

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 949 068**

51 Int. Cl.:

C12N 9/02	(2006.01)
C12N 9/16	(2006.01)
C12P 7/04	(2006.01)
C12N 9/10	(2006.01)
C12N 9/88	(2006.01)
C12P 5/00	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.04.2018 PCT/EP2018/060889**

87 Fecha y número de publicación internacional: **08.11.2018 WO18202578**

96 Fecha de presentación y número de la solicitud europea: **27.04.2018 E 18718492 (4)**

97 Fecha y número de publicación de la concesión europea: **07.06.2023 EP 3619303**

54 Título: **Sesquiterpeno sintasas para la producción de drimenol y mezclas de los mismos**

30 Prioridad:

03.05.2017 WO PCT/CN2017/082803
07.06.2017 EP 17174791

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.09.2023

73 Titular/es:

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

72 Inventor/es:

LI, PAN;
WANG, QI;
HE, XIU-FENG y
HAEFLIGER, OLIVIER

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

Observaciones:

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

ES 2 949 068 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

Sesquiterpeno sintasas para la producción de drimenol y mezclas de los mismos

5 **Campo técnico**

En el presente documento se proporcionan métodos bioquímicos para producir drimenol y compuestos relacionados y derivados y mezclas que comprenden drimenol, cuyo método comprende el uso de nuevos polipéptidos.

10 **Antecedentes**

Los terpenos se encuentran en la mayoría de los organismos (microorganismos, animales y plantas). Estos compuestos están formados por cinco unidades de carbono 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, se encuentran ampliamente en el reino vegetal. Muchas moléculas de sesquiterpenos son conocidas por sus propiedades aromáticas y de fragancia y sus efectos cosméticos, medicinales y antimicrobianos. Se han identificado numerosos hidrocarburos sesquiterpénicos y sesquiterpenoides.

20 La producción biosintética de terpenos implica enzimas llamadas terpeno sintasas. Hay numerosas sesquiterpeno sintasas presentes en el reino vegetal, que utilizan todas ellas el mismo sustrato (farnesil difosfato, FPP), pero con diferentes perfiles de producto. Se han clonado genes y ADNc que codifican sesquiterpeno sintasas y se han caracterizado las enzimas recombinantes correspondientes.

25 Actualmente, las principales fuentes de drimenol son plantas que contienen drimenol de forma natural; sin embargo, los contenidos de drimenol en estas fuentes naturales son bajos. Se han creado enfoques de síntesis química, pero aún son complejos y no rentables. Sigue siendo necesario descubrir nuevos terpenos, terpeno sintasas y métodos más rentables para producir drimenol y derivados del mismo y mezclas que comprenden drimenol.

30 Los documentos WO 2015/169871 y WO 2015/176959 divulgan cada uno un método para producir drimenol que comprende poner en contacto un precursor de farnesil difosfato acíclico con una sesquiterpeno sintasa (drimenol sintasa) y aislar el drimenol. Sin embargo, estos documentos no divulgan ni sugieren una drimenol sintasa que tenga una secuencia de aminoácidos de la SEQ ID NO: 2 de la presente invención.

35 **Sumario**

En el presente documento se proporciona un método para producir drimenol o una mezcla que comprende drimenol que comprende:

40 a. poner en contacto un farnesil difosfato (FPP) acíclico con un polipéptido que tiene actividad sesquiterpeno sintasa, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o comprende la secuencia de aminoácidos de la SEQ ID NO:2; y

45 b. opcionalmente aislar el drimenol.

También se proporciona un método que comprende transformar una célula hospedadora o un organismo hospedador no humano con un ácido nucleico que codifica un polipéptido que tiene actividad drimenol sintasa, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 2 o comprende la SEQ ID NO: 2.

También se proporciona un método que comprende además cultivar un organismo hospedador no humano o una célula hospedadora capaz de producir FPP y transformada para expresar un polipéptido, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad de secuencia con la SEQ ID NO: 2 en condiciones que permitan la producción del polipéptido.

En un aspecto, el drimenol producido a partir de los métodos anteriores se aísla.

60 En otro aspecto, el método comprende además poner en contacto el drimenol con al menos una enzima para producir un derivado de drimenol.

En un aspecto adicional, el método comprende convertir el drimenol en un derivado de drimenol usando síntesis química o síntesis bioquímica.

65 También se proporciona en el presente documento un polipéptido aislado que tiene actividad sesquiterpeno sintasa que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o

99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO:2.

5 Además, en el presente documento se proporciona una molécula de un ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido que tiene actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO: 2.

10 Además se proporciona un ácido nucleico aislado con la secuencia de nucleótidos de la SEQ ID NO: 1 y de la SEQ ID NO: 3.

Además se proporciona una molécula de ácido nucleico aislada que codifica un polipéptido proporcionado en el presente documento.

15 En un aspecto, se proporciona en el presente documento un vector que comprende las moléculas de ácido nucleico descritas en el presente documento. En otro aspecto, el vector es un vector de expresión. En un aspecto adicional, el vector es un vector procariota, vector viral o un vector eucariota.

20 También se proporciona un organismo hospedador no humano o una célula hospedadora que comprende un gen recombinante, comprendiendo dicho gen: (1) una molécula de ácido nucleico descrita anteriormente, o (2) un vector de expresión que comprende dicha molécula de ácido nucleico. En un aspecto, el organismo no humano o la célula hospedadora es una célula procariota o eucariota. En otro aspecto, la célula hospedadora es una célula bacteriana, una célula vegetal, una célula fúngica o una levadura. En un aspecto adicional, la célula bacteriana es *E. coli* y la célula de levadura es *Saccharomyces cerevisiae*

25 Además se proporciona el uso de un polipéptido descrito en el presente documento para producir drimenol o una mezcla que comprende drimenol y uno o más terpenos.

30 En un aspecto, la mezcla producida en los métodos o usos anteriores comprende drimenol y nerolidol.

En un aspecto adicional, se aísla el drimenol y/o nerolidol.

Descripción de los dibujos

35 **Figura 1.** Estructura de (-)-drimenol.

Figura 2. Espectro de masas del (-)-drimenol auténtico.

Figura 3. Espectro de RMN 13C del (-)-drimenol auténtico.

Figura 4. Estructura de rayos X (radiación Cu K α) del (-)-drimenol auténtico.

40 **Figura 5.** Muestra el cromatograma GC/MS de extracto de diclorometano de raíz de *Paeonia anomala*. La flecha indica el pico de drimenol.

Figura 6. Muestra el cromatograma GC/MS del experimento de *PaTPS3* de la expresión de *E. coli* (solamente se muestra la zona de sesquiterpeno).

Figura 7. Muestra el espectro de masas del pico de drimenol en la figura 6.

45 **Figura 8.** Muestra el cromatograma GC/MS del ensayo *in vitro* de *PaTPS3* (solamente se muestra la zona de sesquiterpeno).

Figura 9. Muestra el espectro de masas del pico de drimenol en la figura 8.

Abreviaturas utilizadas

pb	par de bases
kb	kilo base
ADN	ácido desoxirribonucleico
ADNc	ADN complementario
DTT	ditiotreitól
FPP	farnesil difosfato
GC	cromatógrafo de gases
IPTG	isopropil-D-tiogalacto-piranósido
LB	caldo de lisogenia
MS	espectrómetro de masas/espectrometría de masas
MVA	ácido mevalónico
PCR	reacción en cadena de la polimerasa
ARN	ácido ribonucleico
ARNm	ácido ribonucleico mensajero
miRNA	microARN
ARNip	ARN de interferencia pequeño
ARNr	ARN ribosómico

ARNt ARN de transferencia

Definiciones

5 El término "polipéptido" significa una secuencia de aminoácidos de restos de aminoácidos polimerizados consecutivamente, por ejemplo, al menos 15 restos, al menos 30 restos, al menos 50 restos. En algunas realizaciones en el presente documento, un polipéptido comprende una secuencia de aminoácidos que es una enzima, o un fragmento, o una variante de la misma.

10 El término "proteína" se refiere a una secuencia de aminoácidos de cualquier longitud en donde los aminoácidos están unidos por enlaces peptídicos covalentes e incluye oligopéptidos, péptidos, polipéptidos y proteínas de longitud completa ya sea de origen natural o sintético.

15 El término polipéptido "aislado" se refiere a una secuencia de aminoácidos que se extrae de su entorno natural mediante cualquier método o combinación de métodos conocidos en la técnica e incluye métodos recombinantes, bioquímicos y sintéticos.

20 Las expresiones "sesquiterpeno sintasa" o "polipéptido que tiene actividad sesquiterpeno sintasa" se refieren a un polipéptido capaz de catalizar la síntesis de un sesquiterpeno o una mezcla que comprende uno o más sesquiterpenos, por ejemplo, drimenol y/o nerolidol, a partir de un pirofosfato de terpeno acíclico, particularmente farnesil difosfato (FPP).

25 Las expresiones "drimenol sintasa" o "polipéptido que tiene actividad drimenol sintasa" o "proteína drimenol sintasa" se refieren a un polipéptido capaz de catalizar la síntesis de drimenol, en forma de cualquiera de sus estereoisómeros o de una mezcla de los mismos, a partir de un pirofosfato de terpeno acíclico, particularmente farnesil difosfato (FPP). El drimenol puede ser el único producto o puede ser parte de una mezcla de sesquiterpenos.

30 Las expresiones "función biológica", "función", "actividad biológica" o "actividad" se refieren a la capacidad de la drimenol sintasa para catalizar la formación de drimenol o una mezcla de compuestos que comprende drimenol y uno o más terpenos.

Las expresiones "mezcla de terpenos que comprende drimenol" o "mezcla de sesquiterpenos que comprende drimenol" se refieren a una mezcla de terpenos o sesquiterpenos que comprende drimenol y uno o más terpenos o sesquiterpenos adicionales.

35 Las expresiones "secuencia de un ácido nucleico", "ácido nucleico", "molécula de ácido nucleico" y "polinucleótido" se usan indistintamente para indicar una secuencia de nucleótidos. Una secuencia de un ácido nucleico puede ser un desoxirribonucleótido monocatenario o bicatenario o un ribonucleótido de cualquier longitud, e incluir secuencias codificantes y no codificantes de un gen, exones, intrones, secuencias complementarias en sentido y antisentido, ADN genómico, ADNc, miARN, ARNip, ARNm, ARNr, ARNt, secuencias de ácidos nucleicos recombinantes, secuencias de ADN y/o ARN naturales aisladas y purificadas, secuencias de ADN y ARN sintéticas, fragmentos, cebadores y sondas de ácidos nucleicos. El experto en la materia es consciente de que las secuencias de ácido nucleico del ARN son idénticas a las secuencias de ADN con la diferencia de que la timina (T) se sustituye por uracilo (U). La expresión "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 componente de un ácido nucleico más grande.

45 Un "ácido nucleico aislado" o una "secuencia de un ácido nucleico aislada" se refiere a un ácido nucleico o una secuencia de un ácido nucleico que se encuentra en un entorno diferente de aquel en el que el ácido nucleico o la secuencia del ácido nucleico se encuentran de forma natural y puede incluir aquellos que están sustancialmente exentos de material endógeno contaminante. La expresión "de origen natural" como se usa en el presente documento y se aplica a un ácido nucleico se refiere a un ácido nucleico que se encuentra en una célula de un organismo en la naturaleza y que no ha sido modificado intencionalmente por un ser humano en el laboratorio.

50 Las "secuencias de ácidos nucleicos recombinantes" son secuencias de ácidos nucleicos que son el resultado del uso de métodos de laboratorio (por ejemplo, clonación molecular) para reunir material genético de más de una fuente, crear o modificar una secuencia de un ácido nucleico que no se produce de forma natural y que de otro modo no se encontraría en organismos biológicos.

60 "Tecnología de ADN recombinante" se refiere a procedimientos de biología molecular para preparar una secuencia de un ácido nucleico recombinante como se describe, por ejemplo, en *Laboratory Manuals* editado por Weigel y Glazebrook, 2002, Cold Spring Harbor Lab Press; y Sambrook *et al.*, 1989, Cold Spring Harbor, Nueva York, Cold Spring Harbor Laboratory Press.

65 El término "gen" significa una secuencia de ADN que comprende una región, que se transcribe en una molécula de ARN, por ejemplo, un ARNm en una célula, unido operativamente a regiones reguladoras adecuadas, por ejemplo, un promotor. Por lo tanto, un gen puede comprender varias secuencias unidas operativamente, tales como un promotor,

una secuencia líder en 5' que comprende, por ejemplo, secuencias implicadas en el inicio de la traducción, una región codificante de ADNc o ADN genómico, intrones, exones, y/o una secuencia no traducida en 3' que comprende, por ejemplo, sitios de terminación de la transcripción.

5 Un "gen quimérico" se refiere a cualquier gen que normalmente no se encuentra en la naturaleza en una especie, en particular, un gen en el que están presentes una o más partes de la secuencia de un ácido nucleico que no están asociadas entre sí en la naturaleza. Por ejemplo, el promotor no está asociado en la naturaleza con parte o la totalidad de la región transcrita o con otra región reguladora. Se entiende que la expresión "gen quimérico" incluye construcciones de expresión en las que un promotor o una secuencia reguladora de la transcripción se une operativamente a una o más secuencias codificantes o a una secuencia antisentido, es decir, complemento inverso de la hebra en sentido, o secuencia repetida invertida (en sentido y antisentido, por lo que el transcrito de ARN forma ARN bicatenario tras la transcripción). La expresión "gen quimérico" también incluye genes obtenidos a través de la combinación de porciones de una o más secuencias codificantes para producir un nuevo gen.

15 Una "3' UTR" o "secuencia no traducida en 3'" (también conocida como "región no traducida en 3'", o "extremo 3'") se refiere a la secuencia del ácido nucleico que se encuentra cadena abajo de la secuencia codificante de un gen, que comprende, por ejemplo, un sitio de terminación de la transcripción y (en la mayoría, pero no en todos los ARNm de eucariotas) una señal de poliadenilación tal como AAUAAA o variantes de la misma. Después de la terminación de la transcripción, el transcrito de ARNm se puede escindir cadena abajo de la señal de poliadenilación y se puede añadir una cola de poli(A), que está implicada en el transporte del ARNm al sitio de traducción, por ejemplo, el citoplasma.

20 La "expresión de un gen" abarca la "expresión heteróloga" y la "sobreexpresión" e implica la transcripción del gen y la traducción del ARNm en una proteína. La sobreexpresión se refiere a la producción del producto génico medido por los niveles de ARNm, la actividad polipeptídica y/o enzimática en células o en organismos transgénicos que excede los niveles de producción en células o en organismos sin transformar de un acervo genético similar.

30 "Vector de expresión", como se usa en el presente documento, significa una molécula de ácido nucleico modificada por ingeniería utilizando métodos de biología molecular y tecnología de ADN recombinante para el suministro de ADN extraño o exógeno a una célula hospedadora. El vector de expresión normalmente incluye secuencias necesarias para la transcripción adecuada de la secuencia de nucleótidos. La región codificante suele codificar una proteína de interés, pero también puede codificar un ARN, por ejemplo, un ARN antisentido, ARNip y similares.

35 Un "vector de expresión" como se usa en el presente documento incluye cualquier vector recombinante lineal o circular que incluye, entre otros, vectores virales, bacteriófagos y plásmidos. El experto en la materia es capaz de seleccionar un vector adecuado según el sistema de expresión. En una realización, el vector de expresión incluye el ácido nucleico de una realización en el presente documento unido operativamente a al menos una secuencia reguladora, que controla la transcripción, traducción, iniciación y terminación, tal como un promotor transcripcional, operador o potenciador, o un sitio de unión a ribosomas de ARNm y, opcionalmente, que incluye al menos un marcador de selección. Las secuencias de nucleótidos están "unidas operativamente" cuando la secuencia reguladora se relaciona funcionalmente con el ácido nucleico de una realización en el presente documento.

45 "Secuencia reguladora" se refiere a una secuencia de un ácido nucleico que determina el nivel de expresión de las secuencias de ácido nucleico de una realización en el presente documento y es capaz de regular la velocidad de transcripción de la secuencia de ácido nucleico unida operativamente a la secuencia reguladora. Las secuencias reguladoras comprenden promotores, potenciadores, factores de transcripción, elementos promotores y similares.

50 "Promotor" se refiere a una secuencia de un ácido nucleico que controla la expresión de una secuencia codificante al proporcionar un sitio de unión para la ARN polimerasa y otros factores necesarios para una transcripción adecuada, incluidos, entre otros, sitios de unión de factores de transcripción, sitios de unión a proteínas activadoras y represoras. El significado del término promotor también incluye la expresión "secuencia reguladora de promotor". Las secuencias reguladoras de promotor pueden incluir elementos cadena arriba y cadena abajo que pueden influir en la transcripción, en el procesamiento de ARN o en la estabilidad de la secuencia codificante de un ácido nucleico asociada. Los promotores incluyen secuencias de origen natural y sintéticas. Las secuencias codificantes de ácidos nucleicos normalmente se ubican cadena abajo del promotor con respecto a la dirección de la transcripción que comienza en el sitio de inicio de la transcripción.

La expresión "promotor constitutivo" se refiere a un promotor no regulado que permite la transcripción continua de la secuencia de un ácido nucleico a la que está unido operativamente.

60 Como se usa en el presente documento, la expresión "unido/a operativamente" se refiere a un enlace de elementos polinucleotídicos en una relación funcional. Un ácido nucleico está "unido operativamente" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor, o más bien una secuencia reguladora de la transcripción, está unido operativamente a una secuencia codificante si este afecta a la transcripción de la secuencia codificante. Unido/a operativamente significa que las secuencias de ADN que están unidas normalmente son contiguas. La secuencia de nucleótidos asociada a la secuencia promotora puede ser de origen homólogo o heterólogo con respecto a la planta a transformar. La secuencia también puede ser total o parcialmente

5 sintética. Independientemente del origen, la secuencia de un ácido nucleico asociada con la secuencia promotora se expresará o silenciará en función de las propiedades promotoras a las que se une después de unirse al polipéptido de una realización en el presente documento. El ácido nucleico asociado puede codificar una proteína que se desea expresar o suprimir en todo el organismo en todo momento o, como alternativa, en un momento específico o en tejidos, células o compartimiento de células específicos. Dichas secuencias de nucleótidos codifican particularmente proteínas que confieren rasgos fenotípicos deseables a las células hospedadoras o al organismo alterado o transformado con las mismas. Más particularmente, la secuencia de nucleótidos asociada da lugar a la producción de drimenol o una mezcla que comprende drimenol y uno o más terpenos en la célula o el organismo. Particularmente, la secuencia de nucleótidos codifica un polipéptido que tiene actividad drimenol sintasa.

10 "Péptido diana" se refiere a una secuencia de aminoácidos que dirige una proteína o polipéptido a orgánulos intracelulares, es decir, mitocondrias o plástidos o al espacio extracelular (péptido señal de secreción). Una secuencia de un ácido nucleico que codifica un péptido diana puede fusionarse con la secuencia de un ácido nucleico que codifica el extremo aminoterminal, por ejemplo, el extremo aminoterminal, de la proteína o polipéptido, o puede usarse para reemplazar un polipéptido de direccionamiento natural.

15 El término "cebador" se refiere a una secuencia de un ácido nucleico corta que hibrida con una secuencia de un ácido nucleico molde y se usa para la polimerización de una secuencia de un ácido nucleico complementaria al molde.

20 Como se usa en el presente documento, la expresión "célula hospedadora" o "célula transformada" se refiere a una célula (o un organismo) alterada para albergar al menos una molécula de ácido nucleico, por ejemplo, un gen recombinante que codifica una secuencia de un ácido nucleico o proteína deseados que tras la transcripción produce una proteína drimenol sintasa útil para producir drimenol o una mezcla que comprende drimenol y uno o más terpenos. La célula hospedadora es en particular, una célula bacteriana, una célula fúngica o una célula vegetal. La célula hospedadora puede contener un gen recombinante que se ha integrado en los genomas del orgánulo o del núcleo de la célula hospedadora. Como alternativa, el hospedador puede contener el gen recombinante extracromosómicamente.

25 Las secuencias homólogas incluyen secuencias ortólogas o parálogas. En la técnica se conoce métodos para identificar ortólogos o parálogos, incluidos métodos filogenéticos, métodos de similitud de secuencias e hibridación y se describen en el presente documento.

30 Los parálogos son el resultado de la duplicación de genes que da lugar a dos o más genes con secuencias similares y funciones similares. Los parálogos generalmente se agrupan y se forman por duplicaciones de genes en especies de plantas relacionadas. Los parálogos se encuentran en grupos de genes similares usando análisis Blast por pares o durante el análisis filogenético de familias de genes usando programas tal como CLUSTAL. En los parálogos, las secuencias consenso pueden identificarse como características de secuencias dentro de genes relacionados y que tienen funciones similares de los genes.

35 Los ortólogos o secuencias ortólogas, son secuencias similares entre sí porque se encuentran en especies que descienden de un ancestro común. Por ejemplo, se sabe que las especies de plantas que tienen ancestros comunes contienen muchas enzimas que tienen secuencias y funciones similares. El experto en la materia puede identificar secuencias ortólogas y predecir las funciones de los ortólogos, por ejemplo, construyendo un árbol poligénico para una familia de genes de una especie utilizando los programas CLUSTAL o BLAST. Un método para identificar o confirmar funciones similares entre secuencias homólogas es comparar los perfiles de transcripción en células o en organismos hospedadores, tales como plantas, que sobreexpresan o carecen (supresión génica/atenuación génica) de polipéptidos relacionados. El experto en la materia comprenderá que los genes que tienen perfiles de transcripción similares, con más del 50 % de transcripciones reguladas en común, o con más del 70 % de transcripciones reguladas en común, o más del 90 % de transcripciones reguladas en común tendrán funciones similares. Se espera que los homólogos, parálogos, ortólogos y cualquier otra variante de las secuencias en el presente documento funcionen de manera similar haciendo a las células hospedadoras, organismos tales como las plantas que producen proteínas de drimenol sintasa.

40 La expresión "marcador seleccionable" se refiere a cualquier gen que, tras su expresión, pueda utilizarse para seleccionar una célula o células que incluyen el marcador seleccionable. A continuación se describen ejemplos de marcadores seleccionables. El experto sabrá que se aplican diferentes marcadores seleccionables antibióticos, fungicidas, auxotróficos o herbicidas a diferentes especies diana.

45 "Drimenol" para los fines de la presente solicitud se refiere a (-)-drimenol (CAS: 468-68-8) (véanse también las figuras 1-4).

50 El término "organismo" se refiere a cualquier organismo multicelular o unicelular no humano tal como una planta o un microorganismo. Particularmente, un microorganismo es una bacteria, una levadura, un alga o un hongo.

55 El término "planta" se usa indistintamente para incluir células vegetales, incluidos protoplastos vegetales, tejidos vegetales, cultivos de tejidos de células vegetales que dan lugar a plantas regeneradas, o partes de plantas, u órganos

de plantas tales como raíces, tallos, hojas, flores, polen, óvulos, embriones, frutos y similares. Se puede usar cualquier planta para realizar los métodos de una realización en el presente documento.

5 Se entiende que un organismo o célula en particular es "capaz de producir FPP" cuando produce FPP de forma natural o cuando no produce FPP de forma natural pero se transforma para producir FPP, ya sea antes de la transformación con un ácido nucleico como se describe en el presente documento o junto con dicho ácido nucleico. Los organismos o células transformados para producir una mayor cantidad de FPP que el organismo o célula de origen natural también están incluidos en los "organismos o células capaces de producir FPP".

10 Para las descripciones en el presente documento y las reivindicaciones adjuntas, el uso de "o" significa "y/o" a menos que se indique otra cosa. De forma similar, "comprenden", "comprende", "que comprende", "incluyen", "incluye", y "que incluye" son intercambiables y no pretenden ser limitantes.

15 Debe entenderse además que cuando las descripciones de varias realizaciones usan el término "que comprende", los expertos en la materia entenderían que en algunos casos específicos, una realización puede describirse, como alternativa, usando la expresión "que consiste esencialmente en" o "que consiste en".

Descripción detallada

20 En el presente documento se proporciona una molécula de un ácido nucleico que comprende la secuencia de nucleótidos de la SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento inverso de las mismas.

25 Según una realización, la molécula de ácido nucleico consiste en una secuencia de nucleótidos de la SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento inverso de las mismas.

En una realización, el ácido nucleico de una realización en el presente documento puede estar presente de forma natural en plantas de *Paeonia* o en otras especies de plantas, u obtenerse modificando la SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento inverso de las mismas.

30 En otra realización, el ácido nucleico se aísla o procede de una planta de la familia Paeoniaceae. En una realización adicional, el ácido nucleico se aísla o procede de *Paeonia anomala*.

35 También se proporciona, pero no forma parte de la invención, una secuencia de nucleótidos obtenida modificando la SEQ ID NO: 1 o la SEQ ID NO: 3 o el complemento inverso de las mismas que abarca cualquier secuencia que se haya obtenido modificando la secuencia de la SEQ ID NO: 1 o SEQ ID NO: 3, o del complemento inverso de las mismas utilizando cualquier método conocido en la técnica, por ejemplo, introduciendo cualquier tipo de mutaciones tal como mutaciones por eliminación, inserción y/o sustitución. Los ácidos nucleicos que comprenden una secuencia obtenida por mutación de la SEQ ID NO: 1 o SEQ ID NO: 3 o del complemento inverso de las mismas están abarcados por una realización en el presente documento, siempre que las secuencias que comprenden compartan al menos la identidad de secuencia definida de la SEQ ID NO: 1 o SEQ ID NO: 3 o del complemento inverso de las mismas y siempre que codifiquen un polipéptido que tenga actividad drimenol sintasa, como se define en cualquiera de las realizaciones anteriores. Las mutaciones pueden ser cualquier tipo de mutaciones de estos ácidos nucleicos, por ejemplo, mutaciones puntuales, mutaciones por eliminación, mutaciones por inserción y/o mutaciones por cambio de marco de uno o más nucleótidos de la secuencia de ADN de la SEQ ID NO: 1 o SEQ ID NO: 3. En una realización, el ácido nucleico de una realización en el presente documento se puede truncar siempre que codifique un polipéptido como se describe en el presente documento.

50 Puede prepararse un ácido nucleico variante para adaptar su secuencia de nucleótidos a un sistema de expresión específico. Por ejemplo, se sabe que los sistemas de expresión bacterianos expresan polipéptidos de manera más eficiente si los aminoácidos están codificados por codones particulares.

55 Debido a la degeneración del código genético, más de un codón puede codificar la misma secuencia de aminoácidos, múltiples secuencias de ácido nucleico pueden codificar la misma proteína o polipéptido, estando todas estas secuencias de ADN abarcadas por una realización en el presente documento. Cuando sea apropiado, las secuencias de un ácido nucleico que codifican la drimenol sintasa pueden optimizarse para aumentar la expresión en la célula hospedadora. Por ejemplo, los nucleótidos de una realización en el presente documento se pueden sintetizar utilizando codones específicos de un hospedador para mejorar la expresión.

60 En una realización se proporciona en el presente documento una secuencia de un ácido nucleico aislada, recombinante o sintética de SEQ ID NO: 1 o SEQ ID NO: 3 que codifica un polipéptido que tiene actividad drimenol sintasa que comprende la secuencia de aminoácidos de la SEQ ID NO: 2 o fragmentos de la misma que catalizan la producción de drimenol o una mezcla que comprende drimenol y uno o más terpenos en una célula de un precursor de FPP.

65 En el presente documento también se proporcionan secuencias de ADNc, de ADN genómico y de ARN. Cualquier secuencia de un ácido nucleico que codifique la drimenol sintasa o variantes de la misma también se denomina en el

presente documento secuencia codificante de drimenol sintasa.

Según una realización, el ácido nucleico de la SEQ ID NO: 1 o SEQ ID NO: 3 es la secuencia codificante de un gen de drimenol sintasa que codifica una drimenol sintasa obtenida como se describe en los ejemplos.

5 Un fragmento de un polinucleótido de la SEQ ID NO: 1 o SEQ ID NO: 3 se refiere a nucleótidos contiguos que es particularmente de al menos 15 pb, al menos 30 pb, al menos 40 pb, al menos 50 pb y/o al menos 60 pb de longitud del polinucleótido de una realización en el presente documento. En particular, el fragmento de un polinucleótido comprende al menos 25, más particularmente al menos 50, más particularmente al menos 75, más particularmente al menos 100, más particularmente al menos 150, más particularmente al menos 200, más particularmente al menos 300, más particularmente al menos 400, más particularmente al menos 500, más particularmente al menos 600, más particularmente al menos 700, más particularmente al menos 800, más particularmente al menos 900, más particularmente al menos 1000 nucleótidos contiguos del polinucleótido de una realización en el presente documento. Sin intención de limitarse por ninguna teoría, el fragmento de los polinucleótidos en el presente documento puede usarse como cebador de PCR y/o como sonda, o para silenciar genes antisentido o ARNi.

20 Está claro para el experto en la materia que los genes, incluyendo los polinucleótidos de una realización en el presente documento, pueden clonarse en función de la información de secuencia de nucleótidos disponible, tal como se encuentra en el listado de secuencias adjunto, según métodos conocidos en la técnica. Estos incluyen, por ejemplo, el diseño de cebadores de ADN que representan las secuencias flanqueantes de dicho gen, de los cuales uno se genera en orientaciones en sentido y que inicia la síntesis de la cadena en sentido y el otro se crea de manera complementaria inversa y genera la cadena antisentido. Las polimerasas de ADN termoestables, como las que se usan en la reacción en cadena de la polimerasa, se usan comúnmente para realizar tales experimentos. Como alternativa, las secuencias de ADN que representan genes pueden sintetizarse químicamente y posteriormente introducirse en moléculas de vector de ADN que pueden multiplicarse, por ejemplo, mediante bacterias compatibles como, por ejemplo, *E. coli*.

30 En una realización relacionada proporcionada en el presente documento, se proporcionan cebadores y/o sondas de PCR para detectar secuencias de ácido nucleico que codifican una drimenol sintasa, pero no forman parte de la invención. El experto en la materia conocerá los métodos para sintetizar pares de cebadores de PCR degenerados o específicos para amplificar una secuencia de ácido nucleico que codifica la drimenol sintasa o fragmentos de la misma, en función de la SEQ ID NO: 1 o SEQ ID NO: 3. Un kit de detección de secuencias de ácidos nucleicos que codifican la drimenol sintasa puede incluir cebadores y/o sondas específicos para las secuencias de los ácidos nucleicos que codifican la drimenol sintasa, y un protocolo asociado para el uso de los cebadores y/o las sondas para detectar secuencias de ácidos nucleicos que codifican la drimenol sintasa en una muestra. Dichos kits de detección pueden utilizarse para determinar si una planta, organismo o célula se ha modificado, es decir, transformado con una secuencia que codifica la drimenol sintasa.

40 Para probar una función de secuencias de ADN variantes según una realización en el presente documento, la secuencia de interés se une operativamente a un gen marcador seleccionable o explorable y la expresión del gen indicador se analiza en ensayos de expresión transitoria con protoplastos o plantas transformadas de forma estable. El experto en la materia reconocerá que las secuencias de ADN capaces de impulsar la expresión se construyen como módulos. En consecuencia, los niveles de expresión de los fragmentos de ADN más cortos pueden ser diferentes a los del fragmento más largo y pueden ser diferentes entre sí. En el presente documento también se proporcionan equivalentes funcionales de la secuencia de un ácido nucleico que codifica las proteínas de drimenol sintasa proporcionadas en el presente documento, es decir, secuencias de nucleótidos que hibridan en condiciones rigurosas con la secuencia del ácido nucleico de la SEQ ID NO: 1 o SEQ ID NO: 3.

50 El experto en la materia conocerá métodos para identificar secuencias homólogas en otros organismos y métodos para determinar el porcentaje de identidad de secuencia entre secuencias homólogas. A continuación, dichas moléculas de ADN recién identificadas pueden secuenciarse y la secuencia puede compararse con la secuencia del ácido nucleico de la SEQ ID NO: 1 o SEQ ID NO: 3.

55 El porcentaje de identidad entre dos secuencias de péptidos o nucleótidos es función del número de aminoácidos o restos de nucleótidos que son idénticos en las dos secuencias cuando se ha generado una alineación de estas dos secuencias. Los restos idénticos se definen como restos que son iguales en las dos secuencias en una posición dada de la alineación. El porcentaje de identidad de secuencia, como se usa en el presente documento, se calcula a partir de la alineación óptima tomando el número de restos idénticos entre dos secuencias dividiéndolo por el número total de restos en la secuencia más corta y multiplicándolo por 100. La alineación óptima es aquella en la que el porcentaje de identidad es el más alto posible. Pueden introducirse huecos en una o ambas secuencias en una o más posiciones de la alineación para obtener la alineación óptima. A continuación, estos huecos se tienen en cuenta como restos no idénticos para el cálculo del porcentaje de identidad de secuencia. La alineación con el fin de determinar el porcentaje de identidad de la secuencia de aminoácidos o ácidos nucleicos se puede lograr de varias maneras utilizando programas informáticos y, por ejemplo, programas informáticos disponibles públicamente disponibles en internet. Preferentemente, el programa BLAST (Tatiana et al, FEMS Microbiol Lett., 1999, 174:247-250, 1999) ajustado a los parámetros predeterminados, disponible en el sitio web del National Center for Biotechnology Information (NCBI) en

ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi, se puede utilizar para obtener una alineación óptima de secuencias de proteínas o ácidos nucleicos y para calcular el porcentaje de identidad de secuencia.

5 Una realización relacionada proporcionada en el presente documento proporciona una secuencia de un ácido nucleico que es complementaria a la secuencia de ácido nucleico según la SEQ ID NO: 1 o SEQ ID NO: 3, tal como ARN inhibidores, o una secuencia de ácido nucleico que hibrida en condiciones estrictas con al menos una parte de la secuencia de nucleótidos según la SEQ ID NO: 1 o SEQ ID NO: 3. Una realización alternativa de una realización en el presente documento proporciona un método para alterar la expresión génica en una célula hospedadora. Por ejemplo, el polinucleótido de una realización en el presente documento puede potenciarse o sobreexpresarse o inducirse en determinados contextos (por ejemplo, tras la exposición a una determinada temperatura o condiciones de cultivo) en una célula hospedadora o un organismo hospedador.

15 La alteración de la expresión de un polinucleótido proporcionado en el presente documento también puede dar como resultado una expresión ectópica que es un patrón de expresión diferente en un organismo alterado y en un control o de tipo silvestre. La alteración de la expresión se produce a partir de interacciones del polipéptido de una realización en el presente documento con moduladores exógenos o endógenos, o como resultado de la modificación química del polipéptido. El término también se refiere a un patrón de expresión del polinucleótido alterado de una realización en el presente documento que está alterado por debajo del nivel de detección o tiene su actividad completamente suprimida.

20 En una realización, también se proporciona en el presente documento un polinucleótido recombinante o sintético que codifica un polipéptido o polipéptido variante proporcionado en el presente documento.

25 En una realización, se proporciona una molécula de ácido nucleico aislada que codifica un polipéptido que tiene actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO: 2.

30 En una realización en el presente documento, se proporciona un polipéptido aislado que tiene actividad sesquiterpeno sintasa y/o actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO: 2.

Según una realización, el polipéptido consiste en la secuencia de aminoácidos de la SEQ ID NO: 2.

35 En una realización, el polipéptido de una realización en el presente documento puede estar presente de forma natural en plantas de *Paeonia* o en otras especies de plantas, o comprende una secuencia de aminoácidos que es una variante de la SEQ ID NO: 2, obtenida por ingeniería genética o que se encuentra de forma natural en plantas de *Paeonia* o en otras especies de plantas.

40 Según otra realización, el polipéptido se aísla o procede de una planta de la familia Paeoniaceae. En una realización adicional, el polipéptido se aísla o procede de *Paeonia anomala*.

45 En una realización, el al menos un polipéptido que tiene actividad sesquiterpeno sintasa y/o actividad drimenol sintasa usado en cualquiera de las realizaciones descritas en el presente o codificado por el ácido nucleico usado en cualquiera de las realizaciones descritas en el presente documento comprende una secuencia de aminoácidos que es una variante de la SEQ ID NO: 2, obtenida por ingeniería genética. En una realización, el polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando la SEQ ID NO: 1 o la SEQ ID NO: 3 o el complemento inverso de las mismas.

50 También se pretende que los polipéptidos incluyan polipéptidos variantes y truncados siempre que tengan actividad sesquiterpeno sintasa y/o actividad drimenol sintasa.

55 Según otra realización, el al menos un polipéptido que tiene actividad drimenol sintasa usado en cualquiera de las realizaciones descritas en el presente documento o codificado por el ácido nucleico usado en cualquiera de las realizaciones descritas en el presente documento comprende una secuencia de aminoácidos que es una variante de la SEQ ID NO: 2, obtenida por ingeniería genética, siempre que dicha variante tenga actividad drimenol sintasa y tenga el porcentaje de identidad con la SEQ ID NO: 2 necesario como se describe en el presente documento.

60 Según otra realización, el al menos un polipéptido que tiene actividad drimenol sintasa usado en cualquiera de las realizaciones descritas en el presente documento o codificado por el ácido nucleico usado en cualquiera de las realizaciones descritas en el presente documento es una variante de la SEQ ID NO: 2 que se puede encontrar naturalmente en otros organismos, tal como otras especies de plantas, siempre que tenga una actividad drimenol sintasa. Como se usa en el presente documento, el polipéptido incluye un polipéptido o fragmento de péptido que abarca las secuencias de aminoácidos identificadas en el presente documento, así como polipéptidos truncados o variantes siempre que tengan actividad sesquiterpeno sintasa y/o actividad drimenol sintasa y que compartan al menos el porcentaje de identidad con el fragmento correspondiente de la SEQ ID NO: 2 definido.

- 5 Los ejemplos de polipéptidos variantes son proteínas de origen natural que son el resultado de eventos alternos de corte y empalme de ARNm o de la escisión proteolítica de los polipéptidos descritos en el presente documento. Las variaciones atribuibles a la proteólisis incluyen, por ejemplo, diferencias en los extremos N o C al expresarse en diferentes tipos de células hospedadoras, debido a la eliminación proteolítica de uno o más aminoácidos terminales de los polipéptidos de una realización en el presente documento. También están abarcados por una realización en el presente documento polipéptidos codificados por un ácido nucleico obtenido por mutación natural o artificial de un ácido nucleico de una realización en el presente documento, como se describe a continuación.
- 10 Las variantes polipeptídicas resultantes de una fusión de secuencias peptídicas adicionales en los extremos amino y carboxilo terminales también se pueden usar en los métodos de una realización en el presente documento. En particular, tal fusión puede potenciar la expresión de los polipéptidos, siendo útil en la purificación de la proteína o mejorando la actividad enzimática del polipéptido en un entorno o sistema de expresión deseado. Dichas secuencias peptídicas adicionales pueden ser péptidos señal, por ejemplo. Otro aspecto abarca métodos que utilizan polipéptidos variantes, como los obtenidos por fusión con otros oligo o polipéptidos y/o los que se unen a péptidos señal. También se pueden utilizar de forma ventajosa en los métodos de una realización en el presente documento polipéptidos resultantes de una fusión con otra proteína funcional, tal como otra proteína de la ruta de biosíntesis de terpenos.
- 15
- 20 Una variante también puede diferir del polipéptido de una realización en el presente documento por la unión de grupos modificadores que están unidos de forma covalente o no covalente al esqueleto del polipéptido. La variante también incluye un polipéptido que difiere del polipéptido proporcionado en el presente documento por la introducción de sitios de glicosilación ligados a N o ligados a O, y/o una adición de restos de cisteína. El experto en la materia reconocerá cómo modificar una secuencia de aminoácidos y conservar su actividad biológica.
- 25
- 30 Además de las secuencias de genes que se muestran en las secuencias divulgadas en el presente documento, será evidente para el experto en la materia que pueden existir polimorfismos en la secuencia de ADN dentro de una población dada, lo que puede dar lugar a cambios en la secuencia de aminoácidos de los polipéptidos divulgados en el presente documento. Dichos polimorfismos genéticos pueden existir en células de diferentes poblaciones o dentro de una población debido a la variación alélica natural. Las variantes alélicas también pueden incluir equivalentes funcionales.
- 35
- Otras realizaciones también se refieren a las moléculas derivadas por dichos polimorfismos de secuencia a partir de los ácidos nucleicos divulgados concretamente. Estas variaciones naturales generalmente provocan una variación de aproximadamente 1 a 5% en la secuencia de nucleótidos de un gen o en la secuencia de aminoácidos de los polipéptidos divulgados en el presente documento. Como se ha mencionado anteriormente, el ácido nucleico que codifica el polipéptido o variantes del mismo de una realización en el presente documento es una herramienta útil para modificar células u organismos hospedadores no humanos y para modificar células u organismos hospedadores no humanos destinados a ser utilizados en los métodos descritos en el presente documento.
- 40
- Una realización proporcionada en el presente documento proporciona secuencias de aminoácidos de proteínas de drimenol sintasa que incluyen ortólogos y parálogos, así como métodos para identificar y aislar ortólogos y parálogos de las drimenol sintasas en otros organismos. Particularmente, los ortólogos y parálogos de la drimenol sintasa así identificados conservan la actividad drimenol sintasa y son capaces de producir drimenol o una mezcla que comprende drimenol y uno o más terpenos a partir de un precursor de pirofosfato de terpeno acíclico, por ejemplo, FPP.
- 45
- El polipéptido que se va a poner en contacto con un pirofosfato de terpeno acíclico, por ejemplo, FPP, puede obtenerse *in vitro* por extracción de cualquier organismo que lo exprese, utilizando tecnologías convencionales de extracción de proteínas o enzimas. Si el organismo hospedador es un organismo unicelular o una célula que libera el polipéptido de una realización en el presente documento en el medio de cultivo, el polipéptido puede recogerse simplemente del medio de cultivo, por ejemplo por centrifugación, seguido opcionalmente por pasos de lavado y resuspensión en soluciones tampón adecuadas. Si el organismo o la célula acumula el polipéptido dentro de sus células, el polipéptido se puede obtener por ruptura o lisis de las células y, opcionalmente, extracción adicional del polipéptido del lisado celular. Las células intactas, el lisado celular o el polipéptido extraído se pueden usar para ponerlos en contacto con el pirofosfato de terpeno acíclico para la producción de un terpeno o una mezcla de terpenos.
- 50
- 55
- El polipéptido que tiene una actividad drimenol sintasa, ya sea en forma aislada o junto con otras proteínas, por ejemplo, en un extracto de proteína en bruto obtenido de células cultivadas o microorganismos, después, puede suspenderse en una solución tampón a un pH óptimo. Si es adecuado, pueden añadirse sales, DTT, cationes inorgánicos y otros tipos de cofactores enzimáticos, para optimizar la actividad enzimática. El FPP precursor se añade a la suspensión de polipéptidos, que luego se incuba a temperatura óptima, por ejemplo entre 15 y 40 °C, particularmente entre 25 y 35 °C, más particularmente a 30 °C. Después de la incubación, el drimenol producido puede aislarse de la solución incubada mediante procedimientos convencionales de aislamiento, tal como la extracción con disolventes y la destilación, opcionalmente después de la eliminación de los polipéptidos de la solución.
- 60
- 65 Según otra realización, el al menos un polipéptido que tiene actividad drimenol sintasa puede usarse para la producción de drimenol o mezclas de terpenos que comprenden drimenol. En otra realización, la mezcla que comprende drimenol

también puede comprender nerolidol.

Una herramienta particular para realizar el método de una realización en el presente documento es el propio polipéptido como se describe en el presente documento.

5 Según una realización particular, el polipéptido es capaz de producir una mezcla de sesquiterpenos que comprende drimenol. En una realización adicional, el polipéptido es capaz de producir una mezcla de sesquiterpenos que comprende drimenol, en donde el drimenol representa al menos el 20 %, particularmente al menos el 30 %, particularmente al menos el 35 %, particularmente al menos el 90 %, particularmente al menos el 95 %, más particularmente al menos el 98 % de los sesquiterpenos producidos. En otro aspecto proporcionado en este caso, el drimenol se produce con un porcentaje mayor o igual al 95 %, más particularmente al 98 % de selectividad.

15 Según otra realización, la sesquiterpeno sintasa es capaz de producir una mezcla de sesquiterpenos que comprende drimenol y nerolidol. En una realización adicional, la sintasa es capaz de producir una mezcla de sesquiterpenos que comprende drimenol y nerolidol, en donde nerolidol representa al menos del 5 % a aproximadamente el 80 %, en particular, al menos del 10 % a aproximadamente el 80 %, en particular, al menos del 15 % a aproximadamente el 80 %, en particular, al menos del 16 % a aproximadamente el 80 %, en particular, al menos del 50 % a aproximadamente el 80 %, en particular, al menos del 60 % a aproximadamente el 80 %, en particular, al menos del 70 % a aproximadamente el 80 %, particularmente aproximadamente el 79 % de los sesquiterpenos producidos.

20 La funcionalidad o actividad de cualquier proteína, variante o fragmento de sesquiterpeno sintasa o drimenol sintasa, puede determinarse utilizando diversos métodos. Por ejemplo, sobreexpresión transitoria o estable en células vegetales, bacterianas o de levadura para probar si la proteína tiene actividad, es decir, produce uno o más sesquiterpenos como drimenol o una mezcla de sesquiterpenos que comprende drimenol o que comprende drimenol y nerolidol a partir de producir un precursor de pirofosfato de terpeno acíclico, por ejemplo, precursor de FPP. La actividad drimenol sintasa puede evaluarse en un sistema de expresión microbiano, como el ensayo descrito en el ejemplo 2 en el presente documento sobre la producción de drimenol, indicando funcionalidad. Una variante o derivado de un polipéptido de drimenol sintasa de una realización en el presente documento conserva la capacidad de producir drimenol o una mezcla que comprende drimenol a partir de precursores de FPP. Las variantes de secuencia de aminoácidos de las drimenol sintasas proporcionadas en el presente documento pueden tener funciones biológicas deseables adicionales que incluyen, por ejemplo, utilización alterada del sustrato, cinética de reacción, distribución de productos u otras alteraciones.

35 La capacidad de un polipéptido para catalizar la síntesis de un sesquiterpeno en particular (por ejemplo, drimenol) puede confirmarse simplemente, por ejemplo, realizando el ensayo enzimático como se detalla en los ejemplos 1 y 2.

Además se proporciona al menos un vector que comprende las moléculas de ácido nucleico descritas en el presente documento.

40 También se proporciona en el presente documento un vector seleccionado del grupo de un vector procariota, vector viral y un vector eucariota.

Además, aquí se proporciona un vector que es un vector de expresión.

45 En una realización, varias drimenol sintasas que codifican secuencias de ácido nucleico se coexpresan en un solo hospedador, particularmente bajo el control de diferentes promotores. En otra realización, varias proteínas de drimenol sintasa que codifican secuencias de ácidos nucleicos pueden estar presentes en un solo vector de transformación o cotransformarse al mismo tiempo usando vectores separados y seleccionando transformantes que comprenden ambos genes quiméricos. De forma similar, uno o más genes que codifican la drimenol sintasa pueden expresarse en una sola planta, célula, organismo, o microorganismo junto con otros genes quiméricos.

50 Las secuencias de ácido nucleico de una realización en el presente documento que codifican proteínas de drimenol sintasa pueden insertarse en vectores de expresión y/o encontrarse en genes quiméricos insertados en vectores de expresión, para producir proteínas de drimenol sintasa en una célula hospedadora o en un organismo hospedador no humano. Los vectores para insertar transgenes en el genoma de las células hospedadoras son bien conocidos en la técnica e incluyen plásmidos, virus, cósmidos y cromosomas artificiales. Los vectores binarios o de cointegración en los que se inserta un gen quimérico también se pueden usar para transformar células hospedadoras.

60 Una realización proporcionada en el presente documento proporciona vectores de expresión recombinantes que comprenden una secuencia de ácido nucleico de un gen de drimenol sintasa, o un gen quimérico que comprende una secuencia de ácido nucleico de un gen de drimenol sintasa, unido operativamente a secuencias de ácido nucleico asociadas tales como, por ejemplo, secuencias promotoras. Por ejemplo, un gen quimérico que comprende una secuencia de ácido nucleico de SEQ ID NO: 1 o SEQ ID NO: 3 o una variante de las mismas puede unirse operativamente a una secuencia promotora adecuada para la expresión en células vegetales, células bacterianas o células fúngicas, opcionalmente unido a una secuencia de ácido nucleico no traducida en 3'.

65

Como alternativa, la secuencia promotora puede estar ya presente en un vector, de modo que la secuencia de ácido nucleico a transcribir se inserta en el vector cadena abajo de la secuencia promotora. Los vectores se pueden modificar por ingeniería genética para que tengan un origen de replicación, un sitio de clonación múltiple y un marcador seleccionable.

5 En una realización, un vector de expresión que comprende un ácido nucleico como se describe en el presente documento puede usarse como una herramienta para transformar organismos hospedadores no humanos o células hospedadoras adecuados para realizar el método de una realización en el presente documento *in vivo*.

10 Los vectores de expresión proporcionados en el presente documento pueden usarse en los métodos para preparar un organismo hospedador no humano y/o una célula hospedadora transformados genéticamente, en organismos hospedadores no humanos y/o células hospedadoras que albergan los ácidos nucleicos de una realización en el presente documento y en los métodos para producir polipéptidos que tienen actividad drimenol sintasa, como se describe en el presente documento.

15 Los organismos hospedadores no humanos y células hospedadoras recombinantes transformados para albergar al menos un ácido nucleico de una realización en el presente documento de manera que exprese heterológicamente o sobreexpresen al menos un polipéptido de una realización en el presente documento también son herramientas muy útiles para realizar el método de una realización en el presente documento. Por lo tanto, dichos organismos
20 hospedadores no humanos y células hospedadoras se proporcionan en el presente documento.

En una realización, se proporciona una célula hospedadora o un organismo hospedador no humano que comprende al menos una de las moléculas de ácido nucleico descritas en el presente documento o que comprende al menos un vector que comprende al menos una de las moléculas de ácido nucleico.

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

30 En una realización, el organismo hospedador no humano o la célula hospedadora es una célula procariota. En otra realización, el organismo hospedador no humano o la célula hospedadora es una célula bacteriana. En una realización adicional, el organismo hospedador no humano o la célula hospedadora es *Escherichia coli*.

35 En una realización, el organismo hospedador no humano o la célula hospedadora es una célula eucariota. En otra realización, el organismo hospedador no humano o la célula hospedadora es una célula de levadura. En una realización adicional, el organismo hospedador no humano o la célula es *Saccharomyces cerevisiae*.

En una realización adicional, el organismo no humano o célula hospedadora es una célula vegetal.

40 En una realización, el organismo hospedador no humano o la célula hospedadora expresa un polipéptido, siempre que el organismo o la célula se transformen para albergar un ácido nucleico que codifica dicho polipéptido, transcribiéndose este ácido nucleico a ARNm y encontrándose el polipéptido en el organismo o célula hospedadores.

45 Los métodos adecuados para transformar un organismo hospedador no humano o una célula hospedadora se han descrito previamente y también se proporcionan en el presente documento.

50 Para realizar una realización en el presente documento *in vivo*, el organismo hospedador o la célula hospedadora se cultiva en condiciones que dan lugar a la producción de sesquiterpenos tales como drimenol o una mezcla que comprende drimenol. En consecuencia, si el hospedador es una planta transgénica, se pueden proporcionar condiciones óptimas de crecimiento, tal como condiciones óptimas de luz, agua y nutrientes, por ejemplo. Si el hospedador es un organismo unicelular, las condiciones que dan lugar a la producción de drimenol o una mezcla que comprende drimenol 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 drimenol. Se describen ejemplos de condiciones de cultivo óptimas de manera más detallada en los ejemplos.

55 Organismos hospedadores no humanos adecuados para realizar el método de una realización en el presente documento *in vivo* puede ser cualquier organismo multicelular o unicelular no humano. En una realización, el organismo hospedador no humano utilizado para realizar una realización en el presente documento *in vivo* es una planta, un procariota o un hongo. Se puede utilizar cualquier planta, procariota u hongo. Las plantas particularmente
60 útiles son aquellas que producen de manera natural grandes cantidades de terpenos. En otra realización, el organismo hospedador no humano utilizado para realizar el método de una realización en el presente documento *in vivo* es un microorganismo. Se puede utilizar cualquier microorganismo, por ejemplo, el microorganismo puede ser una bacteria o una levadura, tal como *E. coli* o *Saccharomyces cerevisiae*.

65 Algunos de estos organismos no producen FPP de forma natural. Para ser adecuado para realizar el método de una realización en el presente documento, los organismos o células que no producen un precursor de pirofosfato de

terpeno acíclico, por ejemplo, FPP, de manera natural se transforman para producir dicho precursor. Pueden transformarse así antes de la modificación con el ácido nucleico descrito según cualquiera de las realizaciones anteriores o simultáneamente, como se ha explicado anteriormente. También se conocen en la técnica métodos para transformar organismos, por ejemplo microorganismos, para que produzcan un precursor de pirofosfato de terpeno acíclico, por ejemplo, FPP.

También se pueden utilizar células eucariotas superiores aisladas, en lugar de organismos completos, como hospedadores para realizar el método de una realización en el presente documento *in vivo*. Las células eucariotas adecuadas pueden ser cualquier célula no humana, tales como células vegetales o fúngicas.

Además, en el presente documento se proporciona un método para producir drimenol o una mezcla que comprende drimenol que comprende:

- i) poner en contacto farnesil difosfato (FPP) con un polipéptido que tiene actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO: 2; y
- ii) opcionalmente aislar el drimenol.

En un aspecto, se aísla el drimenol.

En otro aspecto proporcionado en este caso, el drimenol se produce con un porcentaje mayor o igual al 20 %, 30 %, 35 %, 40 %, 50 %, 60 %, 80 % o 90 % o incluso el 95 % de selectividad de los sesquiterpenos producidos.

Además, en el presente documento se proporciona un método que comprende transformar un organismo hospedador no humano o una célula hospedadora con un ácido nucleico que codifica un polipéptido que tiene actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO: 2.

En una realización, un método proporcionado en el presente documento comprende cultivar un organismo hospedador no humano o una célula hospedadora capaces de producir FPP y transformados para expresar un polipéptido en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad de secuencia con la SEQ ID NO: 2 en condiciones que permitan la producción del polipéptido.

En otra realización, un método proporcionado en el presente documento comprende poner en contacto un sesquiterpeno tal como drimenol con al menos una enzima para producir un derivado de sesquiterpeno. Los ejemplos de dichos derivados de drimenol incluyen, entre otros, acetato de drimenilo (CAS 40266-93-1), drimenal (CAS 105426-71-9), ácido driménico (CAS 111319-84-7).

Según otra realización particular, el método de cualquiera de las realizaciones descritas anteriormente se realiza *in vivo*. En tal caso, el paso a) comprende cultivar un organismo hospedador no humano o una célula hospedadora capaces de producir FPP y transformados para expresar al menos un polipéptido que comprende una secuencia de aminoácidos que comprende la SEQ ID NO: 2 o una variante funcional de la misma y que tiene un actividad drimenol sintasa, en condiciones que dan lugar a la producción de uno o más sesquiterpenos tales como drimenol o una mezcla que comprende drimenol. El drimenol puede ser el único producto o puede ser parte de una mezcla de sesquiterpenos. En una realización, la mezcla de sesquiterpenos comprende drimenol y nerolidol.

Según una realización adicional, el método comprende adicionalmente, antes de la etapa a), transformar un organismo no humano o célula capaces de producir FPP con al menos un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que comprende la SEQ ID NO: 2 o que codifica un polipéptido que tiene actividad sesquiterpeno sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o un 99 % de identidad de secuencia con la SEQ ID NO: 2, para que dicho organismo exprese dicho polipéptido.

Estas realizaciones de una realización en el presente documento son particularmente ventajosas ya que es posible realizar el método *in vivo* sin aislar previamente el polipéptido. La reacción se produce directamente dentro del organismo o célula transformados para expresar dicho polipéptido.

Una realización en el presente documento proporciona polipéptidos de una realización en el presente documento para utilizarlos en un método para producir drimenol o una mezcla que comprende drimenol poniendo en contacto un precursor de FPP con los polipéptidos de una realización en el presente documento *in vitro* o *in vivo*.

Además se proporciona el uso de un polipéptido como se describe en el presente documento para producir drimenol o una mezcla que comprende drimenol y uno o más terpenos o una mezcla que comprende drimenol y nerolidol. En

una realización, el drimenol y/o nerolidol producidos se aíslan.

Ejemplos

5 Ejemplo 1

Obtención de material vegetal de *Paeonia anomala* y secuenciación del transcriptoma de la raíz.

10 El material vegetal de *Paeonia anomala* se obtuvo de Datong en Qinghai, China. Para establecer si *Paeonia anomala* contenía drimenol, sus raíces se recogieron y extrajeron frescas con diclorometano para análisis químico. El extracto se analizó por GC-MS, estando descritos los parámetros del análisis GC-MS a continuación: Se utilizó un sistema GC Agilent serie 6890 equipado con una columna DB1-ms de 30 m · 0,25 mm · 0,25 µm de espesor de película, PIN 122-0132 (J&W Scientific Inc., Folsom, CA) y acoplado con un espectrómetro de masas de la serie 5975. El gas portador fue helio a un flujo constante de 0,7 ml/min. La inyección estaba en modo dividido (1:5) con la temperatura del inyector establecida en 250 °C. La temperatura del horno se programó desde 50 °C (mantenimiento de 5 min) hasta 300 °C a 5 °C/min, a continuación a 340 °C a 50 °C/min y se mantuvo durante 3 min. La identificación de productos se basó en espectros de masas e índice de retención. Las raíces de *Paeonia anomala* contenían una pequeña cantidad de drimenol (figura 5).

20 Las raíces nuevas de *Paeonia anomala* se utilizaron para el análisis del transcriptoma. El ARN total se extrajo utilizando Column Plant RNAout (TIANDZ, China). Este ARN total se procesó utilizando la técnica Illumina Total RNA-Seq y se secuenció en el secuenciador Illumina MiSeq. Se generaron un total de 9 millones de lecturas de extremos emparejados de 2 × 251 pb. Las lecturas se ensamblaron utilizando el programa informático Trinity (<http://trinityrnaseq.sf.net/>). Se obtuvieron 26457 unigenes con un tamaño promedio de 1109 pb. Los unigenes se anotaron por NCBI Blast (<http://www.ncbi.nlm.nih.gov/>), así como por el programa informático InterProScan (<http://www.ebi.ac.uk/Tools/pfa/ipscan/>). Este enfoque proporcionó las secuencias para 7 nuevas supuestas sesquiterpeno sintasas que incluyen *PaTPS3*. La actividad enzimática de *PaTPS3* se evaluó como se describe en el siguiente ejemplo.

30 Ejemplo 2

Expresión funcional y caracterización de *PaTPS3*.

35 El ARN total extraído con el kit Column Plant RNAout primero se transcribió inversamente en ADNc usando el kit de síntesis de primera cadena SuperScript III (Invitrogen, Shanghai, China). Y luego el producto se usó como molde, se utilizaron cebador directo (5'-ATGCTGTGCAAAGTTCCTCAATC-3') (SEQ ID NO: 4) y cebador inverso (5'-TCACATTGCAATAGGATCGGTG-3') (SEQ ID NO: 5) para amplificar el gen de la biblioteca de ADNc de *P. anomala*

40 Las secuencias de *PaTPS3* se optimizaron siguiendo la frecuencia de codones genéticos de *E. coli* y se sintetizaron. El sitio de restricción de *NdeI* se añadió al extremo 5' de *PaTPS3* mientras *KpnI* se añadió al extremo 3'. *PaTPS3* se subclonó en el plásmido pJ401 (DNA 2.0) para su posterior expresión en *E. coli*.

45 Se cotransformaron células KRX de *E. coli* (Promega) con el plásmido pACYC/ScMVA, que contiene los genes que codifican una vía heteróloga de mevalonato, y el plásmido pJ401-*PaTPS3*. Para construir el plásmido pACYC/ScMVA, se dividieron los ocho genes biosintéticos en 2 operones sintéticos denominados vía de mevalonato (MVA) 'superior' e 'inferior'. Como una vía de MVA superior, se creó un operón sintético que consiste en una acetoacetil-CoA tiolasa de *E. coli* codificada por *atoB*, una HMG-CoA sintasa y una versión truncada de la HMG-CoA reductasa de *Saccharomyces cerevisiae* codificada por *ERG13* y *ERG19*, respectivamente. Este operón transforma el metabolito primario Acetil-CoA en (R)-mevalonato. Como vía de mevalonato 'inferior', se creó un segundo operón sintético que codifica una mevalonato quinasa (*ERG12*, *S. cerevisiae*), una fosfomevalonato quinasa (*ERG8*, *S. cerevisiae*), una fosfomevalonato descarboxilasa (*MVD1*, *S. cerevisiae*), una isopentenil difosfato isomerasa (*idi*, *E. coli*) y una farnesil pirofosfato (FPP) sintasa (*IspA*, *E. coli*). Finalmente, una segunda FPP sintasa de *S. cerevisiae* (*ERG20*) se introdujo en el operón de la vía superior para mejorar la conversión de las unidades isoprenoides C5 (IPP y DMAPP) en pirofosfato de farnesilo (FPP). Cada operón se subclonó en uno de los sitios de clonación múltiple de un plásmido de 55 baja expresión de copias bajo el control de un promotor del bacteriófago T7 (pACYCDuet-1, Invitrogen).

60 Las células cotransformadas se seleccionaron en placas de agar LB que contenían kanamicina (50 µg/ml final) y cloranfenicol (34 µg/ml final). Se usaron colonias individuales para inocular 5 ml de medio LB líquido que contenía kanamicina (25 µg/ml final) y cloranfenicol (34 µg/ml final). Los cultivos se incubaron durante la noche a 37 °C y con agitación de 200 rpm. Al día siguiente, se inocularon 6 ml de medio TB complementado con los mismos antibióticos y glicerol (3 % p/v final) con 0,6 ml de los cultivos de una noche. Después de 4 horas de incubación a 37 °C y agitación a 200 rpm, los cultivos se enfriaron a 25 °C durante una hora. El volumen de los cultivos se ajustó a 2 ml por cada tubo y se añadió IPTG (0,1 mM final), cubierto con 200 µl de dodecano. Los cultivos se incubaron durante otras 48 horas a 25 °C y agitación de 200 rpm. A continuación, los cultivos se extrajeron con 1 ml de acetato de etilo y se añadieron 50 µl de isolongifoleno (patrón interno) a 2 mg/ml como patrón interno antes de analizar las muestras por GC/MS. El análisis GC/MS usó el mismo método que el descrito en el ejemplo 1. El gas portador fue helio a un flujo

5 constante de 0,7 ml/min. La inyección estaba en modo dividido (1:5) con la temperatura del inyector establecida en 250 °C. La temperatura del horno se programó desde 50 °C (mantenimiento de 5 min) hasta 300 °C a 5 °C/min, luego de 340 °C a 50 °C/min y se mantuvo durante 3 min. La identificación de los productos se basó en espectros de masas e índices de retención. El análisis GC/MS reveló que *PaTPS3* produjo drimenol como producto principal con una selectividad del 73 % (sin incluir farnesol ni farnesil acetato) y un título de 18,4 mg/l (figuras 6 y 7).

10 Se realizó un ensayo *in vitro* para confirmar la caracterización anterior *in vivo* de *PaTPS3*. Se transformaron células BL21 (DE3) de *E. coli* con el plásmido pJ401-*PaTPS3*. Las células transformadas se seleccionaron en placas de agar LB que contenían kanamicina (50 µg/ml final). Se usaron colonias individuales para inocular 25 ml de medio LB líquido complementado con el mismo antibiótico. Los cultivos se incubaron a 37 °C y 200 rpm con agitación hasta que se enturbiaron (DO alrededor de 0,5). Después de 5 horas de incubación, los cultivos se enfriaron a 20 °C durante 30 min y se añadió IPTG (0,1 mM final). Los cultivos se incubaron a 20 °C y 200 rpm durante la noche, luego se centrifugaron y se resuspendieron en 5 ml de tampón MOPSO 50 mM (que contenía 10 % de glicerol p/v y DTT 5 mM, pH 7). Las células resuspendidas se rompieron mediante ultrasonidos en hielo durante 10 segundos 3 veces y se centrifugaron a 4 °C, 12000 rpm durante 30 minutos, el sobrenadante (que contiene la proteína en bruto) se usó en un ensayo *in vitro*. Un total de 2 ml de tampón de reacción MOPSO 50 mM (que contiene 10 % de glicerol p/v, MgCl₂ 15 mM, MnCl₂ 0,1 mM, DTT 1 mM, Na₃VO₄ 6 mM, pH 7), 10 µl de la solución de FPP 145 µM y 1 ml de proteína en bruto se mezclaron y se cubrieron con 1 ml de heptano, a continuación se incubaron durante 16 horas para reacción *in vitro*. A continuación, la reacción se extrajo con 1 ml de acetato de etilo y se añadieron 50 µl de isolongifoleno (patrón interno) a 2 mg/ml como patrón interno antes de analizar las muestras por GC/MS. El análisis GC/MS usó el mismo método que el descrito en el ejemplo 1. El gas portador fue helio a un flujo constante de 0,7 ml/min. La inyección estaba en modo dividido (1:5) con la temperatura del inyector establecida en 250 °C. La temperatura del horno se programó desde 50 °C (mantenimiento de 5 min) hasta 300 °C a 5 °C/min, luego de 340 °C a 50 °C/min y se mantuvo durante 3 min. La identificación de los productos se basó en espectros de masas e índices de retención. El análisis GC/MS reveló que *PaTPS3* produjo drimenol con una selectividad de un 21 % junto con 79 % de (E)-nerolidol (figuras 8 y 9).

Listado de secuencias

30 SEQ ID NO: 1
secuencia de ADNc de *PaTPS3*:

ATGTCTGTCAAAGTTCCTCAATCTCAGAATGCTCCTACAGAGGTTGGACGTCGGTCC
GTA AATTTTCATCCTACTGTTTGGGGAGATCGGTTTATCACATAACAATAACCAGTCA
GTTGATGATGATGTGGAAAAGAGATTAACAAAAGA AACTAAAATCCCAAGTGAGGAG
AAAGTTGGTGGATGCTGCTGAAAATACATGTCAGAAGCTTAACACAATCGATGCAA
TCGAGAGATTAGGCTTGGCTTATCATTTCGAAACAGAGATTGAAGAAGCACTGCAA

AATATTTATAATTCCCTCTCAGGTTGTTGGAAATAATGTGGAAGAAGATGACCTCTAC
TCTGTTGCCTTACGCTTTAGGCTTCTCAGACAACAGGGCTACAATATTTTCATCTGATG
TGTTTAACAAATTCAAAGATGATAAAGATAACTTCAAGGTATCTTTAATTGGTGATG
CATCAAGCTTGCTAAGCCTATATGAAGCTGCACACCTTCGAGTACACGGAGAACAC
ATACTGGATGAAGCTCTAACTTTCTCAGTTAATAATCTGGAATCAATGGCAACCCAA
TTAAGTCCACCCCTTGCAACACATGTAACCCATGCACTAAACAGACCCTTCGAAAG
GGCATTCCAAGGCTAGAAGCAAGGCACTACATTTCTGTCTACGAACAAGATCCTTTA
CACGATGAAGATCTATTGAAGCTCGCAAAGTTAGATTTCAACCAATTACAGAAAATT
CACCAGAAGGAGCTAAGCGAGATCTCAAAGTGGTGGAAAGATATAAACTTTGTATC
AAAGCTACCTTTTGCAAGGGACAGAGTGGTGGAGTGCTACTTTTGGATAATGTCAGT
GCATAGCGAGCCCGAGAAGCTGGCTTGCACGAAGGACAGCTGCAAAAATAGCTGCGG
TAACCTCCATTATAGATGATATCTATGATGTGCATGGTACAATGACGAAGTACGCG
TATTTACAGAAGCCGTCAACAGGTGGGATATAAACAACATTGATCAACTCCCGGAG
TACATGAAAATATGTTATAAGGCGCTCTTGGGCGTTTTTGTAGTGAATTAGGGGAAGAG
TTGAAAAACAAGGAAGATCTTACCGCCTCGATCATAACAATTGAACTTATGAAAGA
TCTAGTTGGGAAGTATTTTACTGAATCGAAATGGTTAAGCGAAAAATATGTGCCAC
AATAGAGGAGTATATGCGTGCTGCAGAAGTCACCATAGGTTACAACAATGCTATAA
CTGCATCTTTTGCCACAGCCAAAGCCGGAGATATTGCAACCAAGGAGACCTTTGAAT
GGGTGTTGAGTGAACCTAAAATTGTTAAGGCTTCCTCAGTAATTTGCAGGTTGATGG
ATGACTTATCATCCACAAGTTTGAGCAAAAAGAGAGGACATGTTGCATCTGCTATTG
AATGCTACATGAAGCAACATGATGCTACAGAGGAAAAGGTGCGTGCGGAGTTTAAAT
AAACAAGTCACCGACGCCTGGAAGGTGATAAATCAAGAATGTCTCCACCCAACAGC
CATTCCAATGCCTCTTCTTACATGTGTTCTCAACTATGCACGTGTGGCTGATGTCATG
TACAAGGATGGAGATGCTTATACATTTGCCAGATCTTACTGAAAGATCATTATCG
GCATTGTTACCGATCCTATTGCAATGTGA

SEQ ID NO: 2
Secuencia de aminoácidos de *PaTPS3*:

5

MSVKVPQSQNAPTEVGRRSVNFHPTVWGDRFFTYNNQSVDDDVEKRLTKELKSQVRRK
LVDAEAENTCQKLNTIDAIERLGLAYHFETEIEEALQNIYNSSQVVGNNVEEDDLYSVALR
FRLLRQQGYNISSDVFNKFKDDKDNFKVSLIGDASSLLSLYEAHLRVHGEHILDEALTF
SVNNLESMATQLSPPLATHVTHALNRPLRKGIPRLEARHYISVYEQDPLHDEDLLKLAKL
DFNQLQKIHQKELSEISKWWKDINFVSKLPFARDRVVECYFWIMSVHSEPENWLARRTA
AKIAAVTSIIDDIYDVHGTIDELTLFTEAVNRWDINNIDQLPEYMKICYKALLGVFSELGE
ELEKQGRSYRLDHTIELMKDLVGNYFTESKWLSEKYVPTIEEYMRAAEVTIGYNNAITA
SFATAKAGDIATKETFEWVLSPEKIVKASSVICRLMDDLSSHKFEQKRGHVASAIECYM
KQHDATEEKVRAEFNKQVTDWVKVINQECLHPTAIPMPLLTCLVLYARVADVMYKDG
DAYTFAQILLKDHLSALFTDPIAM

SEQ ID NO: 3
secuencia de ADNc con codones optimizados de *PaTPS3* para expresión en *E. coli*:

10

ATGTCCGTTAAAGTTCCGCAAAGCCAAAATGCCCTACCGAAGTTGGCCGTCGTTCC
 GTCAACTTCCACCCGACGGTCTGGGGTGATCGTTTTATTACCTACAATAACCAGAGC
 GTTGACGACGATGTGGAAAAGCGTTTTGACCAAAGAATTGAAGTCCCAGGTCCGTCG
 TAAACTGGTTGACGCTGCAGAGAACAACCTTGCCAGAACTGAACACCATCGACGCGA
 TCGAGCGCCTGGGTCTGGCTTACCATTTGAGACTGAGATTGAAGAGGCACTGCAGA
 ACATCTACAATTCCAGCCAAGTCGTGGGCAATAATGTAGAGGAAGATGATTTATATA
 GCGTGGCGCTGCGTTTTTCGTCTGCTGCGTCAACAGGGTTATAACATCAGCTCCGATG
 TCTTTAACAAGTTCAAAGATGATAAAGACAATTTCAAGGTTAGCCTGATCGGTGACG
 CAAGCTCTTTGTTATCTCTGTATGAAGCCGCGCATCTGCGCGTGCATGGCGAGCATA
 TCTTGATGAAGCGCTGACCTTTAGCGTTAATAATCTGGAATCGATGGCAACCCAGC
 TGAGCCC GCCGTGGCAACGCACGTTACGCACGCGTTGAACCGCCCGCTGCGCAAG
 GGTATCCCGCGTCTGGAAGCGCGTCAATTACATTTCTGTGTACGAACAAGATCCACTG
 CACGACGAAGATTTGCTTAAACTGGCGAAACTGGATTTTAATCAACTGCAAAAAGATT
 CACCAGAAAGAACTGAGCGAGATTAGCAAATGGTGGAAAGACATTAATTTTCGTCAG
 CAAGCTGCCGTTCCGCCGCGACCGTGTGTGGAGTGTCTATTTCTGGATTATGAGCGT
 TCACAGCGAGCCTGAGAACTGGCTGGCGCGCCGCACCGCGGCTAAGATTGCGGCAG
 TCACGTGATTATCGACGATATCTATGACGTCCACGGCACCATCGATGAACTGACGC
 TGTTACCGAAGCCGTTAACCGCTGGGACATCAACAACATTGATCAGCTGCCGGAAT
 ACATGAAGATCTGCTACAAAGCGCTGCTGGGCGTGTTTCAGCGAGCTGGGTGAAGAA
 CTGGAGAAACAGGGTCGTAGCTATCGCTTGGATCATAACCATTGAGCTGATGAAAGA
 TCTGGTCCGTAATTACTTCACCGAGTCCAAGTGGCTGAGCGAGAAATACGTTCCGAC
 GATCGAAGAGTACATGCGTGCTGCCGAAGTGACCATCGGTTACAACAATGCCATTA
 CGGCATCTTTTGCCACGGCAAAGGCCGGTGATATCGCTACCAAAGAAACCTTTGAAT
 GGGTGTGAGCGAACCGAAGATTGTCAAAGCCTCCAGCGTTATTTGTGCTCTGATGG
 ACGATTTGAGCAGCCATAAGTTTGAGCAAAAAGCGTGGCCACGTCGCGAGCGCGATC
 GAGTGCTATATGAAACAGCACGACGCGACCGAGGAAAAAGTTTCGTGCAGAGTTCAA
 TAAACAAGTCACCGATGCGTGGAAAAGTCATTAACCAAGAGTGTTCACCCGACGG
 CGATCCCGATGCCACTGCTGACCTGTGTGCTCAATTATGCACGTGTTGCGGACGTTA
 TGTATAAGGATGGTGACGCGTATACCTTTGCGCAAATTCTGCTGAAAGACCACCTGA
 GCGCACTGTTACGGACCCGATCGCGATGTAA

5 SEQ ID NO: 4
 cebador directo
 ATGTCTGTCAAAGTTCCTCAATC

10 SEQ ID NO: 5
 cebador inverso
 TCACATTGCAATAGGATCGGTG

REIVINDICACIONES

1. Un método para producir drimenol o una mezcla que comprende drimenol que comprende:
 - 5 a. poner en contacto un farnesil difosfato (FPP) acíclico con un polipéptido que tiene actividad sesquiterpeno sintasa, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o comprende la secuencia de aminoácidos de la SEQ ID NO:2; y
 - 10 b. opcionalmente aislar el drimenol.
2. El método de la reivindicación 1, que comprende transformar una célula hospedadora o un organismo hospedador no humano con un ácido nucleico que codifica un polipéptido que tiene actividad drimenol sintasa, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 2 o comprende la SEQ ID NO: 2.
- 15 3. El método descrito en las reivindicaciones 1 o 2, que comprende además cultivar un organismo hospedador no humano o una célula hospedadora capaces de producir FPP y transformados para expresar un polipéptido, en donde el polipéptido comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad de secuencia con la SEQ ID NO: 2 en condiciones que permitan la producción del polipéptido.
- 20 4. El método descrito en una cualquiera de las reivindicaciones 1 a 3, que comprende convertir el drimenol en un derivado de drimenol usando síntesis química o síntesis bioquímica.
- 25 5. Un polipéptido aislado que tiene actividad sesquiterpeno sintasa que comprende una secuencia de aminoácidos que tiene al menos un 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia de la SEQ ID NO: 2 o que comprende la secuencia de aminoácidos de la SEQ ID NO:2.
6. Una molécula de ácido nucleico aislada
 - 30 a. que comprende una secuencia de nucleótidos que codifica el polipéptido de la reivindicación 5; o
 - b. que comprende la secuencia de nucleótidos de la SEQ ID NO: 1 o SEQ ID NO: 3.
7. Un vector que comprende
 - 35 a. la molécula de ácido nucleico de la reivindicación 6; o
 - b. un ácido nucleico que codifica el polipéptido de la reivindicación 5.
8. El vector de la reivindicación 7, en donde el vector es un vector procarionta, vector viral o un vector eucariota.
- 40 9. Una célula hospedadora o un organismo hospedador no humano que comprende un gen recombinante, comprendiendo dicho gen:
 - a. el ácido nucleico aislado de la reivindicación 6; o
 - 45 b. el vector de las reivindicaciones 7 u 8.
10. El método de las reivindicaciones 2 o 3, en donde la célula o el organismo hospedador no humano es una planta, un procarionta, un hongo o un microorganismo.
- 50 11. El método de la reivindicación 10, en donde el microorganismo es una bacteria o una levadura.
12. El método de la reivindicación 11, en donde la bacteria es *E. coli* y la levadura es *Saccharomyces cerevisiae*.
13. El uso del polipéptido de la reivindicación 7 para producir drimenol o una mezcla que comprende drimenol y uno o más terpenos.
- 55 14. El uso de la reivindicación 13, en donde la mezcla comprende drimenol y nerolidol.
15. El método de la reivindicación 1, en donde la mezcla comprende drimenol y nerolidol.

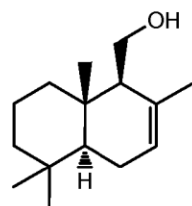


Figura 1

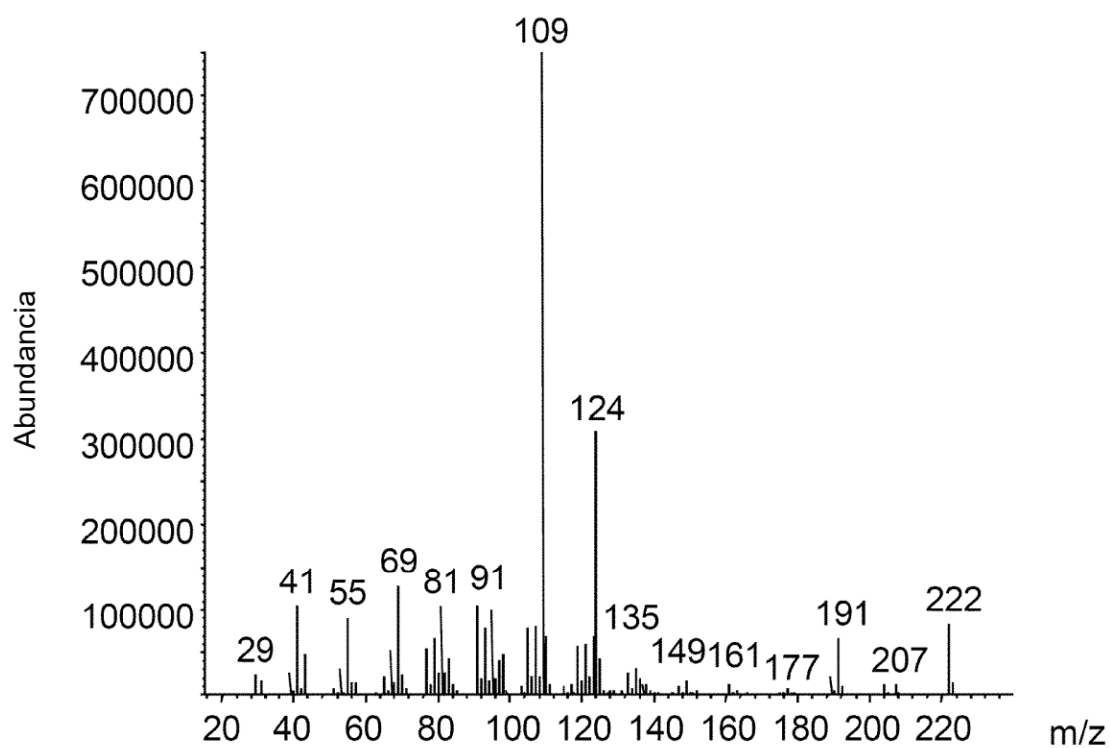


Figura 2

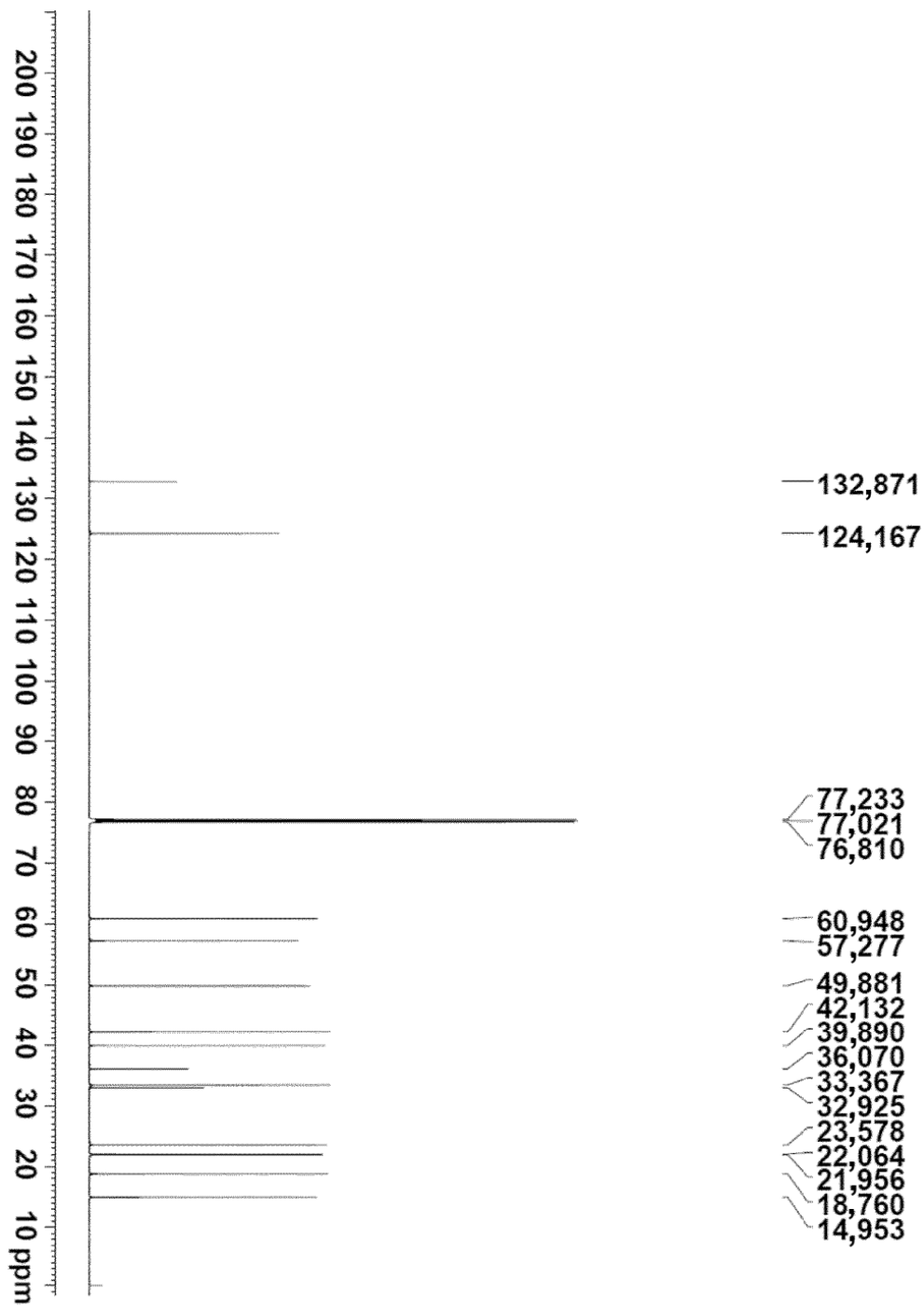


Figura 3

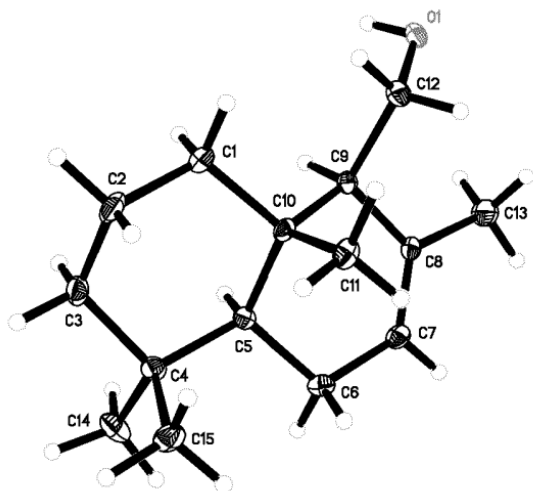


Figura 4

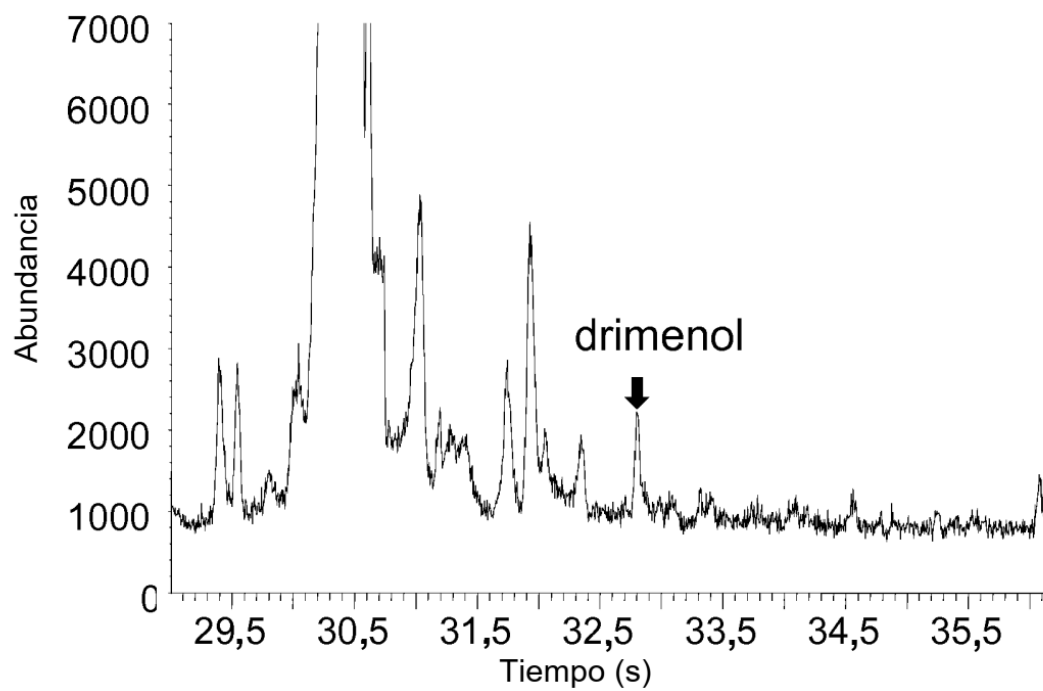


Figura 5

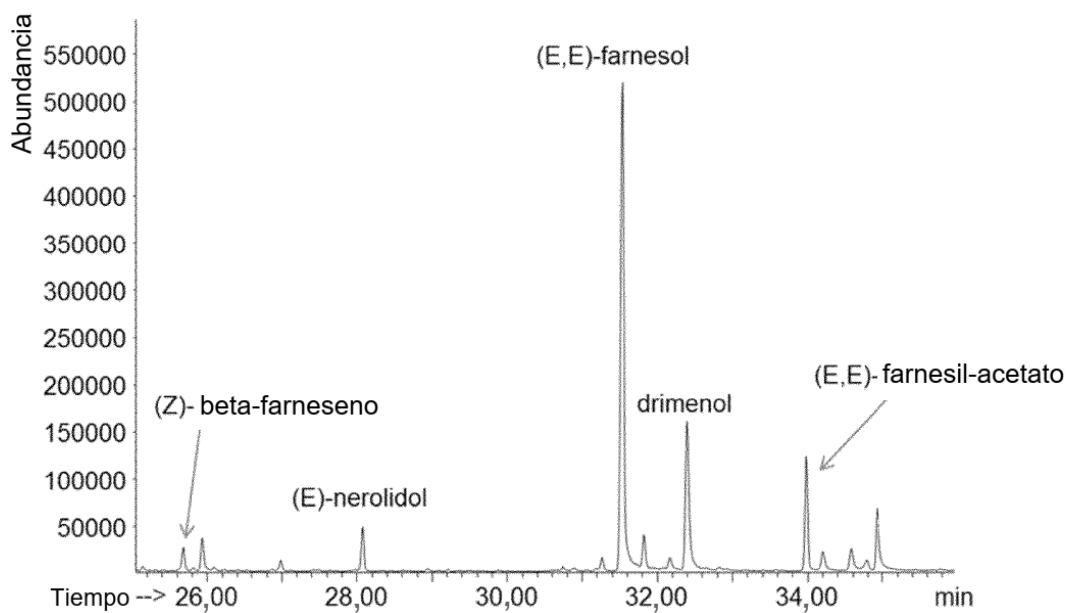


Figura 6

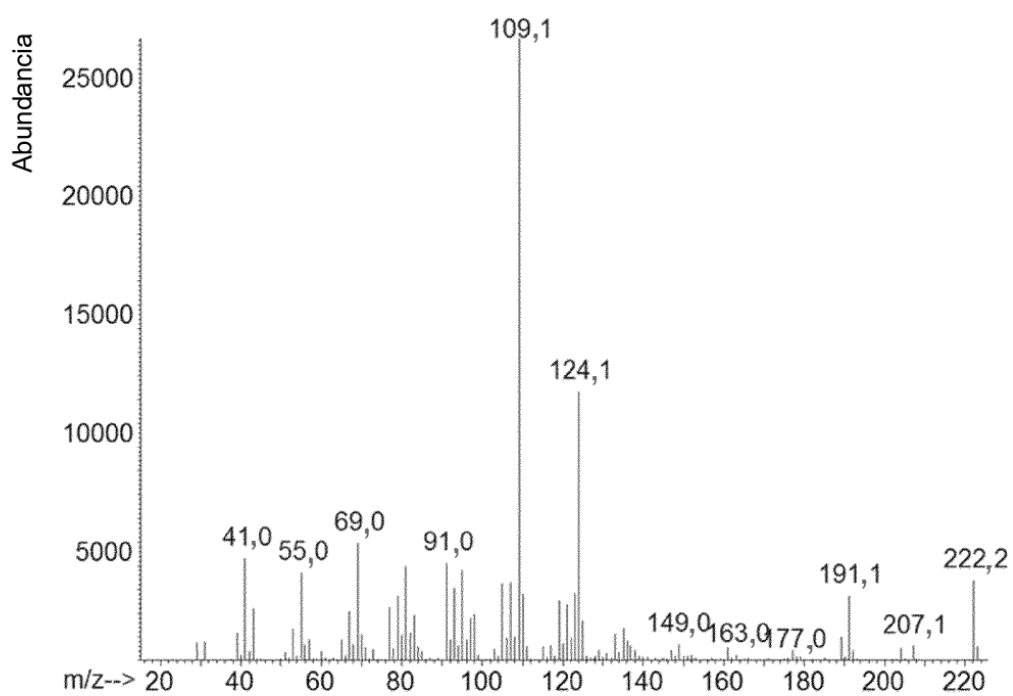


Figura 7

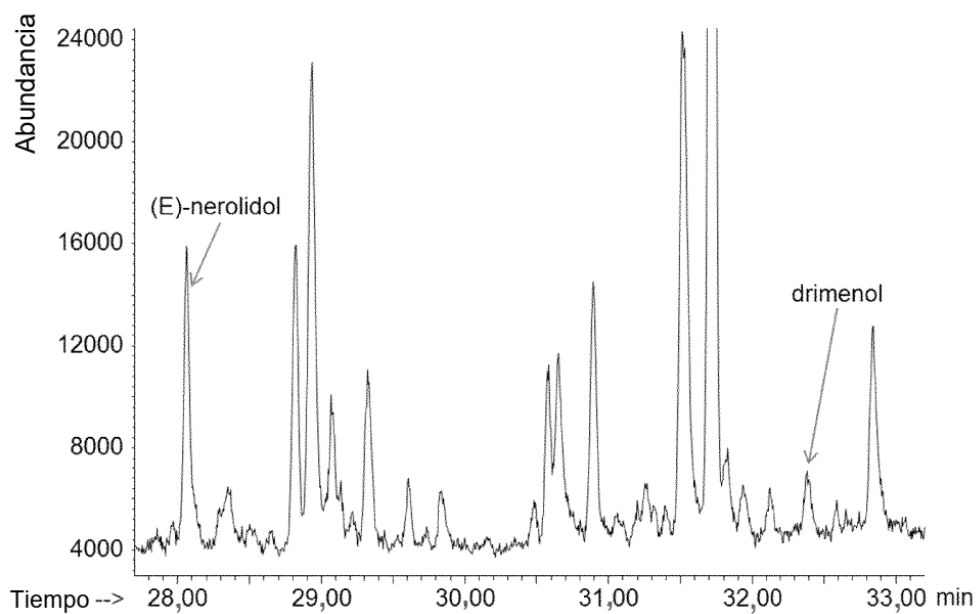


Figura 8

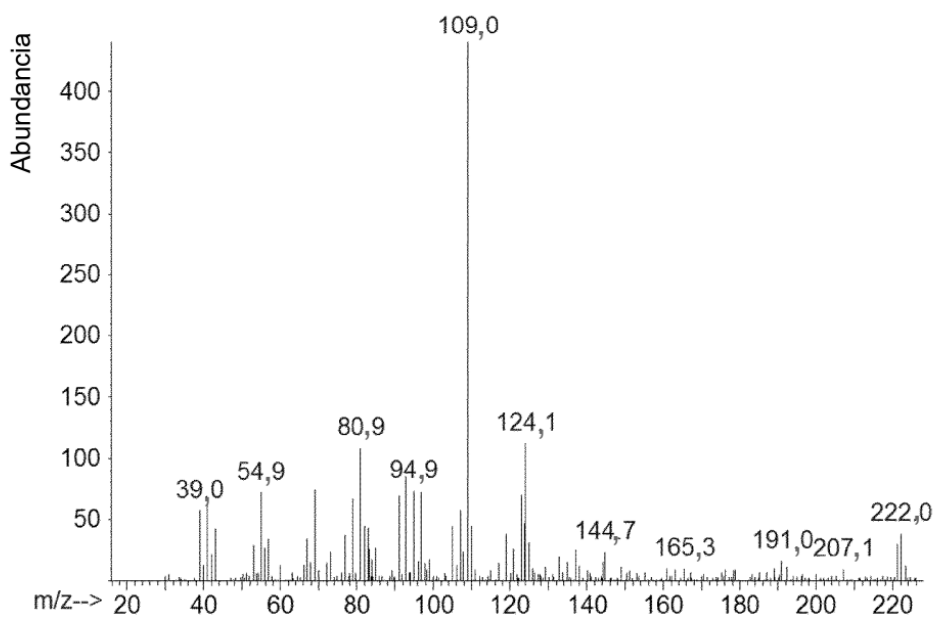


Figura 9