ile Lys Ala Phe Glu Trp Val
 420 425 430

Lys Ser His Pro Lys Met Val Val Ser Ser Asn Lys Val Leu Arg Leu
435 440 445

Val Asp Asp Val Met Ser His Glu Glu Glu Asp Glu Arg Gly His Val
450 455 460

Ala Thr Gly Leu Glu Cys Tyr Gln Lys Thr Tyr Gly Gly Asn Arg Lys
465 470 475 480

Glu Ile Ile Pro Glu Phe Tyr Lys Arg Ile Asp Asp Ala Trp Lys Asp
485 490 495

Val Asn Glu Glu Phe Leu Lys Pro Asp Lys Leu Pro Leu Glu Ile Leu
500 505 510

Met Arg Val Ile Asn Leu Thr Arg Ile Gly Asp Val Val Tyr Lys Tyr
515 520 525

Asp Asp Gly Tyr Thr His Pro Thr Lys Ala Leu Lys Asp His Ile Ile
530 535 540

Ser Leu Phe Val Asn Pro Ile Ile Ile
545 550

<210> 2

<211> 1662

<212> ADN

<213> *Valeriana jatamansi*

5

<400> 2

atgttaagca ctgagagtca agtttttctgt ccgttggcaa attttgagaa aagtttatgg	60
ggaaatcttt tcacctcatt ttctgtggat tatctgacta agaaaacaaa tacagaagaa	120
catgaaggat tattagaaaa agtgagactg atgttttttag atgcatccaa attgaagatt	180
ccagaaaaga tcaatttcat aaatacactt gaaaggtttag gtgtatcata tcatatggag	240
agagagattg aagatcagct tcatcagatg tttgatgctc attctaaatt tcaagatgat	300
attcaacggg ttgatttggt cactttggga atttacttca ggattctcag acaacatggg	360
tataatatct ctagcgatgt tttcaagaag ttgaaagata gcaatggaaa attcaaggaa	420
gaactaaaag atgatgtgat tggcattcta agcttgatg aagctacaca tgtaagaacc	480
cacggtgacg atattctaga tgaagctttc atctatacaa aagctcaact agaacttatg	540
tccaccgcaa gtttaagccc gtttcttggg atgcaagtta cgcattgctt gattcagtct	600
ctccacaaaag ggatcccaag aatcgagtcg cgcaactata tatctgttta tgaagaagat	660
ccaaacaaaa atgatctatt attgaggttc tcaacgattg atttcaatct gctgcaaatg	720
cttcacaagc aagaattgtg tgatgcctca aggtggtgga atgaaatgga gtttgagacg	780
aaactatctt atgcgagaga tcgagtggta gaaggctatt tatggaccct tagcgcatat	840
tacgaaccga aatactcttt ggctogaaga atattaatca aattaatgat attggtatct	900
cttacggatg atacgtatga tgcatatggg acgttagatg aacttcaact ctttacggat	960
gcagtagaaa ggttggtatga gggttccatc aatcagcttc ctgattacat gaagattctc	1020
tataaggctc tgctagattt ttctgaggaa acagaagata tattatgcaa acatggaata	1080
attaatggtt ctcatcgctt taattatggg aaatatgtgt ataaagagat tgtgaattgc	1140
tacaataccg agtacaaatg gttcaacaaa agatacgtgc cggattttga agaatatatg	1200
cagaaagcag tagtgacttc aggttaacaat ttgcttataa cgtgggtctt tcaaggaatg	1260
gatcaagtcg caagtatcaa agcgttcgag tgggttaaaa gtcattccga aatggtagtc	1320
tcgtcgaata aagtcctacg acttggtgac gacgtaatga gccacgagga agaagatgaa	1380
aggggacatg ttgcaacagg ccttgaatgc tatcagaaaa catatggtgg aaatagaaaa	1440
gagatcattc cagaatttta taagaggatt gatgatgctt ggaaagatgt aaatgaagaa	1500
tttttgaaac ccgataaatt accgctagaa atactaatgc gtgttattaa cctcacgaga	1560
attggcgacg ttgtttacaa gtatgacgac gggatatact atccaacgaa agcattaaag	1620
gatcacatca tatcgttggt cgtgaatccc ataactatat ga	1662

<210> 3
 <211> 29
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Cebador

	<220> <221> misc_feature <222> (9)..(9) <223> n es a, c, g o t	
5	<400> 3 gggawwgnw bgttgaakkt tattttgg	29
10	<210> 4 <211> 26 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
15	<220> <221> misc_feature <222> (9)..(9) <223> n es a, c, g o t	
20	<400> 4 gtwscgtgng hgtcgtahgk gtcac	26
	<210> 5 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
25	<400> 5 acccttagcg catattacga accga	25
30	<210> 6 <211> 28 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
35	<400> 6 cgaaccgaaa tactcttgg ctgaaga	28
	<210> 7 <211> 28 <212> ADN <213> Secuencia artificial	
40	<220> <223> Cebador	
	<400> 7 ccactcgaac gcttgatac ttgcgact	28
45	<210> 8 <211> 28 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
50	<400> 8 cctcatcaa ccttctact gcatccgt	28
	<210> 9	

ES 2 792 041 T3

<211> 1979

<212> ADN

<213> *Valeriana jatamansi*

<400> 9

acgcggggga ctatcatata attaagagat taattaagtt tgtgaaaatg ttaagcactg 60

agagtcaagt ttttcgtccg ttggcaaatt ttgagaaaag tttatgggga aatcttttca 120

5

cctcattttc	tgtggattat	ctgactaaga	aaacaaatac	agaagaacat	gaaggattat	180
tagaaaaagt	gagactgatg	tttttagatg	catccaaatt	gaagattcca	gaaaagatca	240
atttcataaa	tacacttgaa	aggttaggtg	tatcatatca	tatggagaga	gagattgaag	300
atcagcttca	tcagatgttt	gatgctcatt	ctaaatttca	agatgatatt	caacggtttg	360
atttgttcac	tttggaatt	tacttcagga	ttctcagaca	acatggttat	aatatctcta	420
gcgatgtttt	caagaagttg	aaagatagca	atggaaaatt	caaggaagaa	ctaaaagatg	480
atgtgattgg	cattctaagc	ttgtatgaag	ctacacatgt	aagaaccac	ggtgacgata	540
ttctagatga	agctttcatc	tatacaaaag	ctcaactaga	atctatgtcc	accgcaagtt	600
taagcccgtt	tcttggtatg	caagttacgc	atgctttgat	tcagtctctc	cacaaagga	660
tcccaagaat	cgagtcgcg	aactatatat	ctgtttatga	agaagatcca	aacaaaaatg	720
atctattatt	gaggttctca	aagattgatt	tcaatctgct	gcaaagtctt	cacaagcaag	780
aattgtgtga	tgcctcaagg	tggtggaatg	aaatggagtt	tgagacgaaa	ctatcttatg	840
cgagagatcg	agtggtagaa	ggctatttat	ggacccttag	cgcatattac	gaaccgaaat	900
actctttggc	tcgaagaata	ttaatcaaata	taatgatatt	ggtatctctt	acggatgata	960
cgtatgatgc	atatggtacg	ttagatgaac	ttcaactctt	tacggatgca	gtagaaaggt	1020
tggatgaggg	ttccatcaat	cagcttcctg	attacatgaa	gattctctat	aaggctctgc	1080
tagatTTTTT	cgaggaaaca	gaagatatat	tatgcaaaca	tggaataatt	aatggttctc	1140
atcgcgttaa	ttatggaaaa	tatgtgtata	aagagattgt	gaattgctac	aataccgagt	1200
acaaatgggt	caacaaaaga	tacgtgccgg	atTTTgaaga	atatatgcag	aaagcagtag	1260
tgacttcagg	taacaatttg	cttataacgt	ggtcttttca	aggaatggat	caagtcgcaa	1320
gtatcaaagc	gttcgagtgg	gttaaaaatc	atccgaaaat	ggtagtctcg	tcgaataaag	1380
tcctacgact	tgttgacgac	gtaatgagcc	acgaggaaga	agatgaaagg	ggacatgttg	1440
caacaggcct	tgaatgctat	cagaaaacat	atggtggaaa	tagaaaagag	atcattccag	1500
aattttataa	gaggattgat	gatgcttgga	aagatgtaaa	tgaagaattt	ttgaaaccgg	1560
ataaattacc	gctaaaaata	ctaatgcgtg	ttattaacct	cacgagaatt	ggcgacgttg	1620
tttacaagta	tgacgacggg	tatactcatc	caacgaaagc	attaaaggat	cacatcatat	1680
cgttgttcgt	gaatcccata	atcatatgat	tatcatcaat	cgttcgacgt	gtgtgcgcgt	1740
ggttttggtt	aataagtaaa	taagggtggac	tgcatcacca	tcccaaactc	gcatgtttgc	1800
ttctgggttt	ctttatttgt	gtttgcgtgt	ttacttgatt	tcttttccgt	tgtttgtaat	1860
aataccgagt	atttcttttc	cgttgtttgt	aataatacgg	agtatgcatt	actctctttt	1920
tcacatgcaa	ttaataagac	gatattatTT	gtaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1979

ES 2 792 041 T3

	<210> 10	
	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador	
	<400> 10	
	caccatgtta agcactgaga gtcaagtt	28
10	<210> 11	
	<211> 27	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador	
15	<400> 11	
	gataatcata tgattatggg attcacg	27

REIVINDICACIONES

1. Un polipéptido que tiene una actividad de pachulol y 7-epi- α -selineno sintasa y que comprende una secuencia de aminoácidos al menos el 70 % idéntica a SEQ ID NO: 1.
- 5 2. El polipéptido de la reivindicación 1, caracterizado porque comprende (1) una secuencia de aminoácidos al menos el 90 % idéntica a SEQ ID NO: 1 o (2) la secuencia de aminoácidos de SEQ ID NO: 1.
3. Un ácido nucleico que codifica un polipéptido según una cualquiera de las reivindicaciones 1 y 2.
4. El ácido nucleico de la reivindicación 3, caracterizado porque comprende
 - (1) una secuencia de nucleótidos al menos el 70 % idéntica a SEQ ID NO: 2 o el complemento de la misma;
 - (2) una secuencia de nucleótidos al menos el 90 % idéntica a SEQ ID NO: 2 o el complemento de la misma; o
 - 10 (3) la secuencia de nucleótidos de SEQ ID NO: 2 o el complemento de la misma.
5. Un vector de expresión que comprende el ácido nucleico de la reivindicación 3 o 4.
6. El vector de expresión de la reivindicación 5, en forma de un vector vírico, un bacteriófago o un plásmido.
7. El vector de expresión de la reivindicación 5 o 6, que incluye el ácido nucleico unido operativamente con al menos una secuencia reguladora que controla la transcripción, inicio o terminación de la traducción, tal como un promotor, operador o potenciador de la transcripción o un sitio de unión ribosómico de ARNm y, opcionalmente, que incluye al menos un marcador de selección.
- 15 8. Un organismo huésped no humano o célula transformado para alojar al menos un ácido nucleico según la reivindicación 3 o 4, de modo que exprese o sobreexpresa de manera heteróloga al menos un polipéptido según la reivindicación 1 o 2.
- 20 9. El organismo huésped no humano de la reivindicación 8, caracterizado porque es una planta, un procariota o un hongo.
10. El organismo huésped no humano de la reivindicación 8, caracterizado porque es un microorganismo, preferentemente una bacteria o levadura.
- 25 11. El organismo huésped no humano de la reivindicación 10, caracterizado porque dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.
12. La célula de la reivindicación 8, caracterizado porque es una célula vegetal o una célula fúngica.
13. Un procedimiento para producir al menos un polipéptido según cualquiera de las reivindicaciones 1 y 2 que comprende
 - a) cultivar un organismo huésped no humano o célula transformado con el vector de expresión de una cualquiera de las reivindicaciones 5 a 7, de modo que aloje un ácido nucleico según la reivindicación 3 o 4 y exprese o sobreexpresa un polipéptido según la reivindicación 1 o 2;
 - 30 b) aislar el polipéptido del organismo huésped no humano o célula cultivado en la etapa a).
14. El procedimiento de la reivindicación 13, caracterizado porque comprende además, antes de la etapa a), transformar un organismo huésped no humano o célula con el vector de expresión de una cualquiera de las reivindicaciones 5 a 7, de modo que aloje un ácido nucleico según la reivindicación 3 o 4 y exprese o sobreexpresa el polipéptido de la reivindicación 1 o 2.
- 35 15. Un procedimiento para preparar un polipéptido variante que tiene una actividad de pachulol y 7-epi- α -selineno sintasa que comprende las etapas de:
 - (a) seleccionar un ácido nucleico según la reivindicación 3 o 4;
 - 40 (b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
 - (c) transformar células u organismos unicelulares huéspedes con la secuencia de ácido nucleico mutante para expresar un polipéptido codificado por la secuencia de ácido nucleico mutante;
 - (d) explorar el polipéptido para determinar al menos una propiedad modificada; y,
 - (e) opcionalmente, si el polipéptido no tiene ninguna actividad deseada de pachulol y 7-epi- α -selineno sintasa variante, repetir las etapas (a) a (d) del procedimiento hasta que se obtenga un polipéptido con una actividad deseada de pachulol y 7-epi- α -selineno sintasa variante;
 - 45 (f) opcionalmente, si se identificó en la etapa (d) un polipéptido que tenía una actividad deseada de pachulol y 7-epi- α -selineno sintasa variante, aislar el ácido nucleico mutante correspondiente obtenido en la etapa (c).

Figura 1

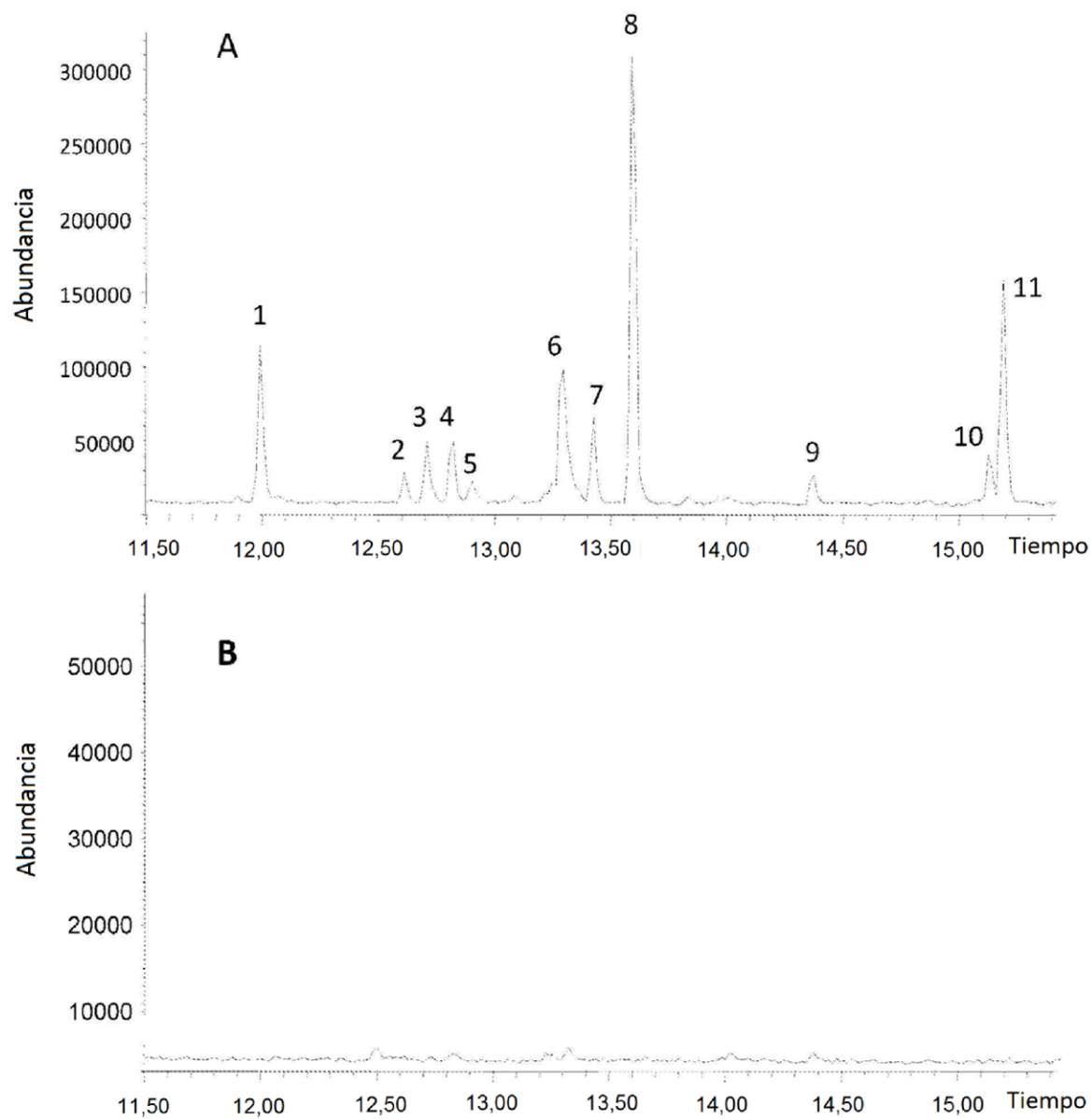


Figura 2

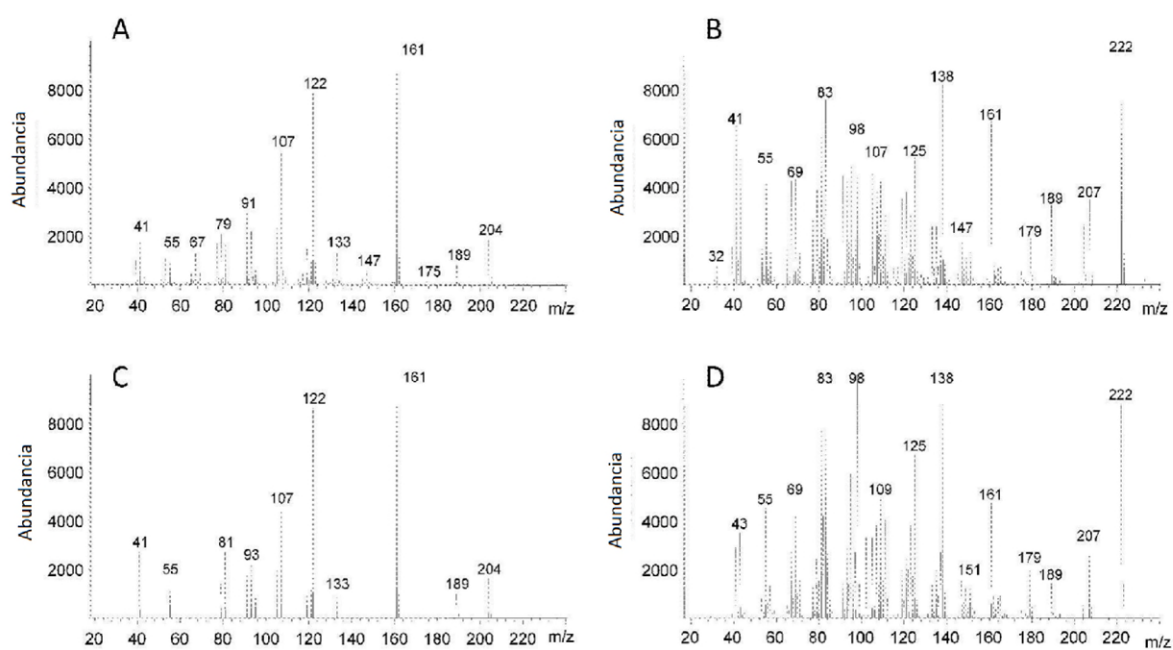


Figura 3

