(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 626 619**

(51) Int. Cl.:
C12N 9/88 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **12.05.2010 PCT/IB2010/052103**

(87) Fecha y número de publicación internacional: **25.11.2010 WO10134004**

(96) Fecha de presentación y número de la solicitud europea: **12.05.2010 E 10721853 (9)**

(97) Fecha y número de publicación de la concesión europea: **15.03.2017 EP 2432877**

(54) Título: **Procedimiento de producción de (+)-zizaeno**

(30) Prioridad:

20.05.2009 WO PCT/IB2009/052118

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
25.07.2017

(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 y
DEGUERRY, FABIENNE

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 626 619 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 (+)-zizaeno

Campo técnico

La presente invención proporciona un procedimiento de producción de (+)-zizaeno, comprendiendo dicho procedimiento poner en contacto al menos un polipéptido con farnesil pirofosfato (FPP). En particular, dicho procedimiento se puede llevar a cabo *in vitro* o *in vivo* para producir (+)-zizaeno, un compuesto que se puede usar como precursor para diversos compuestos útiles 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 o una célula hospedadores no humanos transformados para usarse en el procedimiento de producción de (+)-zizaeno también es un objeto de la presente invención.

Técnica anterior

Los terpenos se encuentran en la mayoría de los organismos (microorganismos, animales y plantas). Estos compuestos están compuestos de cinco unidades de carbono llamadas unidades de isopreno y se clasifican por el número de estas unidades presentes en su estructura. Por tanto, monoterpenos, sesquiterpenos y diterpenos son terpenos que contienen 10, 15 y 20 átomos de carbono respectivamente. Los sesquiterpenos, por ejemplo, son muy conocidos en el reino vegetal. Muchas moléculas de sesquiterpeno son conocidas por sus propiedades de sabor y fragancia y sus efectos cosméticos, medicinales y antimicrobianos. Se han identificado más de 3.000 hidrocarburos de sesquiterpeno y 3.000 sesquiterpenoides y cada año se identifican muchas nuevas estructuras. Los extractos vegetales obtenidos por diferentes medios tales como destilación por arrastre de vapor o extracción por disolvente se usan como fuente de terpenos. Las moléculas de terpeno con frecuencia se usan como tales, pero en algunos casos se usan reacciones químicas para transformar los terpenos en otras moléculas de alto valor.

La producción biosintética de terpenos implica enzimas llamadas terpeno sintasas. Virtualmente hay una infinidad de sesquiterpeno sintasas presentes en el reino vegetal, usando todas el mismo sustrato (farnesil pirofosfato, FPP), pero teniendo diferentes perfiles de producto. Se han clonado los genes y los ADNc que codifican las sesquiterpeno sintasas y se caracterizaron las correspondientes enzimas recombinantes. La biosíntesis de los terpenos en plantas y otros organismos ha sido exhaustivamente estudiada y no se detalla más en el presente documento.

Generalmente, el precio y la disponibilidad de los extractos naturales vegetales son dependientes de la abundancia, la producción de aceite y el origen geográfico de las plantas. Además, la disponibilidad y la calidad de los extractos naturales es muy dependiente del clima y otras condiciones locales que conducen a la variabilidad de año en año, volviendo el uso de tales ingredientes en perfumería de alta calidad muy difícil o incluso imposible algunos años.

El aceite de vetiver es uno de estos extractos naturales. Es un ingrediente de perfumería relativamente caro, que consiste en una mezcla compleja de alcoholes de sesquiterpeno, aldehídos y ácidos que tienen un perfil olfativo complejo. Los constituyentes individuales del aceite de vetiver también podrían ser útiles como ingredientes de perfumería pero su purificación a partir del aceite no es viable a gran escala.

Por lo tanto, sería muy deseable un procedimiento independiente de la planta para producir los constituyentes de aceite de vetiver, pero hasta ahora no está disponible una síntesis química de tales compuestos eficaz respecto al coste.

(+)-Zizaeno es una molécula de sesquiterpeno que se da de manera natural. Se puede usar como precursor para diversos compuestos que son útiles en el campo de la perfumería y aromatizantes, en particular para constituyentes de aceite de vetiver como khusimol, zizaen-12-al y ácido khuzénico. Por lo tanto, sería de gran interés una ruta bioquímica que conduzca a la síntesis de (+)-zizaeno.

El análisis de la composición del aceite de vetiver mostraba que los derivados de zizaeno son constituyentes principales de este aceite y contribuyen al olor de vetiver. Véase, por ejemplo, Weyerstahl y col. (2000), *Flav. Frag. J.*, 15, 395-412. Aun así, el presente documento no proporciona ni siquiera siguiere una secuencia de aminoácidos o nucleótidos que conduzca a la producción de (+)-zizaeno.

La biosíntesis del aceite de vetiver en raíces de vetiver ha sido investigada en Del Giudice y col (2008), “The microbial community of Vetiver root and its involvement into essential oil biogenesis", *Environ. Microbiol.*, 10(10), 2.824-2.841. Esta publicación describe la producción de sesquiterpenos mediante microorganismos aislados de raíces de vetiver y respalda la idea de los microorganismos implicados en la biosíntesis de los sesquiterpenos de vetiver. Nuestros resultados están en contra de las sugerencias hechas en este artículo debido a que demostramos que el zizaeno está producido por una sesquiterpeno sintasa expresada por las propias raíces de vetiver y no por la comunidad microbiana.

En la técnica anterior, nunca se ha revelado una sesquiterpeno sintasa capaz de sintetizar un precursor de los constituyentes del aceite de vetiver, en particular, (+)-zizaeno.

El porcentaje de identidad entre las sesquiterpeno sintasas conocidas y el polipéptido de la invención es muy bajo. La secuencia de proteína más cercana a la (+)-zizaeno sintasa de la invención es una supuesta terpeno sintasa de *Zea mays* (Nº de referencia NCBI ACG24265) que comparte 56 % de identidad de secuencia de aminoácidos con la (+)-zizaeno sintasa de la invención. No se han identificado los productos obtenidos con esta terpeno sintasa potencial. La sintasa completamente caracterizada más cercana es una *(E)-beta*-cariofileno sintasa de *Zea mays* (Nº de referencia NCBI ABY79212), que es únicamente 51 % idéntica a la sintasa de la invención.

Además de la diferencia entre las propias secuencias, también se ha señalado que la estructura y las propiedades de los productos sintetizados por la enzima anteriormente mencionada son muy diferentes de los de (+)-zizaeno. En particular, (*E*)-*beta*-cariofileno no es adecuado como precursor para la producción de constituyentes del aceite de vetiver.

A pesar de exhaustivos estudios de la ciclización del terpeno, el aislamiento y la caracterización de las terpeno sintasas es aún difícil, particularmente en plantas, debido a su baja abundancia, sus frecuentes patrones de expresión transitorios, y la complejidad de su purificación a partir de las mezclas de resinas y compuestos fenólicos en tejidos donde se expresan.

Es un objetivo de la presente invención proporcionar procedimientos para producir (+)-zizaeno de una manera económica, como se indicó anteriormente. Por consiguiente, la presente invención tiene el objetivo de producir (+)-zizaeno mientras que tenga pocos residuos, un procedimiento eficaz de más energía y recurso y mientras que reduzca la dependencia de fueles fósiles. Es un objetivo adicional proporcionar enzimas capaces de sintetizar (+)-zizaeno, el cual es útil como precursor para perfumería y/o ingredientes de aroma.

Abreviaciones usadas

pb	par de bases
kb	kilo base
BSA	albúmina de suero bovino
ADNc	ADN complementario
CTAB	bromuro de cetiltrimetilamonio
DMAPP	dimetilalil difosfato
ADN	ácido desoxirribonucleico
dATP	desoxi adenosina trifosfato
dNTP	desoxi nucleótido trifosfato
DTT	ditiotreitol
EDTA	ácido etilenodiaminotetraacético
FPP	farnesil pirofosfato
GC	cromatografía gaseosa
idi	isopentenil difosfato isomerasa
IPP	isopentenil difosfato
IPTG	isopropil-D-tiogalacto-piranósido
LB	caldo de lisogenia
MOPSO	ácido 3-(N-morfolino)-2-hidroxipropanosulfónico
MS	espectrómetro de masas
mvaK1	mevalonato quinasa
mvaK2	mevalonato difosfato quinasa
PCR	reacción en cadena de la polimerasa
PVP	polivinilpirrolidona
RMCE	intercambio de casete mediado por recombinasa
3'-/5'-RACE	amplificación rápida 3' y 5' de los extremos de ADNc
ARN	ácido ribonucleico
ARNm	ácido ribonucleico mensajero
TE	tris y EDTA
BNL	base de nitrógeno de levadura

Descripción de la invención

La presente invención proporciona un procedimiento para producir biosintéticamente (+)-zizaeno de una manera económica, fiable y reproducible.

Una "sesquiterpeno sintasa" o un "polipéptido que tiene una actividad de sesquiterpeno sintetasa" 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 a partir del precursor de terpeno acíclico FPP.

Como una “(+)-zizaeno sintasa” o como un “polipéptido que tiene una actividad de (+)-zizaeno sintasa”, los investigadores quieren decir en el presente documento un polipéptido capaz de catalizar la síntesis de (+)-zizaeno que comienza a partir de FPP. (+)-Zizaeno puede ser el único producto o puede ser parte de una mezcla de sesquiterpenos.

La capacidad de un polipéptido de catalizar la síntesis de un sesquiterpeno particular (por ejemplo, (+)-zizaeno) se puede confirmar simplemente realizando el ensayo de enzima como se detalla en el Ejemplo 3.

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

Tal como se indica más adelante en el presente documento, “una secuencia de nucleótidos obtenida modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento” abarca cualquier secuencia que se ha obtenido cambiando la secuencia de SEQ ID NO:2, de SEQ ID NO:11 o del complemento de una de estas dos secuencias que usan cualquier procedimiento conocido en la técnica, por ejemplo, introduciendo cualquier tipo de mutaciones tal como mutaciones por deleción, inserción o sustitución. Ejemplos de tales procedimientos se citan en la parte de la descripción en relación a los polipéptidos variantes y los procedimientos para prepararlos.

El porcentaje de identidad entre dos secuencias peptídicas o nucleotídicas es una función del número de residuos de aminoácidos o nucleótidos que son idénticos en las dos secuencias cuando se ha generado un alineamiento de estas dos secuencias. Los residuos idénticos se definen como residuos que son el mismo en las dos secuencias en una posición dada del alineamiento. El porcentaje de la identidad de secuencia, como se usa en el presente documento, se calcula a partir del alineamiento óptimo tomando el número de residuos idénticos entre dos secuencias dividiéndolo por el número total de residuos 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. Se pueden introducir huecos (*gaps*) en una o ambas secuencias en una o más posiciones del alineamiento para obtener el alineamiento óptimo. Entonces, estos huecos se tienen en cuenta como residuos no idénticos para el cálculo del porcentaje de identidad de secuencia.

El alineamiento con el fin de determinar el porcentaje de identidad de secuencia de aminoácidos o ácido nucleicos se puede alcanzar de diversas maneras que usan programas informáticos y, por ejemplo, programas informáticos libres disponibles por toda la red. Preferentemente, el programa BLAST (Tatiana y col, *FEMS Microbiol Lett*., 1999, 174:247-250, 1999) fijado en los parámetros por defecto, disponible en el *National Center for Biotechnology Information* (NCBI) en http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi, se puede usar para obtener un alineamiento óptimo de secuencias peptídicas o nucleotídicas y para calcular el porcentaje de identidad de secuencia.

Por lo tanto, un objeto de la presente invención es un procedimiento para producir (+)-zizaeno que comprende

a) poner en contacto FPP con al menos un polipéptido que tiene una actividad de (+)-zizaeno sintasa y que comprende una secuencia de aminoácidos al menos 70 % idéntica a SEQ ID NO:1;
b) opcionalmente, aislar el (+)-zizaeno producido en la etapa a).

Según una realización preferida, el procedimiento es un procedimiento para producir (+)-zizaeno como producto principal. Según una realización aún más preferida, (+)-zizaeno representa al menos 50 %, preferentemente al menos 60 %, preferentemente al menos 80 %, preferentemente al menos 90 % del producto producido por el procedimiento de la invención.

El procedimiento se puede llevar a cabo *in vitro,* así como *in vivo*, como se explicará con detalle más adelante.

El polipéptido a poner en contacto con FPP *in vitro* se puede obtener por extracción a partir de cualquier organismo que lo exprese, usando tecnologías de extracción de proteína o enzima estándares. Si el organismo hospedador es un organismo unicelular o una célula que libera el polipéptido de la invención en el medio de cultivo, el polipéptido simplemente se puede recoger del medio de cultivo, por ejemplo, por centrifugación, opcionalmente seguido de etapas de lavado y resuspensión en soluciones tampón adecuadas. Si el organismo o célula acumula el péptido dentro de sus células, el polipéptido se puede obtener por alteración o lisis de las células y extracción adicional del polipéptido a partir del lisado celular.

El polipéptido que tiene una actividad de (+)-zizaeno sintasa, o bien en una forma aislada o junto con otras proteínas, por ejemplo, en un extracto de proteína bruto obtenido a partir de células o microorganismos cultivados, a continuación, se puede poner en suspensión en una solución tampón a pH óptimo. Si es adecuado, se pueden añadir sales, BSA y otras clases de cofactores enzimáticos para optimizar la actividad enzimática. Las condiciones apropiadas se describen con más detalle en los Ejemplos de más adelante.

A continuación, se puede añadir el precursor FPP a la suspensión o solución, la cual, a continuación, se incuba a temperatura óptima, por ejemplo, entre 15 y 400C, preferentemente entre 25 y 35 ºC, más preferentemente a 300C.

Después de la incubación, el (+)-zizaeno producido se puede aislar de la solución incubada por procedimientos de aislamiento estándares, tal como la extracción por disolvente y la destilación, opcionalmente después de la separación de los polipéptidos de la solución.

Según 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 una célula hospedadores no humanos capaces de producir FPP y transformados para expresar al menos un polipéptido que comprende una secuencia de aminoácidos al menos 70 % idéntica a SEQ ID NO:1 y que tiene una actividad de (+)-zizaeno sintasa, bajo condiciones conducentes a la producción de (+)-zizaeno.

Según una realización más preferida, el procedimiento comprende, además, antes de la etapa a), transformar un organismo o una célula no humanos capaces de producir FPP con al menos un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos al menos 70 % idéntica a SEQ ID NO:1 y que tiene una actividad de (+)-zizaeno sintasa, de manera que dicho organismo exprese dicho polipéptido.

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 da directamente dentro del organismo o la célula transformados para expresar dicho polipéptido.

Según una realización particular de la invención, el al menos un ácido nucleico que codifica la (+)-zizaeno sintasa comprende una secuencia de nucleótidos 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 SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización más preferida, dicho ácido nucleico comprende la secuencia de nucleótidos SEQ ID NO:2, SEQ ID NO:11 o su complemento. En una realización aún más preferida, dicho ácido nucleico consiste en SEQ ID NO:2, SEQ ID NO:11 o su complemento.

Según una realización más preferida el al menos un ácido nucleico usado en cualquiera de las anteriores realizaciones comprende una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización aún más preferida, dicho al menos un ácido nucleico consiste en una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento.

Según otra realización, el al menos un ácido nucleico se aísla de *Vetiveria zizanoides*.

El organismo o la célula se supone que “expresan” un polipéptido, siempre y cuando el organismo o la célula se transformen para portar un ácido nucleico codificador de dicho polipéptido, este ácido nucleico se transcribe a ARNm y el polipéptido se encuentra en el organismo o la célula hospedadores. El término “expresar” abarca “expresar heterogéneamente” y “sobreexpresar”, refiriéndose esto último a niveles de ARNm, polipéptido y/o actividad de enzima sobre y por encima de lo medido en un organismo o una célula no transformados. Una descripción más detallada de procedimientos adecuados para transformar un organismo o una célula hospedadores no humanos se describirá después en la parte de la memoria que se dedica a tales organismos o células hospedadores no humanos transformados como objetos específicos de la presente invención y en los ejemplos.

Un organismo o una célula particulares se supone que son “capaces de producir FPP” cuando se produce FPP de manera natural o cuando no se produce FPP de manera natural pero se trasforma para producir FPP, o bien antes de la transformación con un ácido nucleico como se describe en el presente documento o junto con dicho ácido nucleico. Por “organismos o células capaces de producir FPP” también se abarca los organismos o las células transformados para producir una cantidad mayor de FPP que el organismo o la célula que se dan de manera natural. Los procedimientos para transformar organismos, por ejemplo, microorganismos, de manera que producen FPP ya son conocidos en la técnica. Tales procedimientos se pueden encontrar, 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), 1.441-1.447 (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 la invención *in vivo*, el organismo o la célula hospedadores se cultivan bajo condiciones conducentes a la producción de (+)-zizaeno. Por consiguiente, si el hospedador es una planta transgénica, se proporcionan las condiciones de crecimiento óptimas, tales como, por ejemplo, condiciones de luz, agua y nutrientes óptimos. Si el hospedador es un organismo unicelular, las condiciones conducentes a la producción de (+)-zizaeno pueden comprender la adición de cofactores adecuados al medio de cultivo del hospedador. Además, se puede seleccionar un medio de cultivo, para maximizar la síntesis de (+)-zizaeno. Las condiciones de cultivo óptimas se describen de una manera más detallada en los siguientes Ejemplos.

Los organismos hospedadores no humanos adecuados para llevar a cabo el procedimiento de la invención *in vivo* pueden ser cualquier organismo multicelular o unicelular. En una realización preferida, el organismo hospedador no humano usado para llevar a cabo la invención *in vivo* es una planta, un procariota o un hongo. Se puede usar cualquier planta, procariota u hongo. Plantas particularmente útiles son aquellas que producen de manera natural altas cantidades de terpenos. En una realización más preferida, la planta se selecciona entre la familia de

Solanaceae, Poaceae, Brassicaceae, Fabaceae, Malvaceae, Asteraceae o Lamiaceae. Por ejemplo, la planta se selecciona entre 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 hospedador no humano usado para llevar a cabo el procedimiento de la invención *in vivo* es un microorganismo. Se puede usar cualquier microorganismo, pero según 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*.

Algunos de estos organismos no producen FPP de manera natural. Para ser adecuados para llevar a cabo el procedimiento de la invención, estos organismos tienen que transformarse para producir dicho precursor. Se pueden transformar así o bien antes de la modificación con el ácido nucleico descrito según cualquiera de las anteriores realizaciones o simultáneamente, como se explicó anteriormente.

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

Según una realización preferida, el al menos un polipéptido que tiene una actividad de (+)-zizaeno 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 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 la SEQ ID NO:1. Según una realización más preferida, dicho polipéptido comprende la secuencia de aminoácidos SEQ ID NO:1. En una realización aún más preferida, dicho polipéptido consiste en SEQ ID NO:1.

Según otra realización preferida, el al menos un polipéptido que tiene una actividad de (+)-zizaeno 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 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 se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización más preferida, el al menos un polipéptido que tiene una actividad de (+)-zizaeno 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 SEQ ID NO:1 obtenida por ingeniería genética, es decir, una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento.

Tal como se usa en el presente documento, el polipéptido se entiende como un polipéptido o fragmento peptídico que abarca las secuencias de aminoácidos identificadas en el presente documento, así como polipéptidos truncados o variantes, siempre y cuando conserven su actividad como se definió anteriormente y que compartan al menos el porcentaje definido de identidad con el fragmento correspondiente de SEQ ID NO:1.

Ejemplos de polipéptidos variantes son proteínas que se dan de manera natural que resultan de sucesos de corte y empalme (*splicing*) de ARNm alterno o de escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a proteólisis incluyen, por ejemplo, diferencias en los terminales N- o C- tras la expresión en diferentes tipos de células hospedadoras, debido a la separación proteolítica de uno o más aminoácidos terminales de los polipéptidos de la invención. La invención también abarca los polipéptidos codificados por un ácido nucleico obtenido por mutación natural o artificial de un ácido nucleico de la invención, descritos más adelante. Por ejemplo, como se detalla en el Ejemplo 4 de más adelante, SEQ ID NO:11 es una variante de SEQ ID NO:2, obtenida por mutación artificial de SEQ ID NO:2, conduciendo a una secuencia de nucleótidos que se optimiza por la expresión en *E. coli* y que codifica la misma (+)-zizaeno sintasa que SEQ ID NO:2 (es decir, SEQ ID NO:1). Las secuencias SEQ ID NO:2 y SEQ ID NO:11 son idénticas al 76 %.

También se pueden usar en los procedimientos de la invención variantes de polipéptidos que resultan de una fusión de secuencias peptídicas adicionales en los extremos terminales amino y carboxilo. En particular, tal fusión puede aumentar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente deseado o sistema de expresión. Tales secuencias peptídicas adicionales pueden ser, por ejemplo, péptidos señal. Por consiguiente, 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. Polipéptidos que resultan de una fusión con otra proteína funcional, tal como otra proteína de la ruta de biosíntesis de terpeno, también se pueden usar ventajosamente en los procedimientos de la invención.

Según otra realización, el al menos un polipéptido que tiene una actividad de (+)-zizaeno 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 *Vetiveria zizanoides*.

Una herramienta importante para llevar a cabo el procedimiento de la invención es el propio polipéptido. Por lo tanto, un polipéptido que tiene una actividad de (+)-zizaeno sintasa y que comprende una secuencia de aminoácidos al menos 70 % idéntica a SEQ ID NO:1 es otro objeto de la presente invención.

Según una realización preferida, el polipéptido es capaz de producir (+)-zizaeno como un producto principal. Según una realización aún más preferida, es capaz de producir una mezcla de sesquiterpenos en la que (+)-zizaeno representa al menos 60 %, preferentemente al menos 80 %, preferentemente al menos 90 % de los sesquiterpenos producidos.

Según una realización más preferida, el polipéptido tiene una actividad de (+)-zizaeno sintasa.

Según una realización preferida, el polipéptido comprende una secuencia de aminoácidos 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 SEQ ID NO:1. Según una realización más preferida, el polipéptido comprende la secuencia de aminoácidos SEQ ID NO:1. Según una realización aún más preferida, el polipéptido consiste en SEQ ID NO:1.

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 se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización más preferida, el polipéptido que tiene una actividad de (+)-zizaeno 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 se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento.

Según otra realización, el polipéptido se aísla de *Vetiveria zizanoides*.

Tal como se usa en el presente documento, el polipéptido se entiende como un polipéptido o fragmento peptídico que abarca las secuencias de aminoácidos identificadas en el presente documento, así como polipéptidos truncados o variantes, siempre y cuando conserven su actividad como se definió anteriormente y que compartan al menos el porcentaje definido de identidad con el correspondiente fragmento de SEQ ID NO:1.

Ejemplos de polipéptidos variantes son proteínas que se dan de manera natural que resultan de sucesos de corte y empalme de ARNm alterno o de escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a proteólisis incluyen, por ejemplo, diferencias en los terminales N- o C- tras la expresión en diferentes tipos de células hospedadores, debido a la separación proteolítica de uno o más aminoácidos terminales de los polipéptidos de la invención. La invención también abarca los polipéptidos codificados por un ácido nucleico obtenido por mutación natural o artificial de un ácido nucleico de la invención, descritos más adelante. Por ejemplo, como se detalla en el Ejemplo 4 de más adelante, SEQ ID NO:11 es una variante de SEQ ID NO:2, obtenida por mutación artificial de SEQ ID NO:2, conduciendo a una secuencia de nucleótidos que se optimiza por la expresión en *E. coli* y que codifica la misma (+)-zizaeno sintasa que SEQ ID NO:2 (es decir, SEQ ID NO:1). Las secuencias SEQ ID NO:2 y SEQ ID NO:11 son idénticas al 76 %.

Los polipéptidos de la invención también abarcan variantes de polipéptido que resultan de una fusión de secuencias peptídicas adicionales en los extremos terminales amino y carboxilo. En particular, tal fusión puede mejorar la expresión de los polipéptidos, ser útil en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente deseado o sistema de expresión. Tales secuencias peptídicas adicionales pueden ser, por ejemplo, un péptido señal. Por consiguiente, la presente invención abarca variantes de los polipéptidos de la invención, tales como las obtenidas por fusión con otros oligo- o polipéptidos y/o las que están unidas a los péptidos señal. Los polipéptidos de la invención también abarcan los polipéptidos resultantes de una fusión con otra proteína funcional, tales como otra proteína de la ruta de biosíntesis de terpeno.

Tal como se mencionó anteriormente, el ácido nucleico codificador del polipéptido de la invención es una herramienta útil para modificar organismos o células hospedadores no humanos destinados a usarse cuando el procedimiento se lleva a cabo *in vivo*.

Por tanto, un ácido nucleico codificador de un polipéptido según cualquiera de las realizaciones anteriormente descritas es también un objeto de la presente invención.

Según una realización preferida, el ácido nucleico comprende una secuencia de nucleótidos 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 SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización más preferida, el ácido nucleico comprende la secuencia de nucleótidos SEQ ID NO:2, SEQ ID NO:11 o su complemento. Según una realización aún más preferida, el ácido nucleico consiste en SEQ ID NO:2, SEQ ID NO:11 o su complemento.

Según otra realización, el ácido nucleico se aísla de *Vetiveria zizanoides*.

El ácido nucleico de la invención se puede definir como que incluye polímeros de desoxirribunucleótido o ribonucleótido en forma de o bien cadena sencilla o doble (ADN y/o ARN). Los términos “secuencia de nucleótidos” también se deberían entender como que comprende una molécula de polinucleótidos o una molécula de oligonucleótidos en la 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 que incluyen las que son sustancialmente libres de material endógeno contaminante. El ácido nucleico de la invención puede ser truncado, siempre y cuando codifique un polipéptido abarcado por la presente invención, como se describió anteriormente.

Según una realización más preferida, el al menos un ácido nucleico según cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento. Preferentemente dicho ácido nucleico consiste en una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO:2, SEQ ID NO:11 o su complemento.

La invención abarca los ácidos nucleicos que comprenden una secuencia obtenida por mutación de SEQ ID NO:2, SEQ ID NO:11 o su complemento, siempre y cuando las secuencias comprendan compartir al menos el porcentaje definido de identidad con los correspondientes fragmentos de SEQ ID NO:2, SEQ ID NO:11 o su complemento y siempre y cuando codifiquen un polipéptido que tenga una actividad de (+)-zizaeno sintasa, como se define en cualquiera de las anteriores realizaciones. Las mutaciones pueden ser cualquier tipo de mutaciones de estos ácidos nucleicos, tales como mutaciones puntuales, mutaciones por deleción, mutaciones por inserción y/o mutaciones por desplazamiento de marco. Un ácido nucleico variante se puede preparar 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 se codifican los aminoácidos por un codón preferido. Debido a la degeneración del código genético, en la que más de un codón puede codificar el mismo aminoácido, las secuencias de ADN múltiples pueden codificar el mismo polipéptido, siendo todas estas secuencias de ADN abarcadas por la invención.

Otra herramienta importante para transformar los organismos o las células hospedadores adecuados para llevar a cabo el procedimiento de la invención *in vivo* es un vector de expresión que comprende un ácido nucleico según cualquier realización de la invención. Por tanto, tal 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 limita, a vectores víricos, 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, la cual controla la iniciación y/o terminación de la transcripción y/o traducción, tal como un promotor, operador o potenciador transcripcional, o un sitio de unión ribosómico del ARNm y, opcionalmente, incluyendo al menos un marcador de selección. Las secuencias de nucleótidos están “operativamente unidas” cuando la secuencia reguladora se relaciona funcionalmente con el ácido nucleico de la invención.

Los vectores de expresión de la presente invención se pueden usar en los procedimientos para preparar un organismo y/o una célula hospedadores genéticamente transformados, en organismos y/o células hospedadoras que portan los ácidos nucleicos de la invención y en los procedimientos para producir o realizar polipéptidos que tienen una actividad de (+)-zizaeno sintasa, como se revela más a continuación.

Los organismos y las células hospedadores no humanos recombinantes transformados para portar al menos un ácido nucleico de la invención de manera que expresen o sobreexpresen de manera heteróloga al menos un polipéptido de la invención también son herramientas muy útiles para llevar a cabo el procedimiento de la invención. Por tanto, tales organismos y células hospedadores no humanos son otro objeto de la presente invención.

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

Los organismos hospedadores no humanos de la invención pueden ser cualquier organismo multicelular o unicelular no humano. En una realización preferida, el organismo hospedador 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 las que producen de manera natural altas cantidades de terpenos. En una realización más preferida, la planta se selecciona entre la familia de *Solanaceae, Poaceae, Brassicaceae, Fabaceae, Malvaceae, Asteraceae o Lamiaceae.* Por ejemplo, la planta se selecciona entre el género *Nicotiana, Solanum, Sorghum, Arabidopsis, Brassica (colza), Medicago* (alfalfa), *Gossypium* (algodón), *Artemisia, Sylvia y Mentha.* Preferentemente, la planta pertenece a las especies de *Nicotiana tabacum.*

En una realización más preferida el organismo hospedador no humano es un microorganismo. Cualquier microorganismo es adecuado para la presente invención, pero según 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*.

Células eucariotas superiores aisladas también pueden ser transformadas, en lugar de organismos completos. Como células eucariotas superiores, los inventores quieren decir en el presente documento 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 hospedador se sometió a ingeniería genética para comprender una, dos o más copias de cada uno de los ácidos nucleicos requeridos en cualquiera de la realización anteriormente descrita. Preferentemente, el término “transformado” se refiere a hospedadores que expresan de manera heteróloga los polipéptidos codificados por el ácido nucleico con el que son transformados, así como que sobreexpresan dichos polipéptidos. Por consiguiente, en una realización, la presente invención proporciona un organismo transformado, en el cual los polipéptidos se expresan en cantidades mayores que en el mismo organismo no transformado así.

Hay diversos procedimientos conocidos en la técnica para la creación de organismos o células hospedadores transgénicos tales como plantas, hongos, procariotas o cultivos de células eucariotas superiores. Vectores de clonación y expresión apropiados para su uso con hospedadores celulares bacterianos, fúngicos, de levadura, vegetales y mamíferos están descritos, por ejemplo, en Pouwels y col., “Cloning Vectors: A Laboratory Manual”, 1985, Elsevier, Nueva York y Sambrook y col., “Molecular Cloning: A Laboratory Manual”, 2º edición, 1989, Cold Spring Harbor Laboratory Press. Vectores de clonación y expresión para plantas superiores y/o células vegetales en particular están disponibles para los expertos en la técnica. Véase, por ejemplo, Schardl y col. *Gene* 61: 1-11, 1987.

Procedimientos para transformar organismos o células hospedadores para portar ácidos nucleicos transgénicos son familiares para los expertos en la técnica. Para la creación de plantas transgénicas, por ejemplo, procedimientos actuales incluyen: electroporación de protoplastos vegetales, transformación mediada por liposoma, transformación mediada por agrobacterium, transformación mediada por polietilenglicol, bombardeo de partícula, microinyección de células vegetales y transformación usando virus.

En una realización, el ADN transformado está integrado en un cromosoma de un organismo y una célula hospedadores no humanos de manera que da como resultado un sistema recombinante estable. Cualquier procedimiento de integración cromosómica conocido en la técnica se puede usar en la práctica de la invención, incluyendo, pero no limitándose a, intercambio de casete mediado por recombinasa (RMCE), inserción cromosómica específica a sitio vírica, adenovirus e inyección pronuclear.

Para llevar a cabo el procedimiento para producir (+)-zizaeno *in vitro,* como se expuso anteriormente en el presente documento, es muy ventajoso proporcionar un procedimiento de producción de al menos un polipéptido que tenga una actividad de (+)-zizaeno 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 o una célula hospedadores no humanos según cualquier realización de la invención;
b) aislar el polipéptido del organismo o la célula hospedadores no humanos en una etapa a).

Según una realización preferida, dicho procedimiento comprende, además, antes de la etapa a), transformar un organismo o una célula hospedadores no humanos con al menos un ácido nucleico según cualquier realización de la invención, de manera que dicho organismo exprese el polipéptido codificado por dicho ácido nucleico.

Se pueden usar un ácido nucleico según cualquiera de las realizaciones anteriormente descritas.

La transformación y el cultivo del organismo o la célula hospedadores no humanos se puede llevar a cabo como se describió anteriormente para el procedimiento de producción de (+)-zizaeno *in vivo*. La etapa b) se puede realizar usando cualquier técnica bien conocida en la técnica para aislar un polipéptido particular a partir de un organismo o una célula.

Una “variante de polipéptido” como se denomina en el presente documento significa un polipéptido que tiene una actividad de (+)-zizaeno 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 deleciones, inserciones o sustituciones.

Las variantes pueden comprender secuencias sustituidas de manera conservadora, significando que un residuo de aminoácidos dado está reemplazado por un residuo que tiene características fisioquímicas similares. Ejemplos de sustituciones conservadoras incluyen sustituciones de un residuo alifático por otro, tal como Ile, Val, Leu, o Ala por otro, o sustituciones de un residuo 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 tales sustituciones se pueden calcular usando matrices de puntuación de sustitución tales como PAM-120, PAM-200, y PAM-250 como se discute en *Altschul, J. Mol. Biol.*, 1991, 219, 555-565. Se conocen bien otras de tales sustituciones conservadoras, por ejemplo, sustituciones de regiones enteras que tienen características de hidrofobicidad similares.

La invención también abarca las variantes peptídicas que se dan de manera natural. Ejemplos de tales variantes son proteínas que resultan de sucesos de corte y empalme de ARNm alterno o de escisión proteolítica de los polipéptidos descritos en el presente documento. Variaciones atribuibles a proteólisis incluyen, por ejemplo, diferencias en los terminales N- o C- tras la expresión en diferentes tipos de células hospedadoras, debido a la separación proteolítica de uno o más aminoácidos terminales de los polipéptidos codificados por las secuencias de la invención.

Las variantes de los polipéptidos de la invención se pueden usar para conseguir, por ejemplo, actividad enzimática aumentada o reducida deseada, regioquímica o estereoquímica modificada, o utilización de sustrato o distribución de producto alterada, afinidad incrementada para el sustrato, especificidad mejorada para la producción de uno o más compuestos deseados, velocidad incrementada de la reacción enzimática, mayor actividad o estabilidad en un ambiente específico (pH, temperatura, disolvente, etc.) o nivel de expresión mejorada en un sistema de expresión deseado. Una variante o un mutante dirigido a sitio se pueden producir mediante cualquier procedimiento conocido en la técnica. Variantes y derivados de polipéptidos nativos se pueden obtener aislando variantes que se dan de manera natural, o la secuencia de nucleótidos de variantes, de otra o la misma línea o especie vegetal, o programando artificialmente mutaciones de secuencias de nucleótidos que codifican los polipéptidos de la invención. Alteraciones de la secuencia de aminoácidos nativa se pueden efectuar por alguno de un número de procedimientos convencionales.

Las variantes de polipéptido resultantes de una fusión de secuencias peptídicas adicionales en los extremos terminales amino y carboxilo de los polipéptidos de la invención pueden ser usadas para aumentar la expresión de los polipéptidos, ser útiles en la purificación de la proteína o mejorar la actividad enzimática del polipéptido en un ambiente o sistema de expresión deseado. Tales secuencias peptídicas adicionales pueden ser, por ejemplo, péptidos señal. Por consiguiente, la presente invención abarca variantes de los polipéptidos de la invención, tal como las obtenidas por fusión con otros oligo- o polipéptidos y/o las que están unidas a péptidos señal. Los polipéptidos de fusión abarcados por la invención también comprenden polipéptidos de fusión resultantes de una fusión de otras proteínas funcionales, tales como otras proteínas de la ruta de biosíntesis de terpeno.

Por lo tanto, en una realización, la presente invención proporciona un procedimiento para preparar un polipéptido variante que tiene una actividad de (+)-zizaeno 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 anteriormente expuestas;
(b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
(c) transformar células u organismos unicelulares hospedadores con la secuencia de ácidos nucleicos mutante para expresar un polipéptido codificado por la secuencia de ácidos nucleicos mutante;
(d) cribado (*screening*) del polipéptido para al menos una propiedad modificada; y
(e) opcionalmente, si el polipéptido no tiene la actividad de (+)-zizaeno sintasa variante deseada, repetir las etapas (a) a (d) del proceso hasta que se obtenga un polipéptido con una actividad de (+)-zizaeno sintasa variante deseada;
(f) opcionalmente, si en la etapa (d) se identificó un polipéptido que tenía una actividad de (+)-zizaeno sintasa variante deseada, aislar el correspondiente ácido nucleico mutante obtenido en la etapa (c).

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

En la etapa (b), se puede crear un gran número de secuencias de ácidos nucleicos mutantes, por ejemplo, por mutagénesis aleatoria, mutagénesis específica a sitio, o transposición (*shuffling*) de ADN. Los procedimientos detallados de transposición genética se encuentran en Stemmer, “DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution”. *Proc. Natl. Acad. Sci. USA*., 1994, 91(22): 10747-1075. En resumen, transposición de ADN se refiere a un proceso de recombinación aleatoria de secuencias conocidas *in vitro*, implicando al menos dos ácidos nucleicos seleccionados para recombinación. Por ejemplo, se pueden introducir mutaciones en loci concretos mediante la síntesis de oligonucleótidos que contienen una secuencia mutante, flanqueada por sitios de restricción que permiten la ligación a fragmentos de la secuencia nativa. Después de la ligación, la secuencia reconstruida resultante codifica un análogo que tiene la inserción, sustitución o deleción de aminoácidos deseada. Alternativamente, se pueden emplear procedimientos de mutagénesis específica a sitio dirigida a oligonucleótido para proporcionar un gen alterado en el que los codones predeterminados se pueden alterar por sustitución, deleción o inserción.

Por consiguiente, el polipéptido que comprende SEQ ID NO:1 se puede recombinar con algún otro ácido nucleico codificador de sesquiterpeno sintasa, por ejemplo, aislado de un organismo distinto de *Vetiveria zizanoides*. Así, se pueden obtener y separar ácidos nucleicos mutantes, los cuales se pueden usar para transformar una célula hospedadora según procedimientos estándares, por ejemplo, tales como los descritos en los presentes ejemplos.

En la etapa (d), el polipéptido obtenido en la etapa (c) se somete a cribado para al menos una propiedad modificada, por ejemplo, una actividad enzimática modificada deseada. Ejemplos de actividades enzimáticas deseadas, por las cuales se puede someter a cribado un polipéptido expresado, incluyen actividad enzimática aumentada o reducida, medida por el valor K_M o $V_{máx}$, regioquímica o estereoquímica modificada y utilización de sustrato o distribución de producto alterada. El cribado de la actividad enzimática se puede realizar según procedimientos familiares a los expertos en la técnica y los descritos en los presentes ejemplos.

La etapa (e) prevé la repetición de las etapas (a) a (d) del proceso, lo cual se puede realizar preferentemente en paralelo. Por consiguiente, creando un número significante de ácidos nucleicos mutantes, se pueden transformar muchas células hospedadoras con diferentes ácidos nucleicos mutantes a la vez, permitiendo el posterior cribado de un número elevado de polipéptidos. Las posibilidades de obtener un polipéptido variante deseado se pueden incrementar así en el criterio del experto en la técnica.

Todas las publicaciones mencionadas en esta solicitud son para revelar y describir los procedimientos y/o materiales con respecto a lo cual se citan las publicaciones.

Descripción de los dibujos

Figura 1: análisis GC-MS de sesquiterpenos producidos por la (+)-zizaeno sintasa de *Vetiveria zizanoides* (VzZS). A, cromatograma iónico total. B, espectro de masas del pico principal (1). C, espectro de masas de un patrón de (+)-zizaeno auténtico. El pico 1 se identificó como el (+)-zizaeno. Los picos 2, 3 y 4 se identificaron como prezizaeno, α-funebreno y β-funebreno respectivamente. Los picos marcados con S eran compuestos de sesquiterpeno no identificados.
Figura 2: Estructura de los principales compuestos de sesquiterpeno producidos por la (+)-zizaeno sintasa de *Vetiveria zizanoides* (VzZS).
Figura 3: Perfil del producto obtenido por la producción *in vitro* (A) e *in vivo* (B) de (+)-zizaeno. El pico principal en cada análisis es (+)-zizaeno.

Realizaciones específicas de la invención o Ejemplos

A continuación, la invención se describirá con más detalle por medio de los siguientes Ejemplos.

Ejemplo 1

Extracción de ARN y construcción de genoteca de ADNc

Se obtuvieron plantas de vetiver (*Vetiveria zizanoides*) de un vivero (*The Austral Plants Company*, Les Avirons, La Isla de Reunión, Francia). Las plantas se cultivaron en macetas en un invernadero en la *Lullier Agronomy research Station* (Suiza) y se propagaron vegetativamente dividiendo matas de seis meses a un año de edad. Para recolectar las raíces, las plantas se sacaron de las macetas y se aclararon con agua del grifo.

Para la preparación de la genoteca de ADNc, se combinaron raíces de diversas plantas: plantas jóvenes (4 a 6 meses después de la propagación), plantas viejas con un sistema radicular denso bien desarrollado (1 a 2 años después de la propagación) y plantas jóvenes secadas a temperatura ambiente durante 24 a 36 horas después de sacarlas de las macetas. Las raíces se separaron de la parte aérea de las plantas y se congelaron en nitrógeno líquido. Primero se trocearon bruscamente en nitrógeno líquido usando un Batidora Waring (Waring Laboratory, Torrington, EEUU) y, a continuación, se pulverizaron a un polvo fino usando un mortero y la mano del mortero. Se extrajo el ARN total siguiendo el procedimiento descrito en Kolosova y col (Kolosova, Miller, Ralph, Ellis, Douglas, Ritland y Bohlmann, “Isolation of high-quality RNA from gymnosperm and angiosperm trees”. *J. Biotechniques*, 36(5), 821-4, 2004) con las siguientes modificaciones. Se usó un volumen de 20 ml de tampón de extracción para 2 gramos de tejido pulverizado y el tampón de extracción se complementó con 2 % (p/v) de PVP (polivinilpirrolidona, Sigma-Aldrich). Para la extracción de CTAB (bromuro de cetiltrimetilamonio, Sigma-Aldrich), el gránulo de ácido nucleico se sometió a resuspensión en 2 ml de tampón TE (Tris-HCl 10 mM, pH 8, EDTA 1 mM) y la extracción se realizó con 2 ml de NaCl 5 M y 1 ml de CTAB al 10 %. Para la precipitación de isopropanol, se disolvió el gránulo de ácido nucleico en 500 µl de TE. El gránulo de ARN final se sometió a resuspensión en 50 µl de agua.

Se preparó una genoteca de ADNc de doble cadena usando el SMART™ PCR cDNA Synthesis Kit (Clontech Laboratories, Mountain View, CA) según las instrucciones del fabricante y usando la transcriptasa inversa SuperScript™ II RNAse H (Invitrogen, Carlsbad, CA) para la etapa de transcripción inversa. Se usó una cantidad de 1 µg de ARN total de tejido subterráneo de vetiver como patrón para la síntesis de ADNc y se realizaron 15 ciclos para el procedimiento de amplificación. Se cargó la genoteca sobre un gel de agarosa al 1 % y se eluyeron los fragmentos de tamaños que oscilaban desde 1,3 a 3 kb. Para la secuenciación se usó 270 ng de esta genoteca de ADNc.

Ejemplo 2

Secuenciación de genoteca de ADNc y amplificación de un ADNc de sesquiterpeno sintasa

Se usó la tecnología de secuenciación paralela a gran escala de fragmentos pequeños de ADN desarrollada por Illumina (San Diego, California) para secuenciar la genoteca completa de ADNc. La preparación del ADN para secuenciar, la secuenciación y la reunión de las lecturas se realizaron por Fasteris SA (Plan-les-Ouates, Suiza). La genoteca de ADNc se trató siguiendo el Genomic Sample Prep Kit (Illumina) y se secuenció sobre el sistema del Genome Analyzer (Illumina). Se obtuvieron un total de 4,2 millones de lecturas de 35 pb (de los cuales 3,6 millones eran secuencias únicas). Estas lecturas se reunieron usando EDENA 2.1.1, un programa informático que encuentra solapamientos entre las lecturas y reúne de nuevo contigos (Hernandez D., François P., Farinelli L., Østerås M., y Schrenzel J., "De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer". *Genome Res*. 18(5), 802-809, 2008). Después de eliminar contigos más cortos de 100 bases, se obtuvieron 4.324 contigos únicos con una longitud máxima de 1.882 pb. Se realizó otra reunión usando el programa Velvet 1.0 ((Zerbino y Birney (2008), "Velvet: algorithms for de novo short read assembly using de Bruijn graphs". *Genome Res*. 18(5), 821-829), proporcionando 9.264 contigos únicos de longitud entre 100 y 2.006 bases.

Se compararon todos los contigos generados frente una base de datos de secuencias de proteína (que contenían una selección de 7.000 secuencias de proteína vegetal) usando el algoritmo Blastx (Altschul y col, *J. Mol. Biol.* 215, 403-410, 1990). Se retuvieron los contigos que mostraban homología de secuencia significante con sesquiterpeno sintasas vegetales y la homología se confirmó más realizando, para cada contigo seleccionado, una investigación de blast frente a las secuencias de proteína no redundantes de NCBI (NCBI; http://www.ncbi.nlm.nih.gov). De este modo, 15 contigos se confirmaron como fragmentos de ADNc codificador de sesquiterpeno.

Uno de los contigos seleccionados (VzCtg306, SEQ ID NO:3) era de 1.090 pb de longitud y las comparaciones de secuencia con terpeno sintasas de longitud completa mostraron que el extremo 3' y el extremo 5' estaban perdidos. Se diseñaron dos cebadores directos (ctg306-3R1 (SEQ ID NO:4) y ctg306-3R2 (SEQ ID NO:5)) y dos cebadores inversos (ctg306-5R1 (SEQ ID NO:6) y ctg306-5R2, (SEQ ID NO:7)) a partir de esta secuencia y se usaron para la Amplificación Rápida de los Extremos de ADNc (RACE). Se usó el SMART™ RACE cDNA Amplification Kit (Clontech Laboratories, Mountain View, CA) con la transcriptasa inversa PrimeScript (TaKaRa Bio, Shiga, Japón). Así, se prepararon un conjunto de ADNc Preparado por 5'RACE SMART™ y un ADNc Preparado por 3'RACE SMART™ cada uno a partir de 1,2 µg de ARN total de raíz de vetiver. Para la 5'RACE, se realizó una primera ronda de PCR con los cebadores UPM (Clontech Laboratories) y el cebador ctg306-5R1 (SEQ ID NO:6) seguido de una segunda ronda de PCR con el cebador NUP (Clontech Laboratories) y el cebador ctg306-5R2 (SEQ ID NO:7). Para el 3'RACE, se realizó una primera ronda de PCR con los cebadores UPM (Clontech Laboratories) y el cebador ctg306-3R1 (SEQ ID NO:4) seguido de una segunda ronda de PCR con el cebador NUP (Clontech Laboratories) y el cebador ctg306-3R2 (SEQ ID NO:5). Las amplificaciones se realizaron en las condiciones detalladas en el manual de fabricante (Clontech):

La combinación del 5' y 3'RACE siguió la reconstitución de un ADNc de 1.925 pb (VzZS, SEQ ID NO:8) que contenía un marco de lectura abierto de 1.668 pb (SEQ ID NO:2) codificador de una proteína de 555 aminoácidos de longitud (SEQ ID NO:1).

Ejemplo 3

Expresión heteróloga y caracterización enzimática

El marco de lectura abierto VzZS de longitud completa (VzZS-ORF, SEQ ID NO:2) se amplificó a partir del conjunto de ADNc Preparado por 5'RACE SMART™ usando el cebador ctg306-inicio (SEQ ID NO:9) y ctg306-parada (SEQ ID NO:10). La amplificación de este ADNc para las construcciones de expresión se realizó usando la *Pfu* ADN polimerasa (Promega, Madison, WI, USA), en un volumen final de 50 µl que contenía 5 µl de *Pfu* ADN polimerasa 10x tampón, 200 µM cada dNTP, 0,4 µM cada cebador directo e inverso, 2,9 unidades de *Pfu* ADN polimerasa y 2,5 µl del ADNc (preparado como se describió anteriormente). Las condiciones de ciclación térmica eran las siguientes: 1,5 min a 95 ºC; 30 ciclos de 45 s a 95 ºC, 30 s a 54 ºC y 4 min a 72 ºC; y 10 min a 72 ºC.

Se insertaron los productos de PCR en el vector pET101/D-TOPO usando el Champion pET101 Directional TOPO Expression Kit (Invitrogen, Carlsbad, CA) siguiendo las instrucciones de fabricación. Se seleccionaron diversos clones y los insertos de plásmido secuenciados para confirmar que la secuencia era idéntica a la secuencia obtenida por RACE.

Se usó el plásmido pET101-VzZS para transformar células de *E. coli* BI21(DE3) (Novagen, Madison, WI). Se usaron colonias únicas de células transformadas para inocular 5 ml de medio LB. Después de 5 a 6 horas de incubación a 37 ºC, los cultivos se transfirieron a una incubadora a 20 ºC y se dejaron 1 hora para equilibrado. A continuación, la expresión de la proteína se indujo por la adición de IPTG 1 mM y el cultivo se incubó durante la noche a 20 ºC. Al día siguiente, las células se recogieron por centrifugación, se sometieron a resuspensión en 0,1 volumen de MOPSO 50 mM pH 7, glicerol al 10 % y se lisaron por sonicación. Los extractos se aclararon por centrifugación (30 min a 20.000 g) y los sobrenadantes que contenían las proteínas solubles se usaron para experimentos adicionales.

Se usaron los extractos de proteína de *E. coli* brutos que contenían la proteína recombinante para la caracterización de las actividades enzimáticas. Se sintetizó farnesil-difosfato (FPP) como se describe por Keller, R.K., y Thompson, R., *J. Chromatogr*. 645(1), 161-167, 1993. Los ensayos se realizaron en 1 a 4 ml de MOPSO 50 mM pH 7, glicerol al 10 %, DTT 1 mM, $MgCl_2$ 10 mM en presencia de 10 a 100 µM de sustrato y 0,1 a 0,5 mg de proteína bruta. Los tubos se incubaron 12 a 24 horas a 30 ºC y se extrajeron dos veces con un volumen de pentano. Después de la concentración bajo un flujo de nitrógeno, los extractos se analizaron por GC y GC-MS y se compararon con extractos de ensayo con proteínas control. El análisis por GC se realizó sobre un sistema Agilent 6890 Series GC equipado con un detector de ionización por llama que usa un diámetro interno de 0,25 mm por columna de capilaridad SPB-1 30 m (Supelco, Bellefonte, PA). El gas vehículo era He a un flujo constante de 1 ml/min. La temperatura de horno inicial era de 80 ºC (mantenida 1 min) seguido de un gradiente de 10 ºC/min hasta 300 ºC. El análisis por GC-MS se realizó en las mismas condiciones y los espectros se registraron en un detector de masas Agilent 5975.

En estas condiciones, la proteína recombinante codificada por el ADNc de VzZS produjo un sesquiterpeno principal que representaba el 75 % de la mezcla de sesquiterpeno producida. El producto principal se identificó como (+)-zizaeno mediante el enfrentamiento del espectro de masas y el índice de retención con patrones auténticos y datos publicados (Joulain, D., y König, W.A., "The Atlas of Spectral Data of Sesquiterpene Hydrocarbons", EB Verlag, Hamburg, 1998). La enzima produce también 6,9 % de prezizaeno, 2,8 % de α-funebreno, 2,7 % de β-funebreno y al menos otros 3 sesquiterpenos a proporciones entre 0,85 y 8,7 % (Figura 2). Así, el ADNc de VzZS aislado de *Vetiveria zizanoides* codificó una (+)-zizaeno sintasa (SEQ ID NO:1) que producía el precursor de hidrocarburo de los sesquiterpenos más abundantes en raíces de vetiver (Khusimol, zizaen-12-al, ácido khuzénico). La enzima también produjo como productos secundarios algunos de los precursores de los constituyentes menores de las raíces de vetiver.

Ejemplo 4

Uso de la proteína VzZS recombinante para la producción *in vivo* de (+)-zizaeno en bacterias

Para la expresión óptima de la VzZS en *E. coli*, se rediseñó la secuencia de ADN del ORF del ADNc de VzCtg306 para tener en cuenta el uso del codón del hospedador y otros parámetros que influyen en la estabilidad del ARNm y su traducción. La secuencia optimizada (VzZS-opt, SEQ ID NO:11) se diseñó y se sintetizó con los sitios de restricción *Nde*I y *Kpn*I y los extremos 3' y 5' (DNA 2.0, Menlo Park, CA, USA) y se subclonaron en el plásmido pETDuet-1 (Novagene, Madison, WI) proporcionando el plásmido pETDuet-VzZS-opt.

Para evaluar la producción *in vivo* de (+)-zizaeno, se transformaron células de *E. coli* con el plásmido pETDuet-VzZS-opt y se evaluó la producción de sesquiterpenos a partir del conjunto de FPP endógeno. Para incrementar la productividad de las células, también se expresaron en las mismas células una FPP sintasa y los genes codificadores de una ruta de mevalonato parcial. Estos últimos genes se organizaron en un operón único y codificaron una mevalonato quinasa (mvaK1), una fosfomevalonato quinasa (mvaK2), una mevalonato difosfato descarboxilasa (MvaD) y una isopentenil difosfato isomerasa (idi) y mevalonato exógeno convertido en isopentenil difosfato (IPP) y dimetilalil difosfato (DMAPP), los dos sustratos de la FPP sintasa.

El gen de la FPP sintasa de levadura se amplificó a partir del ADN genómico de *S. cerevisiae* usando los cebadores FPPy_NcoI (SEQ ID NO:12) y FPPy-Eco (SEQ ID NO:13). El ADN genómico se aisló a partir de *S. cerevisiae* usando el Qiagen RNA/DNA Maxi kit (Qiagen AG, Basel, Suiza). El PCR se realizó con la *Pfu* ADN polimerasa (Promega AG, Dubendorf, Suiza) en un volumen final de 50 µl que contenía 0,4 µl de cada cebador, dNTPs 200 µM, 0,5 µl de ADN polimerasa, 5 µl de ADN genómico de *S. cerevisiae*. La condición de ciclación de PCR era la siguiente: 90 s a 95 ºC; 28 ciclos de 45 s a 95 ºC, 30 s a 54 ºC y 4 min a 72 ºC; 10 min a 72 ºC. El ADN amplificado se ligó como fragmento NdeI-EcoRI en el primer sitio de clonación múltiple (MCS1) del plásmido pACYCDuet-1 (Novagen, Madison, WI) proporcionando el plásmido pACYCDuet-FPPs que portaba el gen de FPP bajo el control de un promotor T7.

Se amplificó un operón que contenía los genes codificadores de mvaK1, mvaK2, MvaD e idi a partir del ADN genómico de *Streptococcus pneumoniae* (ATCC BAA-334, LGC Standards, Molsheim, Francia) con los cebadores MVA-up1-inicio (SEQ ID NO:14) y MVA-up2-parada (SEQ ID NO:15). La PCR se realizó usando la ADN polimerasa PfuUltra™ II Fusión HS (Stratagene, Agilent Technologies Inc., Santa Clara, CA, EEUU). La composición de la mezcla de PCR era según las instrucciones del fabricante. La condición de ciclación térmica era 2 min a 95 ºC; 30 ciclos de 20 s a 95 ºC, 20 s a 58 ºC y 90 s a 72 ºC; y 3 min a 72 ºC. El fragmento de 3,8 kb se purificó sobre un gel de agarosa y se ligó usando el In-Fusion™ Dry-Down PCR Cloning Kit (Clontech Laboratories) en el segunda MCS del plásmido pACYCDuetFPPs digerido con *Nde*I y *Xho*I proporcionando el plásmido pACYCDuet-4506. Las secuencias de los dos insertos se secuenciaron completamente para excluir cualquier mutación.

Se transformaron células de *E. coli* BL21 Star™(DE3) (Invitrogen, Carlsbad, CA) con el plásmido pETDuet-VzZS-opt o se transformaron conjuntamente con el mismo plásmido y con el plásmido pACYCDuet-4506. Las células transformadas se seleccionen sobre placas de LB-agarosa con carbenicilina (50 µg/ml) y cloranfenicol (34 µg/ml). Se usaron colonias únicas para inocular 5 ml de medio LB líquido complementado con los mismos antibióticos. El cultivo se incubó durante la noche a 37 ºC. Al día siguiente se inocularon con 0,2 ml del cultivo de noche 2 ml de medio TB

complementado con los mismos antibióticos. Después de 6 horas de incubación a 37 ºC, el cultivo se enfrió a 28 ºC y IPTG 1 mM, 2 mg/ml de mevalonato (preparado disolviendo mevalonolactona (Sigma) en NaOH 0,5 N a una concentración de 1 g/ml e incubando la solución durante 30 min a 37 ºC) y se añadieron 0,2 ml de decano a cada tubo. Los cultivos se incubaron durante 48 horas a 28 ºC. A continuación, los cultivos se extrajeron dos veces con 2 volúmenes de acetato de etilo, la fase orgánica se concentró a 500 µl y se analizó por GC-MS como se describió anteriormente en el Ejemplo 3. Con las células productoras de la (+)-zizaeno sintasa, la FPP sintasa y las cuatro enzimas de la ruta de mevalonato, se obtuvo una productividad de 0,1 mg/ml y el perfil de producto era idéntico a los perfiles observados con los ensayos *in vitro* (Figura 3).

Este ejemplo muestra que una célula de *E. coli* transformada con una (+)-zizaeno sintasa, como se define en la presente invención, es capaz de producir (+)-zizaeno. Las otras enzimas con las que la célula de *E. coli* se transforma no son esenciales para la producción de (+)-zizaeno. De hecho, el (+)-zizaeno también se produce cuando una célula de *E. coli* se transforma con solo la (+)-zizaeno sintasa, pero en cantidades menores. Las otras enzimas con las que se transforma la célula de *E. coli* se añaden con el único fin de incrementar la cantidad del precursor disponible para la (+)-zizaeno sintasa.

Ejemplo 5

Uso de la proteína VzZS recombinante para la producción *in vivo* de (+)-zizaeno en levadura

Para la producción *in vivo* de sesquiterpenos en células de levadura, una cepa de *Saccharomyces cerevisiae* (YNP5) en la cual se ha regulado a la baja el gen ERG9 (codificador de la escualeno sintasa, la enzima que convierte FPP a escualeno) mediante el reemplazo del promotor ERG9 nativo con el promotor MET3, proporcionando así una cepa con biosíntesis de ergosterol reducida y mayor conjunto de FPP disponible para las sesquiterpeno sintasas (Asadollahi, M.A., Maury, J., Møller., K, Nielsen, K.F., Schalk, M., Clark, A., y Nielsen, J., *Biotechnology and Bioengineering* 99(3), 666-677, 2008).

El ADNc de VzZS se amplificó a partir del plásmido pETDuet-VzZS-opt con los cebadores Ctg306_inicio_opt (SEQ ID NO:16) y Ctg306_parada_opt (SEQ ID NO:17). La PCR se realizó con la *Pfu* ADN polimerasa (Promega) usando las siguientes condiciones de ciclación térmica: 90 s a 94 °C; 35 ciclos de 30 s a 94 °C, 30 s a 55 °C, 4 min a 72°C; y 10 min a 72 °C. El ADNc amplificado se purificó y, para añadir salientes 3' A, se incubaron 15 min a 72 ºC en presencia de dATP 0,2 mM y 1 U de HotStart Taq ADN polimerasa en el tampón apropiado (Qiagen). El ADNc se ligó en el plásmido pYES2.1/V5-His-TOPO® usando el Kit de expresión pYES2.1 TOPO® TA (Invitrogen, Carlsbad, CA). Los plásmidos se seleccionaron para correcta secuencia y orientación del inserto y se usaron para transformar las células de levadura YNP5 usando el S.c. EasyComp™ Transformation Kit (Invitrogen, Carlsbad, CA).

Se usaron una colonia única de cepas de levadura transformada para inocular 20 ml de medio BNL (5 g/l de $(NH_4)_2SO_4$; 3 de g/l de KH_2PO_4; 0.5 g/l de $MgSO_4.7\ H_2O$; 1 ml/l de solución de metal traza) complementado con glucosa al 2 %. El cultivo se incubó durante 24 horas a 28 ºC. Las células se recuperaron por centrifugación y se sometieron a resuspensión en 20 ml de medio BNL complementado con galactosa al 2 %. Después de 1 hora de cultivo, se añadieron metionina a concentración final 0,5 mM y 2 ml de decano al cultivo. Después de 24 horas de incubación a 28 ºC, los cultivos se sometieron a extracción con acetato de etilo y se analizaron por GC-MS como se describe en el Ejemplo 3. La cantidad total de los sesquiterpenos producidos por las células de levadura en estas condiciones se estimaron en 25 mg/l.

LISTADO DE SECUENCIAS

<110> Firmenich SA

<120> Procedimiento para producir (+)-zizaeno

<130> 7640 PCT zizaeno sintasa

<160> 17

<170> PatentIn versión 3.5

<210> 1
<211> 555
<212> PRT
<213> *Vetiveria zizanoides*

<400> 1

Met Ala Thr Thr Ala Ala Phe Cys Leu Thr Thr Thr Pro Ile Gly Glu
1 5 10 15

Pro Val Cys Arg Arg Gln Tyr Leu Pro Thr Val Trp Gly Ser Phe Phe
20 25 30

Leu Thr Tyr Gln Pro Cys Thr Pro Glu Glu Val Gln Ser Met Glu Glu
35 40 45

Arg Ala Leu Ala Lys Lys Thr Glu Val Gly Arg Met Leu Gln Glu Val
50 55 60

Ala Ala Ser Ser Asn Leu Ala Arg Lys Leu Gly Leu Val Asp Glu Leu
65 70 75 80

Glu Arg Leu Gly Val Asp Tyr His Tyr Lys Thr Glu Ile Asn Asp Leu
85 90 95

Leu Gly Ala Ile Tyr Asn Gly Lys Asp Asp Asp Asn Gly Gly Ser Asp
100 105 110

Asp Asp Leu Tyr Ile Thr Ser Leu Lys Phe Tyr Leu Leu Arg Lys His
115 120 125

Gly Tyr Ala Leu Ser Ser Asp Val Phe Leu Lys Phe Arg Asp Glu Gln
130 135 140

Gly Asn Ile Ser Ser Asp Asp Val Lys Cys Leu Ile Met Leu Tyr Asp
145 150 155 160

Ala Ser His Leu Arg Ile His Glu Glu Lys Ile Leu Asp Asn Ile Asn
165 170 175

Ser Phe Thr Lys Ser Cys Leu Gln Ser Val Leu Glu Thr Asn Leu Glu
180 185 190

Pro Ala Leu Gln Glu Glu Val Arg Cys Thr Leu Glu Thr Pro Arg Phe
195 200 205

Arg Arg Val Glu Arg Ile Glu Ala Lys Arg Phe Ile Ser Ala Tyr Glu
210 215 220

Lys Asn Ile Ala Arg Asp Asp Ala Leu Leu Glu Phe Ala Arg Leu Asp
225 230 235 240

Tyr Asn Ile Val Gln Ile Leu Tyr Cys Lys Glu Leu Lys Glu Leu Thr
245 250 255

Val Trp Trp Lys Glu Phe His Ser Arg Thr Asn Leu Thr Phe Ala Arg
260 265 270

Asp Arg Ile Val Glu Met Tyr Phe Trp Val Met Ala Ile Ile Tyr Glu
275 280 285

Pro Cys Tyr Ser Tyr Ser Arg Ile Trp Val Thr Lys Met Phe Leu Ser
290 295 300

Val Ala Leu Leu Asp Asp Ile Tyr Asp Asn Tyr Thr Ser Thr Glu Glu
305 310 315 320

Ser Asn Ile Phe Thr Thr Ala Met Glu Arg Trp Asp Val Lys Ala Thr
325 330 335

Glu Gln Leu Pro Ala Asn Met Arg Thr Phe Tyr Asp Tyr Leu Ile Cys
340 345 350

Thr Thr Asp Glu Val Val Glu Glu Leu Lys Leu Gln Asn Asn Lys Asn
355 360 365

Ala Glu Leu Val Lys Lys Val Leu Ile Asp Ala Ala Lys Cys Tyr His
370 375 380

Ser Glu Val Lys Trp Arg Asp Asp His Tyr Val Pro Asn Asp Val Gly
385 390 395 400

Glu His Leu Gln Leu Ser Met Arg Ser Ile Ala Ala Met His Ser Ile
405 410 415

Asn Phe Val Phe Ile Ser Leu Gly Ala Val Cys Thr Arg Glu Ala Val
420 425 430

Glu Cys Ala Phe Thr Tyr Pro Lys Ile Ile Arg Gly Ile Cys Val His
435 440 445

Ala Arg Ile Ser Asn Asp Ile Ala Ser His Glu Arg Glu Gln Ala Ser
450 455 460

Glu His Met Ala Ser Thr Leu Gln Thr Cys Met Lys Gln Tyr Gly Ile
465 470 475 480

Thr Val Glu Glu Ala Ala Glu Lys Leu Arg Val Ile Asn Glu Glu Ser
485 490 495

Trp Met Asp Ile Val Glu Glu Cys Leu Tyr Lys Asp Gln Tyr Pro Leu
500 505 510

Ala Leu Ser Glu Arg Val Val Ala Phe Ala Gln Ser Ile Cys Phe Met
515 520 525

Tyr Asn Gly Val Asp Lys Tyr Thr Ile Pro Ser Lys Leu Lys Asp Ser
530 535 540

Leu Asp Ser Leu Tyr Val Asn Leu Ile Pro Val
545 550 555

<210> 2
<211> 1668
<212> ADN
<213> *Vetiveria zizanoides*

<400> 2

atggcgacga ctgccgcctt ctgcctcacc accactccga tcggcgagcc agtctgtcgc 60
cggcagtacc tcccaaccgt ctggggcagc ttcttcctca cctaccagcc atgcacgccg 120
gaagaggtcc agtccatgga ggagagggct ctggccaaga agacggaggt ggggcgcatg 180
ttgcaggagg tcgccgcctc cagtaacctc gcacggaagc tgggccttgt cgatgagcta 240
gagcggctcg gggtggacta tcactacaag acggagatca acgacttgct gggtgccatt 300
tataatggca aggacgacga taatggaggt tctgatgacg acctctatat cacatcgctt 360
aagttctatc tgctcaggaa gcatgggtac gctttatctt cagatgtgtt tctgaagttc 420
agagatgagc aaggaaatat ttcaagtgat gatgtgaaat gcctgatcat gttgtatgat 480
gcctcacatt tgaggattca tgaggagaaa attcttgaca acatcaacag tttcaccaag 540
agctgcctcc aatcagtttt agaaacaaat ttggaaccgg ctctccaaga ggaggtgcgg 600
tgcacattgg agacacctcg attcagaagg gttgagagaa tcgaagcgaa acgctttatc 660
tcagcgtacg aaaagaacat agcacgagat gacgccctac tagagtttgc aaggctggac 720

tacaatatcg tgcaaattct ctactgcaag gagctgaaag aacttacagt atggtggaag 780
gagttccatt cacggacaaa tctgacattt gcacgagata gaattgtgga gatgtatttc 840
tgggtcatgg caattattta cgagccttgt tactcgtatt cacggatatg ggttacaaaa 900
atgtttctat ccgtggcatt gttggatgac atctatgaca attatacgag cacagaggag 960
agcaatatct ttactacggc catggaaagg tgggatgtga aggccaccga acaactgcca 1020
gcaaacatga ggacattcta cgattactta atttgtacaa cagatgaggt cgtagaagaa 1080
ttgaaacttc agaataataa gaatgctgaa ttagtcaaga aagtgctgat tgacgccgct 1140
aaatgctacc attcggaggt caaatggcgt gatgaccact acgtccctaa tgatgttgga 1200
gagcacctgc agctttcaat gcgaagcatt gcagctatgc actccatcaa ctttgtcttc 1260
atttcactgg gagctgtgtg tactagggag gcggttgagt gtgctttcac ttatccaaaa 1320
attattagag gtatatgtgt tcacgcacgt attagtaacg atatcgcgtc acatgagcga 1380
gaacaagctt cggagcatat ggcatcaacg ttgcaaactt gcatgaagca gtatgggatt 1440
acagtagagg aagctgctga aaagctcaga gtaataaacg aggagtcatg gatggacatc 1500
gttgaggaat gcctttataa ggaccagtat cccctggcgc tttcggagag ggtggtggct 1560
tttgcacaat caatatgttt catgtacaat ggtgtagata aatacaccat accatcaaaa 1620
ctcaaggaca gtctagactc attgtacgtc aatttgattc cagtttga 1668

<210> 3
<211> 1090
<212> ADN
<213> *Vetiveria zizanoides*

<400> 3

tgtgaaatgc ctgatcatgt tgtatgatgc ctcacatttg aggattcatg aggagaaaat 60

tcttgacaac atcaacagtt tcaccaagag ctgcctccaa tcagttttag aaacaaattt 120

ggaaccggct ctccaagagg aggtgcggtg cacattggag acacctcgat tcagaagggt 180

tgagagaatc gaagcgaaac gctttatctc agcgtacgaa aagaacatag cacgagatga 240

cgccctacta gagtttgcaa ggctggacta caatatcgtg caaattctct actgcaagga 300

gctgaaagaa cttacagtat ggtggaagga gttccattca cggacaaatc tgacatttgc 360

acgagataga attgtggaga tgtatttctg ggtcatggca attatttacg agccttgtta 420

ctcgtattca cggatatggg ttacaaaaat gtttctatcc gtggcattgt tggatgacat 480

ctatgacaat tatacgagca cagaggagag caatatcttt actacggcca tggaaaggtg 540

ggatgtgaag gccaccgaac aactgccagc aaacatgagg acattctacg attacttaat 600

ttgtacaaca gatgaggtcg tagaagaatt gaaacttcag aataataaga atgctgaatt 660

agtcaagaaa gtgctgattg acgccgctaa atgctaccat tcggaggtca aatggcgtga 720

tgaccactac gtccctaatg atgttggaga gcacctgcag ctttcaatgc gaagcattgc 780

agctatgcac tccatcaact ttgtcttcat ttcactggga gctgtgtgta ctagggaggc 840

ggttgagtgt gctttcactt atccaaaaat tattagaggt atatgtgttc acgcacgtat 900

tagtaacgat atcgcgtcac atgagcgaga acaagcttcg gagcatatgg catcaacgtt 960

gcaaacttgc atgaagcagt atgggattac agtagaggaa gctgctgaaa agctcagagt 1020

aataaacgag gagtcatgga tggacatcgt tgaggaatgc ctttataagg accagtatcc 1080

cctggcgctt 1090

<210> 4
<211> 29
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 4
catagcacga gatgacgccc tactagagt 29

<210> 5
<211> 26
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 5
gcaaggctgg actacaatat cgtgca 26

<210> 6
<211> 28
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 6
cggtggcctt cacatcccac ctttccat 28

<210> 7
<211> 28
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 7
aagatattgc tctcctctgt gctcgtat 28

<210> 8
<211> 1925
<212> ADN
<213> *Vetiveria zizanoides*

<400> 8

gaagcaaagc catctgccgt gctatcactc tagcaaatta tactgagtgg ataaacttaa 60
taccacacca gacgttttgc attcatggcg acgactgccg ccttctgcct caccaccact 120
ccgatcggcg agccagtctg tcgccggcag tacctcccaa ccgtctgggg cagcttcttc 180
ctcacctacc agccatgcac gccggaagag gtccagtcca tggaggagag ggctctggcc 240
aagaagacgg aggtggggcg catgttgcag gaggtcgccg cctccagtaa cctcgcacgg 300
aagctgggcc ttgtcgatga gctagagcgg ctcggggtgg actatcacta caagacggag 360
atcaacgact tgctgggtgc catttataat ggcaaggacg acgataatgg aggttctgat 420
gacgacctct atatcacatc gcttaagttc tatctgctca ggaagcatgg gtacgcttta 480
tcttcagatg tgtttctgaa gttcagagat gagcaaggaa atatttcaag tgatgatgtg 540
aaatgcctga tcatgttgta tgatgcctca catttgagga ttcatgagga gaaaattctt 600
gacaacatca acagtttcac caagagctgc ctccaatcag ttttagaaac aaatttggaa 660
ccggctctcc aagaggaggt gcggtgcaca ttggagacac ctcgattcag aagggttgag 720
agaatcgaag cgaaacgctt tatctcagcg tacgaaaaga acatagcacg agatgacgcc 780
ctactagagt ttgcaaggct ggactacaat atcgtgcaaa ttctctactg caaggagctg 840
aaagaactta cagtatggtg gaaggagttc cattcacgga caaatctgac atttgcacga 900
gatagaattg tggagatgta tttctgggtc atggcaatta tttacgagcc ttgttactcg 960
tattcacgga tatgggttac aaaaatgttt ctatccgtgg cattgttgga tgacatctat 1020
gacaattata cgagcacaga ggagagcaat atctttacta cggccatgga aaggtgggat 1080
gtgaaggcca ccgaacaact gccagcaaac atgaggacat tctacgatta cttaatttgt 1140
acaacagatg aggtcgtaga agaattgaaa cttcagaata ataagaatgc tgaattagtc 1200
aagaaagtgc tgattgacgc cgctaaatgc taccattcgg aggtcaaatg gcgtgatgac 1260
cactacgtcc ctaatgatgt tggagagcac ctgcagcttt caatgcgaag cattgcagct 1320
atgcactcca tcaactttgt cttcatttca ctgggagctg tgtgtactag ggaggcggtt 1380
gagtgtgctt tcacttatcc aaaaattatt agaggtatat gtgttcacgc acgtattagt 1440
aacgatatcg cgtcacatga gcgagaacaa gcttcggagc atatggcatc aacgttgcaa 1500
acttgcatga agcagtatgg gattacagta gaggaagctg ctgaaaagct cagagtaata 1560
aacgaggagt catggatgga catcgttgag gaatgccttt ataaggacca gtatcccctg 1620
gcgctttcgg agagggtggt ggcttttgca caatcaatat gtttcatgta caatggtgta 1680
gataaataca ccataccatc aaaactcaag gacagtctag actcattgta cgtcaatttg 1740

attccagttt gacgacatcg catcaagtat taattctagg cttaatataa tgccagtaaa 1800

catcatatgt aagggatatt tactttcgtg aatccaaata atttgagggg tcctgtgttc 1860

ctcttaccaa ggatatgtca tcaagttgaa aaatatagcc agcaaaaaaa aaaaaaaaaa 1920

aaaaa 1925

<210> 9
<211> 26
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 9
caccatggcg acgactgccg ccttct 26

<210> 10
<211> 22
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 10
cgatgtcgtc aaactggaat ca 22

<210> 11
<211> 1671
<212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia rediseñada para *E. coli* obtenida de SEQ ID NO:8

<400> 11

atggctacta cggctgcttt ttgtctgact actactccta ttggtgagcc ggtgtgtcgt 60

cgccaatacc tgccgactgt gtggggtagc ttcttcttga cctaccaacc gtgcaccccg 120

gaagaggtgc aaagcatgga ggagcgtgca ttggccaaaa agaccgaggt tggccgtatg 180

ttgcaagagg tggcggccag cagcaacctg gcccgcaagc tgggtttggt tgacgagctg 240

gagcgtctgg gtgtggacta ccattacaaa accgagatta acgatctgct gggcgcgatc 300

tataatggta aggacgatga caacggcggt agcgacgatg atctgtacat tacgtctctg 360

aaattctatc tgctgcgtaa gcatggttat gcattgagca gcgatgtttt tctgaaattt 420

cgcgacgagc agggtaacat tagctccgac gacgtcaagt gcctgatcat gttgtacgac 480

gcgagccact tgcgtattca tgaggagaaa atcctggata atatcaattc cttcacgaag 540

agctgcctgc aaagcgttct ggaaaccaat ctggaaccgg cgctgcagga agaggttcgc 600

tgtacgctgg agacgccgcg tttccgccgt gtcgaacgta ttgaagcaaa acgcttcatt 660

agcgcgtacg agaagaacat tgcgcgtgac gacgctctgc tggagttcgc gcgcctggac 720

tacaacattg tccagattct gtattgcaaa gagctgaaag aactgacggt gtggtggaag 780

gagttccaca gccgcaccaa tctgacgttt gcacgtgatc gcatcgtgga aatgtacttt 840

tgggtcatgg caattatcta cgagccttgc tactcgtata gccgcatttg ggtcaccaaa 900

atgtttttga gcgtcgcact gctggatgat atttacgaca actatactag cacggaggaa 960

agcaacattt tcaccacggc gatggagcgc tgggacgtga aagcgaccga acaactgccg 1020

gcgaatatgc gtacctttta tgactatctg atctgcacca ccgacgaagt tgttgaagaa 1080

ctgaaactgc agaataacaa gaacgcggaa ctggttaaga aggtgctgat cgacgcggcc 1140

aaatgctatc atagcgaagt taaatggcgc gacgatcact acgttccgaa tgacgttggt 1200

gaacatctgc agctgtcgat gcgttccatt gcggcgatgc acagcattaa tttcgtgttc 1260

atttccctgg gcgccgtgtg tactcgcgag gcagtcgaat gcgcgtttac ttatccgaag 1320

atcattcgcg gtatttgtgt ccatgcccgt atctccaatg acattgcctc ccatgagcgc 1380

gagcaggcat ccgagcacat ggctagcact ctgcaaacct gcatgaagca gtatggcatt 1440

accgttgagg aggcggcaga aaaactgcgt gtgattaacg aggagagctg gatggatatc 1500

gtcgaggagt gcctgtacaa ggatcagtat ccgctggcat tgtctgagcg cgtcgttgca 1560

tttgcccaga gcatttgttt tatgtacaat ggcgtggaca agtacaccat tccgagcaag 1620

ctgaaggata gcctggatag cttgtatgtc aacctgattc cggtttaatg a 1671

<210> 12
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 12
ctagccatgg cttcagaaaa agaaattagg 30

<210> 13
<211> 40
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 13
ccggaattcc tatttgcttc tcttgtaaac tttgttcaag 40

<210> 14
<211> 42
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 14
aaggagatat acatatgaca aaaaaagttg gtgtcggtca gg 42

<210> 15
<211> 43
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 15
ctttaccaga ctcgagttac gccttttca tctgatcctt tgc 43

<210> 16
<211> 25
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 16
atggctacta cggctgcttt ttgtc 25

<210> 17
<211> 28
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador

<400> 17
tcattaaacc ggaatcaggt tgacatac 28

REIVINDICACIONES

1. Un polipéptido que tiene una actividad de (+)-zizaeno sintasa y que comprende una secuencia de aminoácidos al menos 70 % idéntica a SEQ ID NO:1.

2. El polipéptido de la reivindicación 1, **caracterizado porque** comprende una secuencia de aminoácidos al menos 80 %, preferentemente al menos 90 % idéntica a SEQ ID NO:1.

3. El polipéptido de la reivindicación 1, **caracterizado porque** comprende la secuencia de aminoácidos SEQ ID NO:1.

4. El polipéptido de la reivindicación 1, **caracterizado porque** consiste en SEQ ID NO:1.

5. Un ácido nucleico que codifica un polipéptido según una cualquiera de las reivindicaciones 1 a 4, **caracterizado porque** comprende una secuencia de nucleótidos al menos 75 % idéntica a SEQ ID NO:2, SEQ ID NO:11 o su complemento.

6. El ácido nucleico de la reivindicación 5, **caracterizado porque** comprende la secuencia de nucleótidos SEQ ID NO:2, SEQ ID NO:11 o su complemento.

7. El ácido nucleico de la reivindicación 5, **caracterizado porque** consiste en SEQ ID NO:2, SEQ ID NO:11 o su complemento.

8. Un vector de expresión que comprende el ácido nucleico de una cualquiera de las reivindicaciones 5 a 7.

9. El vector de expresión de la reivindicación 8, en la forma de un vector vírico, un bacteriófago o un plásmido.

10. El vector de expresión de la reivindicación 8 o 9, incluyendo 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, operador o potenciador transcripcional o un sitio de unión ribosómico de ARNm y, opcionalmente, incluyendo al menos un marcador seleccionado.

11. Un organismo o una célula hospedadores no humanos transformados para portar al menos un ácido nucleico según una cualquiera de las reivindicaciones 5 a 7, de manera que exprese o sobreexprese de manera heteróloga al menos un polipéptido según una cualquiera de las reivindicaciones 1 a 4.

12. El organismo hospedador no humano de la reivindicación 11, **caracterizado porque** es una planta, un procariota o un hongo.

13. El organismo hospedador no humano de la reivindicación 11, **caracterizado porque** es un microorganismo, preferentemente una bacteria o levadura.

14. El organismo hospedador no humano de la reivindicación 13, **caracterizado porque** dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

15. La célula hospedadora no humana de la reivindicación 11, **caracterizada porque** es una célula vegetal o una célula fúngica.

16. Un procedimiento para la producción de (+)-zizaeno que comprende a) poner en contacto FPP con al menos un polipéptido según una cualquiera de las reivindicaciones 1 a 4; b) opcionalmente, aislar el (+)-zizaeno producido en la etapa a).

17. El procedimiento de la reivindicación 16, **caracterizado porque** la etapa a) comprende cultivar un organismo o una célula hospedadores no humanos según una cualquiera de las reivindicaciones 11 a 14, bajo condiciones conducentes a la producción de (+)-zizaeno.

18. El procedimiento de la reivindicación 17, **caracterizado porque** comprende, además, antes de la etapa a), transformar un organismo o una célula hospedadores no humanos capaces de producir FPP con al menos un ácido nucleico según una cualquiera de las reivindicaciones 5 a 7, de manera que dicho organismo exprese el polipéptido codificado por dicho ácido nucleico.

19. Un procedimiento para producir al menos un polipéptido según cualquiera de las reivindicaciones 1 a 4 que comprende

a) cultivar un organismo o una célula hospedadores no humanos según una cualquiera de las reivindicaciones 11 a 14;
b) aislar el polipéptido del organismo o la célula hospedadores no humanos cultivados en la etapa a).

20. El procedimiento de la reivindicación 19, **caracterizado porque** comprende, además, antes de la etapa a), transformar un organismo o una célula hospedadores no humanos con al menos un ácido nucleico según una

cualquiera de las reivindicaciones 5 a 7, de manera que dicho organismo exprese el polipéptido codificado por dicho ácido nucleico.

21. Un procedimiento de preparación de un polipéptido variante que tiene una actividad de (+)-zizaeno sintasa que comprende las etapas de:

(a) seleccionar un ácido nucleico según una cualquiera de las reivindicaciones 5 a 7;
(b) modificar el ácido nucleico seleccionado para obtener al menos un ácido nucleico mutante;
(c) transformar células u organismos unicelulares hospedadores con la secuencia de ácidos nucleicos mutante para expresar un polipéptido codificado por la secuencia de ácidos nucleicos mutante;
(d) cribado del polipéptido para al menos una propiedad modificada; y,
(e) opcionalmente, si el polipéptido no tiene la actividad de (+)-zizaeno sintasa variante deseada, repetir las etapas (a) a (d) del proceso hasta que se obtenga un polipéptido con una actividad de (+)-zizaeno sintasa variante deseada;
(f) opcionalmente, si en la etapa (d) se identifica un polipéptido que tiene una actividad de (+)-zizaeno sintasa variante deseada, aislar el correspondiente ácido nucleico mutante obtenido en la etapa (c).

Figura 1

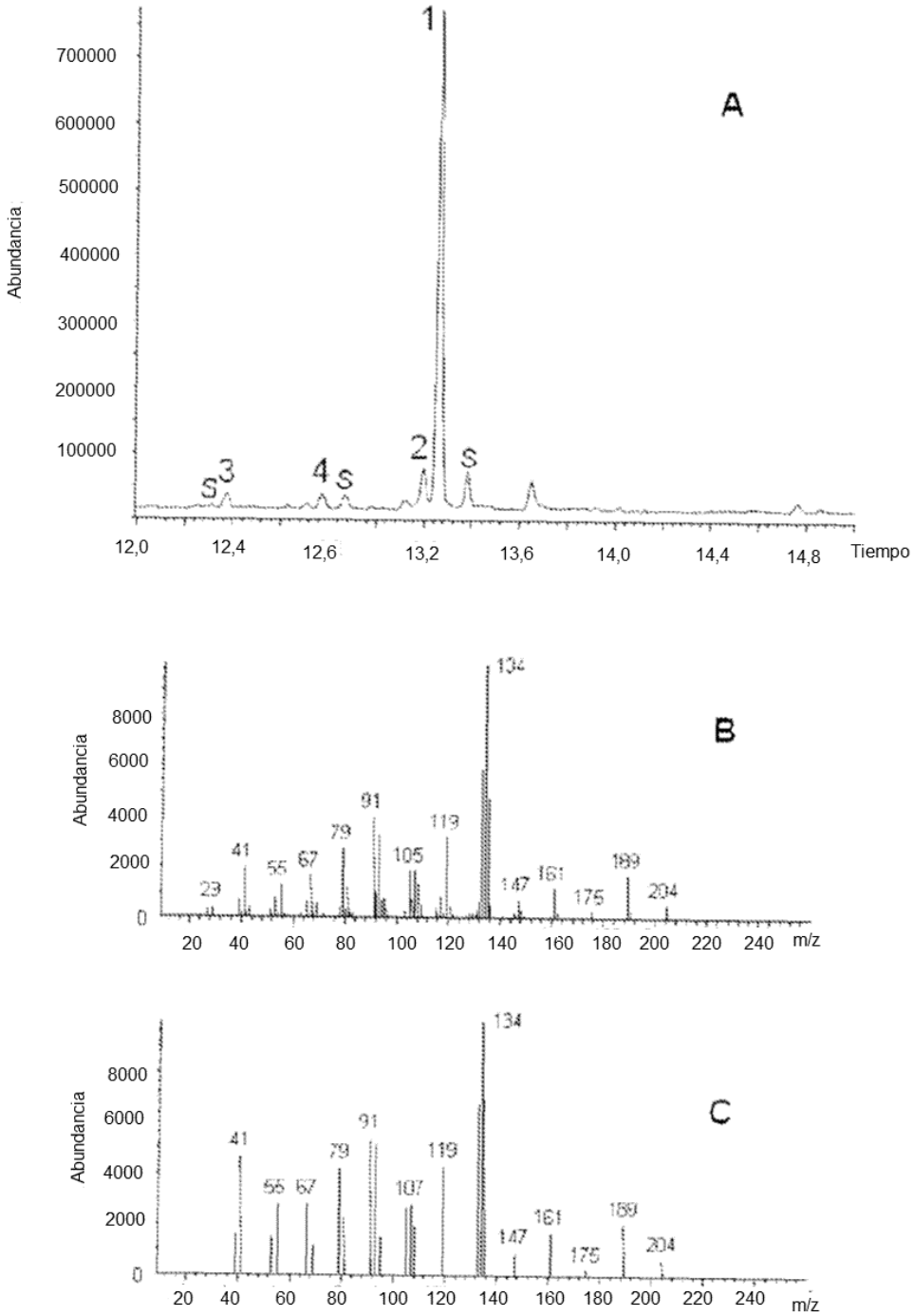
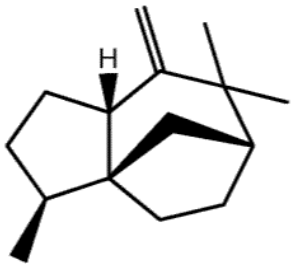
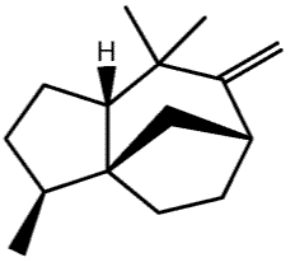

A
Abundancia
700000
600000
500000
400000
300000
200000
100000
1
2
3
4
S
12,0
12,4
12,6
13,2
13,6
14,0
14,4
14,8
Tiempo
B
Abundancia
8000
6000
4000
2000
0
134
91
79
119
41
55
67
105
23
147
161
175
189
204
20
40
60
80
100
120
140
160
180
200
220
240
m/z
C
Abundancia
8000
6000
4000
2000
0
134
91
41
79
119
55
67
107
147
161
175
189
204
20
40
60
80
100
120
140
160
180
200
220
240
m/z

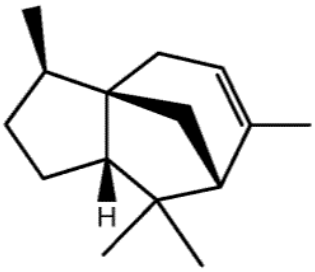
Figura 2



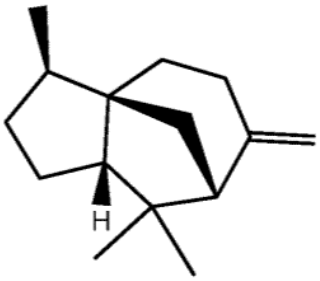
(+)- zizaeno



prezizaeno

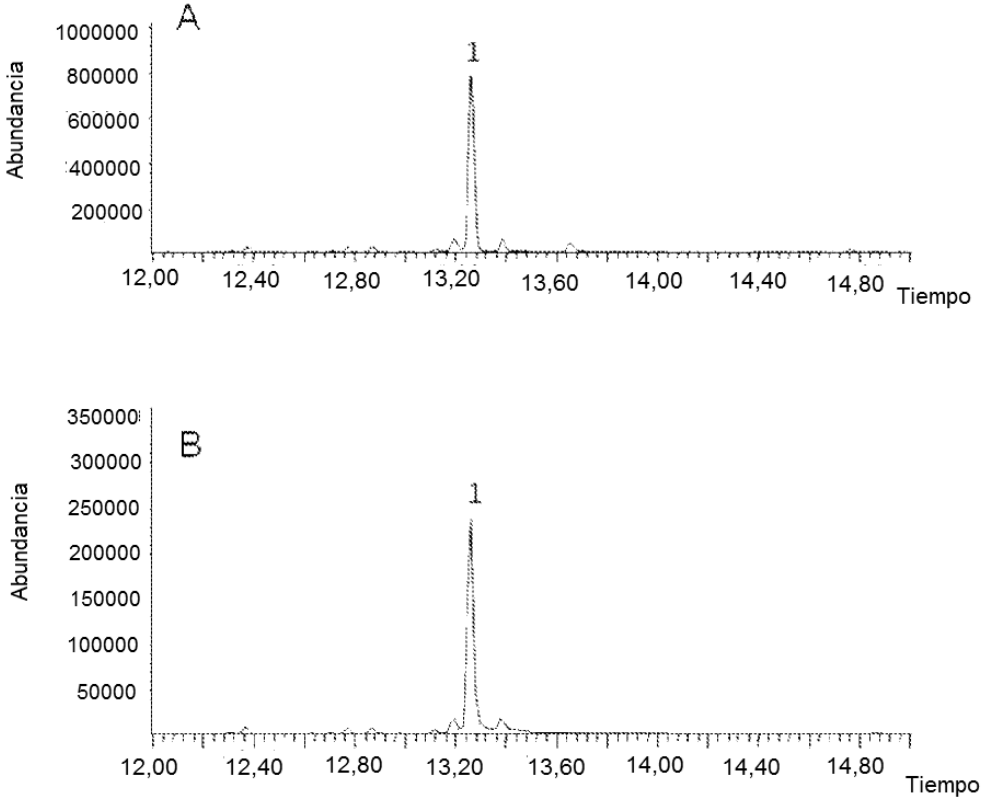


α-funebreno



β-funebreno

Figura 3


A
1
Abundancia
1000000
800000
600000
400000
200000
12,00
12,40
12,80
13,20
13,60
14,00
14,40
14,80
Tiempo
B
1
Abundancia
350000
300000
250000
200000
150000
100000
50000
12,00
12,40
12,80
13,20
13,60
14,00
14,40
14,80
Tiempo