

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 759 724**

51 Int. Cl.:

C12P 7/02 (2006.01)

C12N 9/16 (2006.01)

C12N 15/82 (2006.01)

C12N 9/02 (2006.01)

C12P 7/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.05.2015 PCT/EP2015/059988**

87 Fecha y número de publicación internacional: **12.11.2015 WO15169871**

96 Fecha de presentación y número de la solicitud europea: **06.05.2015 E 15730972 (5)**

97 Fecha y número de publicación de la concesión europea: **11.09.2019 EP 3140410**

54 Título: **Drimenol sintasas y procedimiento de producción de drimenol**

30 Prioridad:

06.05.2014 WO PCT/CN2014/076890

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

12.05.2020

73 Titular/es:

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

72 Inventor/es:

**HAEFLIGER, OLIVIER;
HE, XIU-FENG;
ZHANG, YU-HUA;
SCHALK, MICHEL y
DEGUERRY, FABIENNE**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 759 724 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

Drimenol sintasas y procedimiento de producción de drimenol

Campo técnico

5 El campo se refiere a los procedimientos de producción de drimenol, comprendiendo dicho procedimiento poner en contacto al menos un polipéptido con pirofosfato de farnesilo (FPP, por sus siglas en inglés). En particular, dicho procedimiento puede llevarse a cabo *in vitro* o *in vivo* para producir drimenol, un compuesto muy útil en el campo de la perfumería. También se proporciona en el presente documento una secuencia de aminoácidos de un polipéptido útil en los procedimientos proporcionados en el presente documento. Un ácido nucleico que codifica el polipéptido de una realización en el presente documento y un vector de expresión que contiene dicho ácido nucleico también se proporcionan en el presente documento. Se proporciona además en el presente documento un organismo hospedador no humano o una célula transformada para su uso en el procedimiento de producción de drimenol.

Antecedentes

15 Los terpenos se encuentran en la mayoría de los organismos (microorganismos, animales y plantas). Estos compuestos están constituidos por cinco unidades de carbono llamadas unidades isopreno y se clasifican por el número de estas unidades presentes en su estructura. De esta manera, los monoterpenos, sesquiterpenos y diterpenos son terpenos que contienen 10, 15 y 20 átomos de carbono, respectivamente. Los sesquiterpenos, por ejemplo, se encuentran ampliamente en el reino vegetal. Muchas moléculas de sesquiterpeno se conocen por sus propiedades de sabor y fragancia y sus efectos cosméticos, medicinales y antimicrobianos. Se han identificado numerosos hidrocarburos de sesquiterpeno y sesquiterpenoides.

20 La producción biosintética de los terpenos implica enzimas denominadas terpeno sintasas. Existe prácticamente una infinidad de sesquiterpeno sintasas presentes en el reino vegetal, usando todas el mismo sustrato (pirofosfato de farnesilo, FPP) pero que tienen diferentes perfiles de producto. Los genes y los ADNc que codifican las sesquiterpeno sintasas se han clonado y las enzimas recombinantes correspondientes se han caracterizado.

25 Actualmente, la fuente principal de drimenol son las plantas que contienen drimenol de forma natural y los contenidos de drimenol en estas fuentes naturales son bajos. Las concentraciones de drimenol se han cuantificado en hojas de varias poblaciones de *Drimys winteri* y *Drimys andina* (Munoz-Concha y col., Biochem. System. Ecol., 35: 434-438, (2017)). Los ácidos nucleicos que codifican enzimas drimenol sintasa se han aislado de especies vegetales tales como *Valeriana officinalis* y/o *Persicaria hydropiper* y se han usado para sintetizar drimenol *in vitro* o células hospedadoras u organismos genéticamente modificados (documento WO 2013/058655; Pyle ("Molecular Cloning and Characterization of Sesquiterpene Synthases from Valeriana officinalis", Universidad de Calgary, Tesis, (2012)). Se han desarrollado enfoques de síntesis química pero son todavía complejos y no son rentables.

Sumario

Se proporciona en el presente documento un procedimiento para la producción de drimenol que comprende:

- 35 i) poner en contacto difosfato de farnesilo (FPP) con un polipéptido que tiene actividad drimenol sintasa y que tiene al menos el 70 % de identidad de secuencia a una secuencia de SEQ ID NO: 2 para producir el drimenol; y
ii) aislar opcionalmente el drimenol.

Se proporciona además en el presente documento un polipéptido aislado que tiene actividad drimenol sintasa que comprende una secuencia de aminoácidos que tiene al menos el 70 % de identidad de secuencia a una secuencia de aminoácidos de SEQ ID NO: 2.

40 También se proporciona en el presente documento una molécula de ácido nucleico aislada que codifica un polipéptido que tiene al menos el 70 % de identidad de secuencia a una secuencia de SEQ ID NO: 2.

También se proporciona un vector que comprende la molécula de ácido nucleico anteriormente mencionada y un organismo o célula hospedadores no humanos que comprenden la molécula de ácido nucleico o el vector anteriormente mencionados.

Breve descripción de los dibujos

Figura 1. Análisis de GCMS de las hojas de *Drimys lanceolata* y *Drimys winteri*.

Figura 2. Análisis de GCMS de los sesquiterpenos producidos por DITps589 recombinante en ensayos *in vitro*. A. Cromatograma de iones totales del perfil de sesquiterpeno de una incubación de la proteína DITps589 recombinante con FPP. B. Control negativo realizado en las mismas condiciones con células de *E. coli* transformadas con un plásmido vacío. C. Espectro de masa del pico a 11,76 minutos. D. Espectro de masa de un patrón auténtico de (-)-drimenol.

Figura 3. Análisis de GCMS de los sesquiterpenos producidos *in vivo* por DITps589 recombinante en células de bacterias manipuladas por ingeniería genética. A. Cromatograma de iones totales. B. Espectro de masa del pico a 11,49 minutos. C. Espectro de masa de un patrón auténtico de (-)-drimenol. El compuesto que eluye a 10,98

minutos es el farnesol producido por la enzima DITps589 o que resulta de la hidrólisis de FPP en exceso producida por las células de *E. coli*.

Figura 4. Estructura del (-)-drimenol producido por la DITps589 sintasa recombinante.

Figura 5. Cromatogramas de GC quiral/FID de (-)-drimenol producido por la enzima recombinante (parte superior), drimenol racémico obtenido químicamente (parte intermedia) y (-)-drimenol auténtico (parte inferior).

Figura 6. Cromatograma de iones totales del análisis de GCMS de los sesquiterpenos producidos en ensayo *in vitro* por las proteínas recombinantes SCH51-3228-9 (A), SCH51-998-28 (B) o SCH52-13163-6 (C).

Figura 7. Cromatograma de iones totales del análisis de GCMS de los sesquiterpenos producidos *in vivo* mediante células de bacterias manipuladas por ingeniería genética que expresan las proteínas recombinantes diferentes SCH51-3228-9 (A), SCH51-998-28 (B) o SCH52-13163-6 (C). El farnesol detectó los resultados provenientes de la hidrólisis de FPP en exceso, producido por las células de *E. coli*, y pudo producirse en parte por las proteínas recombinantes.

Descripción detallada

Para las descripciones en el presente documento y las reivindicaciones adjuntas, el uso de "o" significa "y/o" a no ser que se establezca de otro modo. De manera similar, "comprenden", "comprende", "que comprende", "incluyen", "incluye", y "que incluyen" son intercambiables y no se pretende que sean limitantes.

Debe entenderse además que cuando las descripciones de las diversas realizaciones usan la expresión "que comprende", aquellos expertos en la materia podrían entender que en algunos casos específicos, una realización puede describirse alternativamente usando la frase "que consiste esencialmente en" o "que consiste en".

En un aspecto, se proporciona en el presente documento un procedimiento para producir drimenol que comprende:

- i) poner en contacto difosfato de farnesilo (FPP) con un polipéptido que tiene actividad de drimenol sintasa y que tiene al menos un 70 %, particularmente un 75 %, particularmente un 80 %, particularmente un 85 %, particularmente un 90 %, particularmente un 95 %, particularmente un 96 %, particularmente un 97 %, particularmente un 98 % o particularmente un 99 % o más de identidad de secuencia a una secuencia de SEQ ID NO: 2 para producir el drimenol; y
- ii) opcionalmente aislar el drimenol.

En un aspecto, el drimenol, se aísla.

Se proporciona además en el presente documento un polipéptido aislado que tiene actividad drimenol sintasa que comprende una secuencia de aminoácidos que tiene al menos un 70 %, particularmente un 75 %, particularmente un 80 %, particularmente un 85 %, particularmente un 90 %, particularmente un 95 %, particularmente un 96 %, particularmente un 97 %, particularmente un 98 % o más particularmente un 99 % o más de identidad a la secuencia de aminoácidos de SEQ ID NO: 2.

Se proporciona además en el presente documento una molécula de ácido nucleico aislada que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 70 %, particularmente un 75 %, particularmente un 80 %, particularmente un 85 %, particularmente un 90 %, particularmente un 95 %, particularmente un 96 %, particularmente un 97 %, particularmente un 98 % o más particularmente un 99 % o más de identidad a la secuencia de aminoácidos de SEQ ID NO: 2.

Se proporciona además en el presente documento una molécula de ácido nucleico que comprende la secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 y SEQ ID NO: 3. Se desvelan además en el presente documento moléculas de ácido nucleico que comprenden la secuencia del grupo que consiste en SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13 y SEQ ID NO: 15.

Se proporciona además en el presente documento un procedimiento como se indica en la reivindicación 1 que comprende las etapas de transformar una célula hospedadora o un organismo no humano con un ácido nucleico que codifica un polipéptido que tiene al menos un 70 %, particularmente un 75 %, particularmente un 80 %, particularmente un 85 %, particularmente un 90 %, particularmente un 95 %, particularmente un 96 %, particularmente un 97 %, particularmente un 98 % o particularmente un 99 % o más de identidad de secuencia de la secuencia de SEQ ID NO: 2 y cultivar la célula hospedadora o el organismo bajo condiciones que permitan la producción del polipéptido.

Se proporciona además al menos un vector que comprende las moléculas de ácido nucleicos descritas.

Se proporciona además en el presente documento un vector seleccionado del grupo de un vector procariótico, un vector vírico y un vector eucariótico.

Se proporciona además en el presente documento un vector que es un vector de expresión.

Como una "drimenol sintasa" o como un "polipéptido que tiene una actividad de drimenol sintasa", se entiende en el presente documento un polipéptido capaz de catalizar la síntesis del drimenol, en forma de cualquiera de sus estereoisómeros o una mezcla de los mismos, comenzando a partir de difosfato de farnesilo (FPP). El drimenol puede

ser el único producto o puede ser una mezcla de sesquiterpenos.

La capacidad de un polipéptido para catalizar la síntesis de un sesquiterpeno particular (por ejemplo drimenol) puede ser simplemente confirmada por la realización del ensayo de enzima como se detalla en los Ejemplo 2 a 5.

5 Como se pretende más adelante en el presente documento, "una secuencia de nucleótidos obtenida modificando SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas", abarca cualquier secuencia que se haya obtenido cambiando la secuencia de las SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 o SEQ ID NO: 12 o del complemento de las mismas usando cualquier procedimiento conocido en la técnica, por ejemplo mediante la introducción de cualquier tipo de mutaciones tales como mutaciones por delección, inserción, o sustitución. Los ejemplos de tales procedimientos se citan en la parte de la descripción con relación a los polipéptidos variantes y a los procedimientos para prepararlos.

15 El porcentaje de identidad entre dos secuencias peptídicas o nucleotídicas es una función del número de aminoácidos o de los 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 los mismos 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 al tomar el número de restos idénticos entre dos secuencias dividiéndolo entre el número total de restos en la secuencia más corta y multiplicando por 100.

20 La alineación óptima es la alineación en la que el porcentaje de identidad es el más alto posible. Pueden introducirse espacios vacíos dentro de una o ambas secuencias en una o más posiciones de la alineación para obtener la alineación óptima. Estos espacios vacíos se toman en cuenta después como restos no idénticos para el cálculo del porcentaje de identidad de secuencia. La alineación para el propósito de determinar el porcentaje de identidad de secuencia de aminoácido o de ácido nucleico puede lograrse de diversas formas usando programas de ordenador y por ejemplo los programas de ordenador públicamente disponibles que están disponibles en la red mundial. Preferentemente, el programa BLAST (Tatiana y col, *FEMS Microbiol Lett.*, 1999, 174:247-250, 1999) ajustado a los parámetros por defecto, disponible del National Center for Biotechnology Information (NCBI) en <http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi>, puede usarse para obtener una alineación óptima de las secuencias peptídicas o nucleotídicas y para calcular el porcentaje de identidad de secuencia.

Abreviaturas usadas

30	pb	par de bases
	kb	kilo base
	BSA	albúmina sérica bovina
	ADN	ácido desoxirribonucleico
	ADNc	ADN complementario
	DTT	ditiotreitól
35	FID	Detector de ionización de llama
	FPP	pirofosfato de farnesilo
	GC	cromatografía de gases
	IPTG	isopropil-D-tiogalacto-piranosido
	LB	caldo de lisogenia
40	EM	espectrómetro de masas
	MVA	ácido mevalónico
	PCR	reacción en cadena de la polimerasa
	RMCE	intercambio de casete mediado por recombinasa
	3'-/5'-RACE 3' y 5'	amplificación rápida de los extremos de ADNc
45	ARN	ácido ribonucleico
	ARNm	ácido ribonucleico mensajero
	miARN	micro-ARN
	ARNip	ARN de interferencia pequeño
	ARNr	ARN ribosómico
50	ARNt	ARN de transferencia

Definiciones

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

El término polipéptido "aislado" se refiere a una secuencia de aminoácidos que se retira de su ambiente natural mediante cualquier procedimiento o combinación de procedimientos conocidos en la técnica e incluye procedimientos recombinantes, bioquímicos y sintéticos.

El término "proteína" se refiere a una secuencia de aminoácidos de cualquier longitud en la que los aminoácidos están enlazados por enlaces peptídicos covalentes, e incluye oligopéptido, péptido, polipéptido y proteína de longitud completa, ya sea de origen natural o sintéticos.

5 Los términos "drimenol sintasa" o "proteína drimenol sintasa" se refieren a una enzima que es capaz de convertir el difosfato de farnesilo (FPP) en drimenol.

Los términos "función biológica", "función", "actividad biológica" o "actividad" se refieren a la capacidad de la drimenol sintasa de catalizar la formación del drimenol a partir de FPP.

10 Las expresiones "secuencia de ácido nucleico", "ácido nucleico", y "polinucleótido" se usan intercambiamente significando una secuencia de nucleótidos. Una secuencia de ácido nucleico puede ser un desoxirribonucleótido de monocatenario o bicatenario, o un ribonucleótido de cualquier longitud, e incluyen las secuencias codificantes y no codificantes de un gen, exones, intrones, secuencias complementarias en sentido y anti-sentido, ADN genómico, ADNc, miARN, ARNip, ARNm, ARNr, ARNt, secuencias recombinantes de ácido nucleico, secuencias de ADN y/o ARN de origen natural, aisladas y purificadas, secuencias sintéticas de ADN y ARN, fragmentos, cebadores y sondas de ácido nucleico. El experto en la materia está enterado de que las secuencias de ácido nucleico del ARN son idénticas a las secuencias del ADN con la diferencia de que la timina (T) se reemplaza por uracilo (U).

15 Un "ácido nucleico aislado" o una "secuencia aislada de ácido nucleico" se definen como un ácido nucleico o una secuencia de ácido nucleico que está en un ambiente diferente de aquel en el cual aparece de manera natural el ácido nucleico o la secuencia de ácido nucleico. La expresión "de origen natural" como se usa en el presente documento como se aplica a un ácido nucleico se refiere a un ácido nucleico que se encuentra en una célula en la naturaleza. Por ejemplo, una secuencia de ácido nucleico que está presente en un organismo, por ejemplo en las células de un organismo, que puede aislarse de una fuente en la naturaleza, y que no se ha modificado intencionalmente por un ser humano en el laboratorio, es de origen natural.

20 "Secuencias recombinantes de ácido nucleico" son las secuencias de ácido nucleico que resultan del uso de procedimientos de laboratorio (clonación molecular) para reunir material genético proveniente de más de una fuente, creando una secuencia de ácido nucleico que no es de origen natural y que de otro modo no podría encontrarse en organismos biológicos.

25 "Tecnología de ADN recombinante" se refiere a los procedimientos de biología molecular para preparar una secuencia recombinante de ácido nucleico como se describe, por ejemplo, en Laboratory Manuals editado por Weigel y Glazebrook, 2002 Cold Spring Harbor Lab Press; y Sambrook y col., 1989 Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

30 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, operativamente enlazado a las regiones reguladoras adecuadas, por ejemplo, un promotor. Un gen puede de este modo comprender varias secuencias operativamente enlazadas, tales como un promotor, una secuencia líder 5' que comprende, por ejemplo, las secuencias implicadas en el inicio de la traducción, una región de codificación del ADNc o el ADN genómico, intrones, exones, y/o una secuencia 3' no traducida que comprende, por ejemplo, sitios de terminación de la transcripción.

35 Un "gen quimérico" se refiere a cualquier gen, que no se encuentra normalmente en la naturaleza en una especie, en particular, un gen en el cual una o más partes de la secuencia de ácido nucleico están presentes, que no están asociadas una con la otra en la naturaleza. Por ejemplo, el promotor no está asociado en la naturaleza con parte o toda la región transcrita o con otra región reguladora. La expresión "gen quimérico" se entiende que incluye las construcciones de expresión en las cuales un promotor o una secuencia reguladora de la transcripción está operativamente enlazado a una o más secuencias codificantes o a un antisentido, es decir, el complemento inverso de la cadena en sentido, o la secuencia de repetición invertida (sentido y antisentido, con el cual el transcrito de ARN forma el ADN bicatenario después de la transcripción).

40 Una "3' UTR" o "secuencia 3' no traducida" (también denominada "región 3' no traducida", o "extremo 3'") se refiere a la secuencia de ácido nucleico encontrada en dirección 3' de la secuencia de codificación ARNm eucarióticos) una señal de poliadenilación tal como AAUAAA o variantes de la misma. Después de la terminación de la transcripción, el transcrito de ARNm puede escindirse en dirección 3' de la señal de poliadenilación y puede añadirse una cola poli(A) que está implicada en el transporte del ARNm hacia el sitio de traducción, por ejemplo, el citoplasma.

45 "Expresión de un gen" implica la transcripción del gen y la traducción de ARNm en una proteína. La sobreexpresión se refiere a la producción del producto génico según se mide por los niveles de ARNm, el polipéptido y/o la actividad enzimática en las células u fondo genético similar.

50 "Vector de expresión" como se usa en el presente documento, significa una molécula de ácido nucleico manipulada por ingeniería genética usando procedimientos de biología molecular y la tecnología de ADN recombinante, para la administración de ADN extraño o exógeno dentro de una célula hospedadora. El vector de expresión incluye típicamente las secuencias requeridas para la transcripción apropiada de la secuencia de nucleótidos. La región de codificación habitualmente codifica una proteína de interés, pero puede también codificar un ARN, por ejemplo, un

ARN antisentido, ARNip y similares.

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. La persona experta es capaz de seleccionar un vector adecuado de acuerdo con el sistema de expresión. En una realización, el vector de expresión incluye el ácido nucleico de una realización en el presente documento, operativamente enlazado al menos a una secuencia reguladora, la cual controla la transcripción, la traducción, el inicio y la terminación, tales como un promotor transcripcional, operador o potenciador, o un sitio de unión ribosómica del ARNm y, opcionalmente, que incluye al menos un marcador de selección. Las secuencias nucleotídicas están "operativamente enlazadas" cuando la secuencia reguladora se refiere funcionalmente al ácido nucleico de una realización en el presente documento.

"Secuencia reguladora" se refiere a una secuencia de ácido nucleico que determina el nivel de expresión de 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 operativamente enlazada a la secuencia reguladora. Las secuencias reguladoras comprenden promotores, potenciadores, factores de la transcripción, elementos promotores y similares.

"Promotor" se refiere a una secuencia de ácido nucleico que controla la expresión de una secuencia codificante proporcionando un sitio de unión para la ARN-polimerasa y otros factores requeridos para la transcripción apropiada, incluyendo, sin limitación, sitios de unión al factor de transcripción, sitios de unión a la proteína represora y activadora. El significado del término promotor también incluye la expresión "secuencia reguladora promotora". Las secuencias reguladoras promotoras pueden incluir los elementos en dirección 5' y en dirección 3' que pueden influir en la transcripción, el procesamiento del ARN o la estabilidad de la secuencia de ácido nucleico de codificación, asociada. Los promotores incluyen las secuencias naturalmente derivadas y sintéticas. Las secuencias de ácido nucleico codificantes se localizan habitualmente en dirección 3' del promotor con respecto a la dirección de la transcripción del sitio que inicia al inicio de la transcripción.

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

Como se usa en el presente documento, la expresión "operativamente enlazado" se refiere a un enlace de los elementos polinucleotídicos en una relación funcional. Un ácido nucleico está "operativamente enlazado" cuando éste se coloca dentro de una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor, o más bien una secuencia reguladora de la transcripción, está operativamente enlazado a una secuencia codificante si afecta la transcripción de la secuencia codificante. Operativamente enlazado significa que las secuencias de ADN que están enlazadas están típicamente contiguas. La secuencia nucleotídica asociada a la secuencia promotora puede ser de origen homólogo o heterólogo con respecto a la planta que va a ser transformada. La secuencia puede también ser completa o parcialmente sintética. Independientemente del origen, la secuencia de ácido nucleico asociada a la secuencia promotora será expresada o silenciada de acuerdo con las propiedades del promotor al cual está enlazada después de enlazarse al polipéptido de una realización en el presente documento. El ácido nucleico asociado puede codificar una proteína que se desea expresar o suprimir a lo largo de todo el organismo en cualquier momento o, alternativamente, en un momento específico o en tejidos, células o compartimientos celulares específicos. Tales secuencias nucleotídicas codifican particularmente proteínas que confieren rasgos fenotípicos deseables a las células u organismos hospedadores alterados o transformados con las mismas. Más particularmente, la secuencia nucleotídica asociada conduce a la producción del drimenol en el organismo. Particularmente, la secuencia nucleotídica codifica la drimenol sintasa.

"Péptido diana" se refiere a una secuencia de aminoácido que dirige una proteína, o polipéptido hacia los orgánulos intracelulares, es decir, las mitocondrias, o los plástidos, o hacia el espacio extracelular (péptido de señal de secreción). Una secuencia de ácido nucleico que codifica un péptido diana puede estar fusionada a la secuencia de ácido nucleico que codifica el extremo amino terminal, por ejemplo, el extremo N-terminal, de la proteína o el polipéptido, o puede usarse para reemplazar un polipéptido nativo de dirección a la diana.

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

Como se usa en el presente documento, la expresión "célula hospedadora" o "célula transformada" se refiere a una célula (u organismo) alterada para albergar al menos una molécula de ácido nucleico, por ejemplo, un gen recombinante que codifica una proteína o secuencia de ácido nucleico deseada que tras la transcripción produce una proteína drimenol sintasa útil para producir drimenol. La célula hospedadora es particularmente una célula bacteriana, una célula fúngica o una célula vegetal. La célula hospedadora puede contener un gen recombinante que se ha integrado dentro de los genomas nucleares o de los orgánulos de la célula hospedadora. Alternativamente, el hospedador puede contener el gen recombinante extra-cromosómicamente. Las secuencias homólogas incluyen secuencias ortólogas o parálogas. Los procedimientos de identificar los ortólogos o parálogos incluyendo los procedimientos filogenéticos, la similitud de secuencia y los procedimientos de hibridación, se conocen en la técnica y se describen en el presente documento.

Los parálogos resultan de la duplicación génica que da lugar a dos o más genes con secuencias similares y funciones similares. Los parálogos típicamente se agrupan y se forman por duplicaciones de genes dentro de especies vegetales

relacionadas. Los parálogos se encuentran en grupos de genes similares usando el análisis Blast por pares o durante el análisis filogenético de familias de genes utilizando programas tales como CLUSTAL. En los parálogos, las secuencias consenso pueden identificarse características a las secuencias dentro de genes relacionados, y que tienen funciones similares de los genes.

- 5 Los ortólogos, o secuencias ortólogas, son secuencias similares entre sí debido a que éstas se encuentran en especies que descienden de un ancestro común. Por ejemplo, se sabe que las especies vegetales que tienen ancestros comunes contienen muchas enzimas que tienen secuencias y funciones similares. El experto en la materia puede identificar las secuencias ortólogas y predecir las funciones de los ortólogos, por ejemplo, mediante la construcción de un árbol filogenético para una familia de genes de una especie utilizando los programas CLUSTAL o BLAST. Un procedimiento para identificar o confirmar las funciones similares entre las secuencias homólogas es mediante la comparación de los perfiles de transcritos en plantas que sobre-expresan o que carecen de (en organismos con genes inactivados/genos con actividad disminuida) de los polipéptidos relacionados. La persona experta en la materia entenderá que los genes que tienen perfiles de transcritos similares, con más del 50 % de transcritos regulados en común, o con más del 70 % de transcritos regulados en común, o más del 90 % de transcritos regulados en común tendrán funciones similares. Los homólogos, parálogos, ortólogos y cualquier otra variante de las secuencias desveladas en el presente documento se espera que funcionen de una manera similar al hacer que las plantas produzcan proteínas drimenol sintasa.

Una realización proporcionada en el presente documento proporciona las secuencias de aminoácidos de las proteínas drimenol sintasa.

- 20 La expresión "marcador seleccionable" se refiere a cualquier gen que tras la expresión puede usarse para seleccionar una célula o células que incluyen el marcador seleccionable. Los ejemplos de marcadores seleccionables se describen más adelante. El experto en la materia sabrá que diferentes marcadores seleccionables de antibióticos, fungicidas, auxótrofos o herbicidas son aplicables a diferentes especies diana.

"Drimenol" para fines de la presente solicitud se refiere a (-)-drimenol (CAS: 468-68-8).

- 25 El término "organismo" se refiere a cualquier organismo multicelular o unicelular no humano tales como una planta, o un microorganismo. Particularmente, un micro-organismo es una bacteria, una levadura, un alga, o un hongo. El término "planta" se usa intercambiamente para incluir células vegetales incluyendo protoplastos vegetales, tejidos vegetales, cultivos de tejidos de células vegetales que dan origen a plantas regeneradas, o partes de las plantas, u órganos de las plantas, tales como raíces, tallos, hojas, flores, polen, óvulos, embriones, frutos y similares. Cualquier planta puede usarse para llevar a cabo los procedimientos de una realización en el presente documento.

- 30 El polipéptido que va a ponerse en contacto con difosfato de farnesilo (FPP), *in vitro*, puede obtenerse por extracción a partir de cualquier organismo que lo expresa, usando tecnologías convencionales de extracción de proteínas o de enzimas. Si el organismo hospedador es un organismo unicelular o una célula que libera el polipéptido al medio de cultivo, el polipéptido puede simplemente recogerse del medio de cultivo, por ejemplo mediante centrifugación, seguido opcionalmente de etapas de lavado y re-suspensión en soluciones tamponantes adecuadas. Si el organismo o la célula acumulan el polipéptido dentro de sus células, el polipéptido puede obtenerse mediante desintegración o lisis de las células y extracción posterior del polipéptido a partir del lisado celular.

- 35 El polipéptido que tiene una actividad de drimenol sintasa, ya sea en una forma aislada o junto con otras proteínas, por ejemplo en un extracto bruto de proteínas obtenido a partir de las células cultivadas o los microorganismos, puede después suspenderse en una solución tamponante a un pH óptimo. Si es adecuado, pueden añadirse sales, DTT, cationes inorgánicos y otros tipos de cofactores enzimáticos, con el fin de optimizar la actividad enzimática. El FPP se añade a la suspensión del polipéptido, que después 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 de aislamiento convencionales, tales como extracción con disolvente y destilación, opcionalmente después de la retirada de los polipéptidos de la solución.

- 40 De acuerdo con otra realización particular, 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 hospedador no humano o una célula capaces de producir FPP y transformarlo para expresar al menos un polipéptido que comprende una secuencia de aminoácidos al menos un 70 % idéntica a una secuencia de SEQ ID NO: 2 y que tiene actividad drimenol sintasa, en condiciones que conducen a la producción del drimenol.

- 45 De acuerdo con una realización más particular, el procedimiento comprende además, antes de la etapa a), la transformación de 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 un 70 % idéntica a una secuencia de SEQ ID NO: 2 y que tiene una actividad drimenol sintasa, de modo que el organismo expresa dicho polipéptido.

- 55 Estas realizaciones proporcionadas en el presente documento son particularmente ventajosas ya que es posible llevar a cabo el procedimiento *in vivo* sin aislar previamente el polipéptido. La reacción se produce directamente dentro del organismo o la célula transformados para expresar dicho polipéptido.

De acuerdo con una realización más particular al menos un ácido nucleico usado en cualquiera de las realizaciones anteriores comprende una secuencia de nucleótido que se ha obtenido mediante la modificación de SEQ ID NO: 1 y SEQ ID NO: 3. También se desvelan en el presente documento ácidos nucleicos que comprenden una secuencia de nucleótido que se ha obtenido modificando SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas. El al menos un ácido nucleico puede aislarse de una planta de la familia *Winteraceae* o la familia *Canellaceae*, particularmente de *Drimys winteri* o *Drimys lanceolata*.

Se entiende que el organismo o célula "expresa" un polipéptido, con la condición de que el organismo o la célula se transforme para albergar un ácido nucleico que codifica 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 heterológamente" y "sobre-expresar", refiriéndose al último a los niveles de ARNm, el polipéptido y/o la actividad enzimática sobre y por encima de lo que se mide en un organismo o célula no transformados. Una descripción más detallada de los procedimientos adecuados para transformar un organismo hospedador no humano o célula no humana se describirán posteriormente en la parte de la memoria descriptiva que está dedicada a tales células u organismos hospedadores no humanos transformados.

Se entiende que un organismo o célula particular es "capaz de producir FPP" cuando produce FPP de manera natural o cuando no produce FPP de manera 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 las células transformados para producir una mayor cantidad de FPP que el organismo o célula de origen natural, también se abarcan por los "organismos o células capaces de producir FPP". Los procedimientos para transformar los organismos, por ejemplo microorganismos, de modo que éstos produzcan FPP ya se conocen en la técnica.

Para llevar a cabo una realización en el presente documento *in vivo*, el organismo o la célula hospedadores se cultivan en condiciones que conducen a la producción del drimenol. Por consiguiente, si el hospedador es una planta transgénica, se proporcionan las condiciones óptimas de crecimiento, tales como condiciones óptimas de luz, agua y nutrientes, por ejemplo. Si el hospedador es un organismo unicelular, las condiciones que conducen a la producción del drimenol pueden comprender la adición de cofactores adecuados al medio de cultivo del hospedador. Además, puede seleccionarse un medio de cultivo, para elevar al máximo de esta manera la síntesis del drimenol. 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 una realización en el presente documento *in vivo* pueden ser cualquiera de los organismos multicelulares o unicelulares no humanos. En una realización particular, el organismo hospedador no humano usado para llevar a cabo una realización en el presente documento *in vivo* es una planta, un procarionta o un hongo. Puede usarse cualquier planta, procarionta u hongo. Las plantas particularmente útiles son aquellas que producen de manera natural altas cantidades de terpenos. En una realización más particular, el organismo hospedador no humano usado para llevar a cabo el procedimiento de una realización en el presente documento *in vivo*, es un microorganismo. Puede usarse cualquier microorganismo pero de acuerdo con una realización aún más particular dicho microorganismo es una bacteria o levadura. Lo más particularmente, 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 una realización en el presente documento, estos organismos tienen que transformarse para producir dicho FPP. Éstos pueden transformarse de esta manera antes de la modificación con el ácido nucleico descrito de acuerdo con cualquiera de las realizaciones anteriores o simultáneamente, como se explicó anteriormente.

También pueden usarse células eucarióticas superiores aisladas, en vez de los organismos completos, como hospedadores para llevar a cabo el procedimiento de una realización en el presente documento *in vivo*. Las células eucarióticas adecuadas pueden ser cualquier célula no humana, pero son particularmente células vegetales o de hongos.

En todas las realizaciones, el polipéptido consiste en una secuencia de aminoácidos que tiene al menos al menos un 70 %, particularmente al menos un 75 %, particularmente al menos un 80 %, particularmente al menos un 85 %, particularmente al menos un 90 %, particularmente al menos un 95 %, particularmente al menos un 96 %, particularmente al menos un 97 %, particularmente al menos un 98 % y aún más particularmente al menos un 99 % de identidad de secuencia a una secuencia de aminoácidos de SEQ ID NO: 2.

También se desvela en el presente documento un polipéptido que tiene actividad drimenol sintasa que comprende una secuencia de aminoácidos que es una variante de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante la manipulación por ingeniería genética, con la condición de que dicha variante mantenga su actividad drimenol sintasa, como se define anteriormente y tenga el porcentaje requerido de identidad a SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14. En otros términos, dicho polipéptido particularmente comprende una secuencia de aminoácidos codificada por una secuencia nucleotídica que se ha obtenido mediante la modificación de SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas. También se desvela en el presente documento un polipéptido que tiene una actividad drimenol sintasa que consiste

en una secuencia de aminoácidos que es una variante de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante manipulación por ingeniería genética, es decir, una secuencia de aminoácidos codificada por una secuencia de nucleótido que se ha obtenido mediante la modificación de SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas.

También se desvela en el presente documento un polipéptido que tiene una actividad drimenol que es una variante de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 que puede encontrarse de manera natural en otros organismos, tales como otras especies vegetales, con la condición de que mantenga su actividad drimenol sintasa como se define anteriormente y tenga el porcentaje requerido de identidad a las SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14.

Los ejemplos de polipéptidos variantes son proteínas de origen natural que resultan de los eventos de corte y empalme de ARNm alternativos a partir de la escisión proteolítica de los polipéptidos descritos en el presente documento. Las variaciones atribuibles a la proteólisis incluyen, por ejemplo, las diferencias en N- o C-terminal después de la expresión en diferentes tipos de células hospedadoras, debido a la retirada proteolítica de uno o más aminoácidos terminales de los polipéptidos del presente documento. Los polipéptidos codificados por un ácido nucleico obtenido mediante mutación natural o artificial de un ácido nucleico de una realización en el presente documento, como se describe más adelante, también se desvelan en el presente documento.

Las variantes polipeptídicas resultantes de una fusión de secuencias peptídicas adicionales en los amino- y carboxilo-terminal, también pueden usarse en los procedimientos desvelados. 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 o sistema de expresión deseado. Tales secuencias peptídicas adicionales pueden ser péptidos señal, por ejemplo. En consecuencia, se desvelan en el presente documento los procedimientos que usan los polipéptidos variantes, tales como aquellos obtenidos mediante fusión con otros oligo- o polipéptidos y/o aquellos que están enlazados a los péptidos señal. Los polipéptidos que resultan de una fusión con otra proteína funcional, tal como otra proteína proveniente de la vía de la biosíntesis del terpeno, también pueden usarse ventajosamente en los procedimientos descritos anteriormente.

El al menos un polipéptido que tiene una actividad drimenol sintasa usado en cualquiera de las realizaciones anteriormente descritas o codificadas por el ácido nucleico según se desvela puede aislarse de una planta de la familia *Winteraceae* o de la familia *Canellaceae*, particularmente de *Drimys winteri* o *Drimys lanceolata*.

Una herramienta importante para llevar a cabo el procedimiento de una realización en el presente documento es el propio polipéptido. Un polipéptido que tiene una actividad de drimenol sintasa y que comprende una secuencia de aminoácidos al menos un 70 % idéntica a SEQ ID NO:2 se proporciona de esta manera en el presente documento.

El polipéptido puede ser capaz de producir una mezcla de sesquiterpenos en la que 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 el presente documento, el drimenol se produce con más de o igual al 95 %, más particularmente el 98 % de selectividad.

De acuerdo con la invención, el polipéptido comprende una secuencia de aminoácidos al menos un 70 %, particularmente al menos un 75 %, particularmente al menos un 80 %, particularmente al menos un 85 %, particularmente al menos un 90 %, particularmente al menos un 95 %, particularmente al menos un 96 %, particularmente al menos un 97 %, particularmente al menos un 98 % y aún más particularmente al menos un 99 % idéntica a una secuencia seleccionada de SEQ ID NO: 2.

De acuerdo con una realización más particular, el polipéptido comprende una secuencia de aminoácidos de SEQ ID NO: 2.

De acuerdo con otra realización particular, el polipéptido consiste en una secuencia de aminoácidos al menos un 70 %, particularmente al menos un 75 %, particularmente al menos un 80 %, particularmente al menos un 85 %, particularmente al menos un 90 %, particularmente al menos un 95 %, particularmente al menos un 96 %, particularmente al menos un 97 %, particularmente al menos un 98 % y aún más particularmente al menos un 99 % idéntica a una secuencia de SEQ ID NO: 2.

De acuerdo con una realización más particular, el polipéptido consiste en un aminoácido de SEQ ID NO: 2.

También se desvela en el presente documento un polipéptido que comprende una secuencia de aminoácidos que es una variante de SEQ ID NO: 2, ya sea obtenida mediante manipulación por ingeniería genética o encontrada de manera natural en plantas de *Drimys* u otras especies vegetales. En otros términos, cuando el polipéptido variante se obtiene mediante manipulación por ingeniería genética, dicho polipéptido comprende una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se ha obtenido modificando SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas.

También se desvela en el presente documento un polipéptido que tiene una actividad drimenol sintasa que consiste

- 5 en una secuencia de aminoácidos que es una variante de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante manipulación por ingeniería genética, es decir, una secuencia de aminoácidos codificada por una secuencia nucleotídica que se ha obtenido mediante modificando SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de las mismas.
- El polipéptido puede aislarse de una planta de la familia *Winteraceae* o la familia *Canellaceae*, particularmente de *Drimys winteri* o *Drimys lanceolata*.
- 10 Como se mencionó anteriormente, el ácido nucleico que codifica el polipéptido de una realización en el presente documento es una herramienta útil para modificar las células u organismos hospedadores no humanos destinados a usarse cuando el procedimiento se lleva a cabo *in vivo*.
- Un ácido nucleico que codifica un polipéptido que tiene una actividad drimenol sintasa que comprende una secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia con una secuencia de aminoácidos de SEQ ID NO: 2 por lo tanto también se proporciona en el presente documento.
- 15 El ácido nucleico comprende una secuencia nucleotídica al menos un 50 %, particularmente al menos un 55 %, particularmente al menos un 60 %, particularmente al menos un 65 %, particularmente al menos un 70 %, particularmente al menos un 75 %, particularmente al menos un 80 %, particularmente al menos un 85 %, particularmente al menos un 90 %, más particularmente al menos un 95 % particularmente al menos un 96 %, particularmente al menos un 97 %, particularmente al menos un 98 %, y aún más particularmente al menos un 99 % idéntica a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas. De acuerdo con una realización más particular, el ácido nucleico comprende la secuencia nucleotídica seleccionada del grupo que consiste en SEQ ID NO: 1 o SEQ ID NO: 3.
- 20 El ácido nucleico puede consistir en una secuencia nucleotídica al menos un 70 %, particularmente al menos un 75 %, particularmente al menos un 80 %, particularmente al menos un 85 %, particularmente al menos un 90 %, particularmente al menos un 95 %, particularmente al menos un 96 %, particularmente al menos un 97 %, particularmente al menos un 98 % y aún más particularmente al menos un 99 % o más de identidad a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas.
- 25 De acuerdo a una realización incluso más particular, el ácido nucleico consiste en una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas.
- 30 El ácido nucleico de una realización en el presente documento puede definirse incluyendo los polímeros de desoxirribonucleótido o ribonucleótido ya sea en la forma monocatenaria o bicatenaria (ADN y/o ARN). La expresión "secuencia nucleotídica" debe entenderse también que comprende una molécula polinucleotídica o una molécula oligonucleotídica en forma de un fragmento separado o como un componente de un ácido nucleico más grande. Los ácidos nucleicos de una realización en el presente documento también abarcan ciertas secuencias nucleotídicas aisladas que incluyen aquellas que están sustancialmente libres de material endógeno contaminado.
- 35 El ácido nucleico de una realización en el presente documento puede estar presente ya sea de forma natural en las plantas de las especies de *Drimys* u otras especies, o puede obtenerse mediante la modificación de SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas. Particularmente dicho ácido nucleico consiste en una secuencia nucleotídica que se ha obtenido mediante modificación de SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas.
- 40 Los ácidos nucleicos que comprenden una secuencia obtenida por la mutación de SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas están abarcados por una realización en el presente documento, con la condición de que las secuencias que las comprenden comparten al menos el porcentaje definido de identidad con las secuencias correspondientes de SEQ ID NO: 1 o SEQ ID NO: 3 o el complemento de las mismas y con la condición de que codifiquen un polipéptido que tiene una actividad Drimenol sintasa, como se define en cualquiera de las realizaciones anteriores. Las mutaciones pueden ser cualquier tipo de mutaciones de estos ácidos nucleicos, tales como mutaciones puntuales, mutaciones de delección, mutaciones de inserción y/o mutaciones de desplazamiento de marco. Un ácido nucleico variante puede prepararse para adaptar su secuencia nucleotídica a un sistema de expresión específico. Por ejemplo, se sabe que los sistemas de expresión bacterianos expresan de manera más eficiente los polipéptidos si los aminoácidos se codifican por codones particulares.
- 45 Debido a la degeneración del código genético, más de un codón puede codificar la misma secuencia de aminoácidos, múltiples secuencias de ácidos nucleicos pueden codificar la misma proteína o polipéptido. Cuando sea apropiado, las secuencias de ácido nucleico que codifican la drimenol sintasa pueden optimizarse para una expresión amentada en la célula hospedadora.
- 50 Por ejemplo, los nucleótidos de una realización en el presente documento pueden sintetizarse utilizando codones particulares por un hospedador para la expresión mejorada.
- 55 Otra herramienta importante para transformar las células u organismos hospedadores para llevar a cabo el procedimiento de una realización en el presente documento *in vivo* es un vector de expresión que comprende un ácido nucleico de acuerdo con cualquier realización de una realización del presente documento. Un vector tal por lo tanto

también se proporciona en el presente documento.

Los vectores de expresión proporcionados en el presente documento pueden usarse en los procedimientos para preparar un organismo y/o célula hospedadores genéticamente transformados, en células y/u organismos hospedadores que albergan los ácidos nucleicos de una realización en el presente documento y en los procedimientos para la elaboración de polipéptidos que tienen una actividad drimenol sintasa, como se desvela adicionalmente más adelante.

Las células y los organismos hospedadores no humanos recombinantes transformados para albergar al menos un ácido nucleico de una realización en el presente documento de modo que exprese o sobre-exprese heterológicamente al menos un polipéptido de una realización en el presente documento son también herramientas muy útiles para llevar a cabo el procedimiento de una realización en el presente documento. Tales células y organismos hospedadores no humanos se proporcionan también por lo tanto en el presente documento.

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

Los organismos hospedadores no humanos de una realización en el presente documento pueden ser cualquier organismo multicelular o unicelular no humano. En una realización particular, el organismo hospedador no humano es una planta, un procarionta o un hongo. Cualquier planta, procarionta u hongo es adecuado para transformarse de acuerdo con los procedimientos proporcionados en el presente documento. Las plantas particularmente útiles son aquellas que producen de manera natural altas cantidades de terpenos.

En una realización más particular el organismo hospedador no humano es un microorganismo. Cualquier microorganismo es adecuado para su uso el presente documento, pero de acuerdo a una realización aún más particular dicho microorganismo es una bacteria o levadura. Lo más particularmente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

Las células eucarióticas superiores aisladas también pueden transformarse, en vez de organismos completos. Como células eucarióticas superiores, puede entenderse aquí cualquier célula eucariótica no humana excepto las células de levadura. Las células eucarióticas superiores particulares son células vegetales o células de hongos.

Una variante puede también diferir del polipéptido desvelado en el presente documento mediante el acoplamiento de los grupos modificadores que están covalente o no covalentemente enlazados a la cadena principal polipeptídica. La variante también incluye un polipéptido que difiere del polipéptido descrito en el presente documento por los sitios de glucosilación N-enlazados u O-enlazados introducidos y/o una adición de los restos de cisteína. El experto en la materia reconocerá cómo modificar una secuencia de aminoácidos y preservar la actividad biológica. La funcionalidad o la actividad de cualquier proteína, variante o fragmento drimenol sintasa, puede determinarse usando diversos procedimientos. Por ejemplo, la sobreexpresión transitoria o estable en células vegetales, bacterianas o de levadura puede usarse para probar si la proteína tiene o no actividad, es decir, produce el drimenol a partir de FPP. La actividad drimenol sintasa puede evaluarse en un sistema de expresión microbiano, tales como el ensayo descrito en el Ejemplo 2 o 3 en el presente documento sobre la producción del drimenol, indicando funcionalidad. Una variante o derivado de un polipéptido de drimenol sintasa desvelado en el presente documento conserva una capacidad de producir drimenol a partir de FPP. Las variantes de secuencia de aminoácidos de drimenol sintasas desveladas en el presente documento pueden tener funciones biológicas deseables, adicionales, incluyendo, por ejemplo, utilización de sustrato, cinética de reacción, distribución de producto alteradas y otras alteraciones.

Una realización el presente documento proporciona los polipéptidos para su uso en un procedimiento para producir drimenol poniendo en contacto un FPP con los polipéptidos ya sea *in vitro* o *in vivo*.

En el presente documento se desvela también un polinucleótido aislado, recombinante o sintético que codifica un polipéptido o polipéptido variante desvelado en el presente documento.

En una realización, se proporciona una secuencia de ácido nucleico aislada, recombinante o sintética de SEQ ID NO: 1 o SEQ ID NO: 3 que codifica una drimenol sintasa que tiene la secuencia de aminoácidos que es al menos un 70 %, un 75 %, un 80 %, un 85 %, un 90 %, un 92 %, un 95 %, un 96 %, un 97 %, un 98 % o un 99 % idéntica a un aminoácido de SEQ ID NO: 2 que cataliza la producción del drimenol en una célula a partir de FPP. Las realizaciones desveladas en el presente documento incluyen, pero no se limitan a secuencias de ADNc, ADN genómico y ARN. Cualquier secuencia de ácido nucleico que codifique la drimenol sintasa se denomina en el presente documento una secuencia que codifica la drimenol sintasa.

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

Un fragmento de un polinucleótido de SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 se refiere a nucleótidos contiguos que tienen particularmente 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 del presente documento. Particularmente 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. Sin desear quedar limitados, el fragmento de los polinucleótidos en el presente documento puede usarse como un cebador de PCR y/o como una sonda o para el silenciamiento de genes anti-sentido o ARNi.

Está claro para la persona experta en la materia que los genes, incluyendo los polinucleótidos de una realización en el presente documento, pueden clonarse con base en la información de la secuencia de nucleótidos disponible, tal como se encuentra en el listado de secuencias adjunto, por los procedimientos conocidos en la técnica. Éstos incluyen por ejemplo, el diseño de los cebadores de ADN que representan las secuencias flanqueantes de tal gen, de los cuales uno es generado en orientaciones en sentido y que inician la síntesis de la cadena en sentido, y el otro se crea de una manera complementaria inversa, y genera la cadena anti-sentido. Las ADN polimerasas termoestables tales como aquellas usadas en la reacción en cadena de la polimerasa se usan comúnmente para llevar a cabo tales experimentos. Alternativamente, las secuencias de ADN que representan los genes pueden sintetizarse químicamente y posteriormente introducirse en las moléculas de vector de ADN que pueden multiplicarse, por ejemplo, por bacterias compatibles tales como por ejemplo, *E. coli*.

Se desvelan los cebadores y/o sondas de PCR para detectar las secuencias de ácido nucleico que codifican una drimenol sintasa. El experto en la materia estará enterado de los procedimientos 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, con base en las SEQ ID NO: 1 o SEQ ID NO: 3. Un kit de detección para secuencias de ácido nucleico que codifican la drimenol sintasa puede incluir cebadores y/o sondas específicos para las secuencias de ácido nucleico que codifican la drimenol sintasa y un protocolo asociado para el uso de los cebadores y/o las sondas para detectar las secuencias de ácido nucleico que codifican la drimenol sintasa en una muestra. Tales kits de detección pueden usarse para determinar si una planta ha sido o no modificada, es decir, transformada con una secuencia que codifica la drimenol sintasa.

En el presente documento se desvelan las secuencias de ácido nucleico obtenidas por las mutaciones de SEQ ID NO: 1 o SEQ ID NO: 3, tales mutaciones pueden realizarse rutinariamente. Está claro para la persona experta en la materia que mutaciones, deleciones, inserciones, y/o sustituciones de uno o más nucleótidos puedan introducirse dentro de la secuencia de ADN de SEQ ID NO: 1 o SEQ ID NO: 3. En general, una mutación es un cambio en la secuencia de ADN de un gen que puede alterar la secuencia de aminoácidos del polipéptido producido.

Para probar una función de las secuencias de ADN variantes desveladas en el presente documento, la secuencia de interés se enlaza operativamente a un gen marcador seleccionable o rastreable y la expresión del gen indicador se prueba en ensayos de expresión transitorios con protoplastos o en plantas establemente transformadas. El experto en la materia reconocerá que las secuencias de ADN capaces de impulsar la expresión se construyen como módulos. Por consiguiente, los niveles de expresión de fragmentos de ADN más cortos pueden ser diferentes del uno del fragmento más largo y pueden ser diferentes entre sí. Se desvelan además el presente documento también los equivalentes funcionales de la secuencia de ácido nucleico que codifica las proteínas drimenol sintasa, es decir, las secuencias de nucleótidos que se hibridan en condiciones rigurosas a la secuencia de ácido nucleico de SEQ ID NO: 1 o SEQ ID NO: 3. El experto en la materia estará enterado de los procedimientos para identificar las secuencias homólogas en otros organismos y los procedimientos (identificados en la sección de Definiciones en el presente documento), para determinar el porcentaje de identidad secuencial entre secuencias homólogas. Tales moléculas de ADN recién identificadas pueden secuenciarse después y la secuencia puede compararse con la secuencia de ácido nucleico de SEQ ID NO: 1 o SEQ ID NO: 3 y probarse para equivalencia funcional. Se desvelan en el presente documento moléculas de ADN que tienen al menos un 70 %, particularmente un 75 %, particularmente un 80 %, particularmente un 85 %, particularmente un 90 %, particularmente un 95 %, particularmente un 96 %, particularmente un 97 %, particularmente un 98 %, o más particularmente un 99 % o más de identidad secuencial a la secuencia nucleotídica de SEQ ID NO: 1 o SEQ ID NO: 3.

También se desvela una secuencia de ácido nucleico que es complementaria a la secuencia de ácido nucleico de acuerdo con SEQ ID NO: 1 o SEQ ID NO:3, tales como ARN inhibitorios, o la secuencia de ácido nucleico que se hibrida en condiciones rigurosas a al menos parte de la secuencia nucleotídica de acuerdo con NO: 1 o SEQ ID NO: 3.

Se desvela un procedimiento para alterar la expresión génica en una célula hospedadora. Por ejemplo, el polinucleótido de una realización en el presente documento puede aumentarse o sobre-expresarse o inducirse en ciertos contextos (por ejemplo, después de mordeduras o picaduras de insecto o después de la exposición a una cierta temperatura) en una célula hospedadora o un organismo hospedador.

La alteración de la expresión de un polinucleótido desvelado el presente documento también da como resultado la "expresión ectópica" que es un patrón de expresión diferente en un organismo alterado y en un organismo control o de tipo silvestre. La alteración de la expresión se produce a partir de las interacciones del polipéptido con moduladores

exógenos o endógenos, o como resultado de la modificación química del polipéptido. El término también se refiere a un patrón de expresión alterado del polinucleótido de una realización en el presente documento que se altera por debajo del nivel de detección o actividad completamente suprimida.

5 Se desvelan varias secuencias de ácido nucleico que codifican la drimenol sintasa co-expresadas en un único hospedador, particularmente bajo el control de diferentes promotores. Alternativamente, varias secuencias de ácido nucleico que codifican para la proteína drimenol sintasa pueden estar presentes sobre un vector de transformación único o ser co-transformadas al mismo tiempo utilizando vectores separados y seleccionando los transformantes que comprenden ambos genes quiméricos. De manera similar, uno o más genes que codifican para la drimenol sintasa pueden expresarse en una sola planta junto con otros genes quiméricos, por ejemplo que codifican para otras proteínas que aumentan la resistencia a las plagas de insectos, u otros.

10 Las secuencias de ácido nucleico de una realización en el presente documento que codifican las proteínas drimenol sintasa pueden insertarse en los vectores de expresión y/o contenerse en los genes quiméricos insertados en los vectores de expresión, para producir las proteínas drimenol sintasa en una célula hospedadora u organismo hospedador. Los vectores para insertar los transgenes dentro del genoma de las células hospedadoras son bien conocidos en la técnica e incluyen plásmidos, virus, cósmidos y cromosomas artificiales. Los vectores binarios o de co-integración dentro de los cuales se inserta un gen quimérico también se usan para transformar las células hospedadoras.

15 Se desvelan los vectores de expresión recombinantes que comprenden una secuencia de ácido nucleico de un gen de drimenol sintasa o un quimérico que comprende una secuencia de ácido nucleico de un gen de drimenol sintasa, operativamente enlazado a las 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 de SEQ ID NO: 3 puede enlazarse operativamente a una secuencia promotora adecuada para la expresión en células vegetales, células bacterianas o células fúngicas, opcionalmente enlazadas a una secuencia de ácido nucleico no traducida 3'.

20 Alternativamente, la secuencia promotora puede ya estar presente en un vector de modo que la secuencia de ácido nucleico que va a transcribirse se inserta dentro del vector en dirección 3' de la secuencia promotora. Los vectores típicamente se manipulan por ingeniería genética para tener un origen de replicación, un sitio de clonación múltiple y un marcador seleccionable.

Ejemplos

30 Ejemplo 1

Material vegetal de *Drimys lanceolata* y *Drimys winteri* y secuenciación del transcriptoma de hoja

35 Las plantas *Drimys winteri* y *Drimys lanceolata* se obtuvieron de Bluebell Nursery (Leicestershire, RU). Para el análisis de la composición en las moléculas de terpeno, las hojas se recogieron y se extrajeron con disolvente usando MTBE (metil terc-butil éter). El extracto se analizó por GCMS usando un sistema GC Agilent Serie 6890, conectado a un detector de masa Agilent 5975. La GC estuvo equipada con un diámetro interno de 0,25 mm por 30 m de columna capilar DB-1 ms (Agilent). El gas vehículo fue He a un flujo constante de 1 ml/min. La temperatura inicial del horno fue de 50 °C (retención de 1 min) seguida de un gradiente de 10 °C/min a 300 °C. La inyección se realizó en un inyector dividido/no dividido ajustado a 260 °C y usado en el modo sin división. La identificación de los productos estuvo basada en la comparación de los espectros de masa y los índices de retención con los patrones auténticos y las bases de datos internas de los espectros de masa. Las hojas de las dos plantas contenían cantidades significativas de los compuestos drimano-sesquiterpeno incluyendo (-)-drimenol, poligodial y epipoligodial (Figura 1).

40 De este modo se tomaron hojas pequeñas de *D. winteri* y *D. lanceolata* para el análisis del transcriptoma. El ARN total se extrajo usando el Reactivo de ARN Vegetal Concert^{MR} (Invitrogen). Este ARN total se procesó usando la técnica Illumina Total ARN-Seq y el secuenciador Illumina HiSeq 2000. Un total de 101 y 105 millones de lecturas apareadas de 2 x 100 pb se generaron para *D. winteri* y *D. lanceolata*, respectivamente. Las lecturas se ensamblaron usando el ensamblador genómico *de novo* Velvet (<http://www.ebi.ac.uk/~zerbino/velvet/>) y el software Oases (<http://www.ebi.ac.uk/~zerbino/oases/>). Para *D. winteri* se ensamblaron 40.586 cóntigos con un tamaño promedio de 1.080 pb y para *D. lanceolata* se ensamblaron 28.255 cóntigos con un tamaño promedio de 1.179 pb. Los cóntigos se buscaron usando el algoritmo tBlastn (Altschul y col, J. Mol. Biol. 215, 403-410, 1990) y usando como búsqueda las secuencias de aminoácidos de las sesquiterpeno sintasas conocidas. Este enfoque proporcionó las secuencias para 37 sesquiterpeno sintasas putativas nuevas. La actividad enzimática de estas sintasas se evaluó como se describe en los siguientes ejemplos para las sintasas que muestran actividad drimenol sintasa.

Ejemplo 2. Expresión y caracterización funcionales de DITps589 de *D. lanceolata*

45 Las secuencias de ADN de una de las sesquiterpeno sintasas seleccionadas DITps589 se optimizaron por codón, se sintetizaron *in vitro* y se clonaron en el plásmido de expresión pJ444-SR (DNA2.0, Menlo Park, CA, EE.UU.).

La expresión heteróloga de las sintasas DITps589 se realizó en células de *E. coli* KRX (Promega). Se usaron colonias

únicas de células transformadas con el plásmido de expresión pJ444SR-DITps589 para inocular 5 ml de medio LB. Después de 5 a 6 horas de incubación a 37 °C, los cultivos se transfirieron a una incubadora a 20 °C y se dejaron 1 hora para el equilibrado. La expresión de la proteína se indujo después por la adición de IPTG 1 mM y 0,2 % de L-ramnosa y el cultivo se incubó toda la noche a 20 °C. Al siguiente día, las células se recogieron por centrifugación, se resuspendieron en 0,1 volúmenes de MOPSO 50 mM, pH 7, 10 % de glicerol y se lisaron sometiendo a ultrasonidos. 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.

Los extractos brutos de proteína de *E. coli* que contenían la proteína recombinante se usaron para la caracterización de las actividades enzimáticas. Los ensayos se realizaron en 2 ml de MOPSO 50 mM, pH 7, 10 % de glicerol, DTT 1 mM, MgCl₂ 15 mM en presencia de 80 μM de difosfato de farnesilo (FPP, Sigma) y 0,1 a 0,5 mg de proteína bruta. Los tubos se incubaron de 12 a 24 horas a 30 °C y se extrajeron dos veces con un volumen de pentano. Después de la concentración en un flujo de nitrógeno, los extractos se analizaron mediante GC y GC-MS y se compararon con los extractos provenientes de los ensayos con las proteínas control. El análisis de los productos formados por las enzimas se realizó mediante GCMS como se describe en el ejemplo 1. Se realizó un control negativo en las mismas condiciones usando células de *E. coli* transformadas con un plásmido pJ444 vacío. En estas condiciones, la enzima recombinante DITps589 produjo (-)-drimenol como un producto mayor con una selectividad superior al 98 % (Figura 2). La identidad del (-)-drimenol se confirmó mediante coincidencia del espectro de masa y el tiempo de retención de un patrón de drimenol auténtico, aislado de Aceite de Madera de Sándalo Occidental (*Amyris balsamifera*).

Ejemplo 3. Producción *in vivo* de (-)-drimenol en células de *E. coli* usando DITps589

Para evaluar la producción *in-vivo* de (-)-drimenol en células heterólogas, las células de *E. coli* se transformaron con el plásmido de expresión pJ444SR-DITps589 y se evaluó la producción de los sesquiterpenos a partir del combinado endógeno de FPP. Para aumentar la productividad de las células, también se expresaron en las mismas células una FPP sintasa heteróloga y una de las enzimas provenientes de la ruta del mevalonato (MVA) heteróloga completa. La construcción del plásmido de expresión que contiene un gen de la FPP sintasa y el gen para una ruta de MVA completa se describió en la patente WO2013064411 o en Schalk y col (2013) J. Am. Chem. Soc. 134, 18900-18903. En resumen, se preparó un plásmido de expresión que contenía dos operones compuestos por los genes que codifican las enzimas para una ruta completa del mevalonato. Se sintetizó *in vitro* un primer operón sintético que consistía en genes de una acetacetil-CoA tiolasa (atoB) de *E. coli*, una HMG-CoA sintasa (mvaS) de *Staphylococcus aureus*, una HMG-CoA reductasa (mvaA) de *Staphylococcus aureus* y una FPP sintasa (ERG20) de *Saccharomyces cerevisiae* (ADN2.0, Menlo Park, CA, EE.UU.) y se ligó en el vector pACYCDuet-1 digerido con *NcoI*-*Bam*HI (Invitrogen) produciendo pACYC-29258. Un segundo operón que contiene un mevalonata quinasa (MvaK1), una fosfomevalonato quinasa (MvaK2), una mevalonato difosfato descarboxilasa (MvaD) y una isopentenil difosfato isomerasa (idi) se amplificó a partir del ADN genómico de *Streptococcus pneumoniae* (ATCC BAA-334) y se ligó dentro segundo sitio de clonación múltiple de pACYC-29258 proporcionando el plásmido pACYC-29258-4506. Este plásmido contiene de este modo los genes que codifican todas las enzimas de la ruta biosintética que conduce desde la acetil-coenzima A hasta FPP.

Células de *E. coli* KRX (Promega) se co-transformaron con el plásmido pACYC-29258-4506 y el plásmido pJ444SR-DITps589. Las células transformadas se seleccionaron sobre placas de LB-agarosa con carbenicilina (50 μg/ml) y cloranfenicol (34 μg/ml). Se usaron colonias únicas para inocular 5 ml del medio LB líquido suplementado con los mismos antibióticos. El cultivo se incubó toda la noche a 37 °C. Al siguiente día se inocularon 2 ml de medio TB suplementado con los mismos antibióticos con 0,2 ml del cultivo toda la noche. Después de 6 horas de incubación a 37 °C, el cultivo se enfrió hasta 28 °C y se añadieron IPTG 0,1 mM y 0,2 % de ramnosa a cada tubo. Los cultivos se incubaron durante 48 horas a 28 °C. Los cultivos se extrajeron después dos veces con 2 volúmenes de MTBE, la fase orgánica se concentró a 500 μL y se analizó mediante GC-MS como se describe anteriormente en el Ejemplo 1.

En estas condiciones *in-vivo*, la enzima recombinante DITps589 produjo (-)-drimenol como producto principal con la misma selectividad aparente que en el ensayo *in vitro* descrito en el ejemplo 2 (Figura 3). Usando estas células de *E. Coli* manipuladas por ingeniería genética se usaron cultivos más grandes (1 l) para purificar el sesquiterpeno producido por la enzima en cantidad suficiente para confirmar la estructura mediante análisis RMN y la medición de la rotación específica siendo la estructura del (-)-drimenol mostrada en la figura 4. La enantiopureza se confirmó mediante análisis de GC quiral sobre un sistema de GC Varian CP-3800 equipado con una columna ChiraSil (Agilent) y usando temperatura de gradiente de horno de 3,0 °C/min desde 125 hasta 180 °C (Figura 5).

Ejemplo 4 (Ejemplo de Referencia). Expresión y caracterización funcionales de SCH51-3228-9 y SCH51-3228-11 de *D. winteri*.

SCH51-3228-9 y SCH51-3228-11 son otras dos secuencias de ADN que codifican putativamente las sesquiterpeno sintasas y se aislaron a partir de las secuencias de transcriptoma de *Drymis winteri*. Las secuencias de aminoácidos deducidas comparten un 92,6 y un 95,1 % de identidad, respectivamente, con la secuencia de aminoácidos DITps589. Las dos secuencias se optimizaron en codón, se sintetizaron *in vitro* (Invitrogen) y se clonaron entre los sitios de reconocimiento de enzimas de restricción *NdeI* y *Kpml* del plásmido de expresión (Invitrogen) pETDuet-1 (Novagen).

La expresión heteróloga de las sintasas SCH51-3228-9 y SCH51-3228-11 se realizó en células de *E. coli* BL21 (DE3) (Invitrogen). Se usaron colonias únicas de células transformadas con los plásmidos de expresión pETDuet-SCH51-

5 3228-9 o el pEDTDuet-SCH51-3228-11 para producir las enzimas recombinantes como se describe en el ejemplo 2. Los extractos de proteína bruta de *E coli* que contenía las proteínas recombinantes se usaron para la caracterización de las actividades enzimáticas como se describe en el ejemplo 2 excepto por las condiciones de análisis de GCMS que se realizaron como sigue. El análisis de GCMS se realizó sobre un sistema de GC Agilent de la Serie 6890 conectado a un detector de masa Agilent 5975. La GC se equipó con una columna capilar DB-1ms de un diámetro interno de 0,25 mm por 30 m (Agilent). El gas vehículo fue helio a un flujo constante de 1 ml/min. La temperatura inicial del horno fue de 50 °C (retención de 5 minutos) seguido de un gradiente de 5 °C/min a 300 °C. La inyección se realizó en el modo dividido a 250 °C con una relación de división de 5:1.

10 Las dos enzimas recombinantes produjeron (-)-drimenol como el producto principal con alta selectividad. La identidad del (-)-drimenol se confirmó mediante comparación del espectro de masas y el tiempo de retención de un patrón de drimenol auténtico aislado de Aceite de Madera de Sándalo Occidental (*Amyris balsamifera*) (Figura 6).

Usando el sistema de células de *E. coli* completas y el procedimiento descrito en el ejemplo 3 (excepto para las condiciones de análisis de GCMS que fueron como se describe anteriormente) el drimenol pudo producirse también en cultivos bacterianos *in vivo* usando las proteínas recombinantes SCH51-3228-9 y SCH51-3228-11 (Figura 7).

15 **Ejemplo 5 (Ejemplo de Referencia). Expresión y caracterización funcionales de SCH51-998-28 de *D. winteri* y SCH52-13163-6 de *D. lanceolata*.**

De manera similar al ejemplo 4, los dos ADNc SCH51-998-28 y SCH52-13163-6 se optimizaron y se clonaron en el plásmido de expresión pETDuet.

20 Las proteínas recombinantes se produjeron en células de *E. coli* BL21 (DE3) (Invitrogen) y los ensayos *in vitro* usando FPP como sustrato se realizaron como se describe en los ejemplos 2 y 4. Estos ensayos mostraron actividad drimenol sintasa para SCH51-998-28 y SCH52-13163-6 (figura 6). Usando células de *E. coli* que sobre-expresan FPP a partir de una ruta de mevalonato recombinante (ejemplos 2 y 4), el drimenol también pudo producirse *in vivo* usando las proteínas SCH51-998-28 y SCH52-13163-6 (Figura 7).

Ejemplo 6. Comparación de secuencias de las drimenol sintasas de especies de *Drimys*.

25 Las secuencias de aminoácidos de las drimenol sintasas de *Drimys winteri* y *Drimys lanceolata* se alinearon usando el programa de alineamiento Múltiple ClustalW (Thompson y col, 1994, Nucleic Acid Res. 22(22), 4673-4680) y las identidades de secuencia se calcularon a base de esta alineación. Porcentaje de identidad (%) entre las diferentes drimenol sintasas de especies de *Drimys*:

	DITps589	SCH51-3228-11	SCH51-3228-9	SCH51-998-28	SCH52-13163-6
DITps589	ID	95,1	92,6	70,5	88
SCH51_3228_11	95,1	ID	97,1	70,6	87,6
SCH51_3228_9	92,6	97	ID	71	90,1
SCH51_998_28	70,5	70,6	0,71	ID	72,5
SCH52_13163_6	88	87,6	90,1	72,5	ID

30 - Listado de Secuencias -

SEQ ID NO: 1
SCH52-ctg589, Marco de lectura abierto de CITps589 de *D. lanceolata*, secuencia de ADN de tipo silvestre.

ES 2 759 724 T3

ATGGATCTTATTAATCCCTCCCCAGCGGCTTCCACCCTCCCTCTCCCAGTTGATGGAGATT
AGAAGTTGTTAGGCGATCTGCCGGGTTTCATCCGACTATCTGGGGCGATCACTTCCTCTCCT
ACAAGCCCGATCCAAAGAAAATAGATGCATGGAATAAAAGGGTTGAAGAGCTGAAGGAAG
AAGTGAAGAAGATATTAAGCAATGCAAAAGGGACGGTGAAGAGCTGAATTTGATTGATG
ATCTCGTACACCTTGGGATTAGTTATCATTTTGAGAAGGAGATTGATGATGCTCTACAACAC
ATCTTTGATACCCATCTTIGATGATTTTCCTAAGGATGATCTATATGTCGCCGCCTCCGATTT
GGCGTCTTAAGGAAACAGGGGCACCGTGTCTCCAGATGTATTCAAAAAATTCAAAGATG
AGCAGGGGAATTTCAAGGCAGAGTTGAGCACCGATGCGAAAGGTTTGCTATGTTTAAATGA
TGTGGCTTATCTCAGCACAAGAGGGGAAGATATCTTGGATGAAGCCATTCCTTTCACTGAG
GAGCACCTTAGGTCTTGTATTAGCCATGTAGATTCTCATATGGCAGCAAAAATTGAACATTC
TCTCGAGCTTCCCCTTCATCATCGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCT
ATGAAGGAGACAAGGAAAGGAACGAAGTTGTCCTTGAGCTTGCCAATTTAGATTTCAATCT
GATTCAAATCTTGCACCAAAGAGAGCTGAGAGACATCACAATGTGGTGAAGGAGATTGA
CCTTGCAGCAAAGCTGCCTTTTATTAGGGATAGGTTGGTGGAGTGCTACTACTGGATCATG
GGGGTCTATTTTGAACCAATATACTCGAGGGCTAGGGTTTTTTCCACCAAATGACAATGTT
GGTCTCAGTTGTGGACGACATATATGATGTGTATGCTACCGAGGATGAGCTTCAACTATTC
ACTGATGCCATCTATAGGTGGGATGCTGATGACATTGATCAGCTGCCTCAGTACTTGAAAG
ATGCTTTTATGGTACTCTACAACACTGTGAAGACTCTAGAAGAAGAAGCTTGAACCAGAAGG
AAACTCTTATCGTGGATTCTATGTAAAAGATGCAATGAAGGTTTTGGCAAGGGATTACTTT
GTGGAGCACAAATGGTATAACAGAAAAATTGTGCCATCCGTAGAGGAATACTTGAAAATTT
CTTGCATCAGTGTGGCCGTTTCATATGGCTACAGTTCAGTGTATTGCTGGGATGTATGAAATT
GCAACCAAAGAGGCATTCGAATGGTTGATGACTGAGCCCAAAGCTTGTATTGATGCATCTC
TGATTGGTCGTCTCCTTGATGACATGCAGTCCACCTCGTTTGAGCAACAGAGAGGCCACGT
GTCATCAGCAGTACAGTGTACATGGCTGAATATGGTGTAAACAGCGGAAGAAGCATGTGAA
AAGCTCCGAGATATGGCTGCAATTGCTTGGAAAGATGTGAACGAGGCATGCCTTAGGCCCA
CGGTTTTCCCTATGCCTATCCTTTTGCCTTCTATCAACTGGCACGTGTGGCAGAAGTCATCT
ACCTACGTGGAGATGGATACACGCACGCTGGGGGTGAGACCAAGAAACACATCACGGCCA
TGCTTGTTAAGCCAATTGAAGTCTGA

SEQ ID NO: 2

CITps589 de *D. lanceolata*, secuencia de aminoácidos.

ES 2 759 724 T3

MDLINPSAASTLPLPVDGDSEVRRSAGFHPTIWGDHFLSYKDPKIDAWNKRVEELKEEVK
KILSNAKGTVEELNLIDDLVHLGISYHFEKEIDDALQHIFDTHLDDFPKDDLIVAAALRFGVLRKQ
GHRVSPDVFKKFKDEQGNFKAELSTDAKGLLCLNDVAYLSTRGEDILDEAIPFTEEHLRSCISHV
DSHMAAKIEHSLELPLHHRIPRENRYHISVYEGDKERNEVVLELANLDFNLIQILHQRELRDIT
MWWKEIDLAAKLPFIRDRLVECYWIMGVYFEPIYSRARVFTSKMTMLVSVVDDIYDVYATE
DELQLFTDAIYRWDADDIDQLPQYLKDAFMVLYNTVKTLLEEELEPEGNSYRGFYVKDAMKVL
ARDYFVEHKWYNRKIVPSVEEYKISCISVAVHMATVHCIAGMYEIATKEAFEWLMTEPKLVID
ASLIGRLLDDMQSTSFEQQRGHVSSAVQCYMAEYGVTAEEACEKLRDMAAIAWKDVNEACL
PTVFPMPILLPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVKPIEV

SEQ ID NO: 3

DITps589_opt, Secuencia de ADN optimizada en codón de DITps589 de *D. lanceolata*.

ATGGACCTGATTAACCCGAGCCCTGCTGCATCCACCCTGCCACTGCCAGTCGATGGTGATA
GCGAAGTTGTGCGCCGTAGCGCGGGTTTCCATCCGACCATCTGGGGTGACCACTTTCTGTCT
TATAAGCCGGACCCGAAAAAGATTGATGCGTGGAACAAGCGTGTTGAGGAACTGAAAGAA
GAGGTCAAAAAGATTTTGAGCAATGCGAAAGGCACGGTTGAGGAACTGAATTTGATTGAC
GACCTGGTACACCTGGGTATTAGCTATCACTTTGAGAAAGAAATCGACGACGCGCTGCAGC
ATATCTTCGATACGCACCTGGATGATTTCCCGAAAGATGACCTCTACGTGGCTGCGCTGCGT
TTGGCGTCCTGCGTAAGCAAGGCCATCGTGTGACCCCGACGTCTTTAAGAAATTCAAAG
ACGAGCAAGGCAACTTCAAAGCGGAGCTGTCAACCGATGCAAAGGGCCTGTTGTGCCTGA
ACGATGTGGCGTACCTGAGCACCCGTGGTGAGGATATCCTGGACGAAGCGATCCCGTTCAC
GGAAGAACATTTGCGCTCGTGCATTAGCCACGTTGATAGCCACATGGCAGCGAAGATTGAG
CACTCTCTGGAGCTGCCGCTGCACCATCGCATTCGCGTTTAGAGAATCGCCATTACATCTC
CGTGTACGAGGGTGACAAAGAGCGTAATGAAGTCGTTCTGGAGTTGGCTAACTTGGACTTT
AATCTTATCCAGATCCTGCACCAGCGGAGCTGCGCGACATCACGATGTGGTGGAAAGAAA
TTGATCTGGCCGCAAAGCTGCCGTTTATTCGTGACCGTCTGGTGGAGTGTTACTATTGGATT

ES 2 759 724 T3

ATGGGCGTGTA CTTTCGAGCCGATCTACAGCCGTGCGCGCGTGTTTAGCACCAAGATGACCA
TGCTGGTTAGCGTGGTGGATGACATCTATGATGTCTACGCTACGGAAGATGAGTTGCAGCT
GTTTACCGACGCCATTTACAGATGGGACGCCGATGACATTGATCAACTGCCGCAATATCTG
AAAGACGCCTTTATGGTTCTGTACAACACCGTCAAAACCTGGAAGAAGAAGTGGAGCCGG
AAGGTAACCTTATCGTGGTTTCTACGTTAAAGATGCGATGAAAGTTCTGGCGCGTGACTAT
TTCGTTGAGCATAAGTGGTACAATCGTAAGATCGTCCCGTCCGTTGAAGAGTACTTGAAGA
TTAGCTGTATCAGCGTCGCAGTCCACATGGCGACCGTGCACTGTATCGCCGGCATGTATGA
GATCGCCACGAAAGAAGCATTTCGAGTGGCTGATGACCGAGCCGAAACTGGTGATTGACGC
AAGCCTGATTGGTCGCCTGCTGGACGATATGCAGAGCACGAGCTTTGAGCAGCAGCGCGGT
CATGTTAGCTCCGCAGTTCAATGCTACATGGCTGAGTACGGTGTGACTGCCGAAGAAGCAT
GCGAGAAGCTGCGTGATATGGCGGCCATTGCGTGGAAAGATGTGAATGAAGCATGCCTGC
GCCCCACCGTTTTCCCGATGCCGATTTTACTGCCTAGCATCAACCTGGCACGTGTGGCGGA
AGTTATCTATCTGCGTGGCGACGGTTATACGCACGCGGGTGGTGAGACTAAGAAGCACATC
ACCGCGATGCTGGTCAAGCCGATCGAAGTGTA

SEQ ID NO: 4

SCH51-3228-9 de *D. winteri*, Marco de lectura abierto, secuencia de ADN de tipo silvestre.

ATGGCTTCCACCCTCCCTCTCCCAGCTTATGGAGATTCAGAAGTTGTTAGGCGATCTGCCGG
GTTTCATCCGACGATCTGGGGCGATCACTTCCCTCCTACAAGCCTGATCCAACGAAAATA
GATGAATGGAATAAAAGGGTTGAAGAGCTGAAGGAAGAAGTGAAGAAGATATTAAGCAAT
GCAAAAGGGACAGTGGAAGAGCTGAATTTGCTTGATGATCTCGTACACCTTGGGATTAGTT
ATCATTTTGAGAAGGAGATTGATGATGCTTTACAACAAATCTTTGATACCCATCTTGATGTT
TTTCTAAGGATGATCTATATGCCACCGCTCTCCGATTTGGCGTCTTAAGGAAACAGGGGC
ACCGTGTTTCTCCAGATGTATTCAAAAAATTCAAAGATGAGCAGGGGAATTTCAAGGCAGA
GTTGAGCACCGATGCGAAGGGTTTGCTATGTTTATATGATGTGGCTTATCTCAGCACAAGA
GGGGAAGATATCTTGATGAAGCCATTCCTTTACTAAGGAGCACCTTAGGTCTTGTATTA
GCCATGTCGATTCTCATATGGCAGCAAAAATTGAGCATTCTCTAGAGCTTCCCCTTCATCAT
CGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCTATGAAGGAGACAAGGAAAGG
AATGAAGTTGTCCTTGAGCTTGCCAAATTAGATTTCAATCTGATTCAAATCTTGCACCAAAG
AGAGCTGAGGGACATCACAACGTGGTGGAAAGGAGATTGACCTTGACGCAAAAGCTACCTTTT

ES 2 759 724 T3

ATTAGGGATAGGTTGGTGGAGTGCTACTATTGGATCATGGGAGTCTATTTTGAACCAATAT
ACTCAAGGGCTAGAGTTTTTTCGACCAAAATGACAATCTTGGTCTCAGTTGTGGACGACAT
ATATGATGTATATGCTACAGAGGATGAGCTCCAACTTTTCACTGATGCAATCTATAGGTGG
GATGCTGAGGACATTGAGCAGCTTCCACAGTACTTCAAAGATGCTTTTCTTGTACTCTATAA
CACTGTGAAGGACCTAGAAGAGGAATTGGAACCAGAAGGAAACTCTTATCGTGGATACTA
TGAAAAGATGCGATGAAGGTTTTGGCAAGGGATTACTTTGTGGAGCACAAATGGTATAAC
AGAAAAATTGTGCCATCAGTAGAGGACTACCTGCGAATTTCTTGCATTAGTGTGCCGTTC
TATGGCCACAGTTCATTGTATTGCTGGGATGTATGAAATTGCAACCAAAGAGGCATTCGAA
TGGTTGAAGACGGAACCTAACTTGTATAGATGCATCACTGATTGGGCGTCTCCTCGATG
ACATGCAGTCCACCTCGTTTGTAGCAACAGAGAGGTCATGTGTCATCAGCGGTACAGTGTTA
CATGATCCAATATGGGGTATCACACGAAGAAGCGTGTGAGAAGTTGCGAGAAATGGCTGC
AATTGCGTGGAAGATGTAAACCAAGCATGCCTTAGGCCCACTGTTTTCCCTATGCCTATTC
TTCTGCCCTCCATCAACCTTGCACGTGTGGCAGAAGTGATTTACCTACGCGGAGATGGATAT
ACACATGCGGGTGGTGTAGACCAAAAAACATATCACGGCCATGCTTGTGATCCAATCAAAG
TCTGA

SEQ ID NO: 5
SCH51-3228-9 de *D. winteri*, secuencia de aminoácidos.

MASTLPLPAYGDSEVVRRSAGFHPTIWGDHFLSYKPDPTKIDWKNRVEELKEEVKKILSNAKG
TVEELNLLDDLVLHLSYHFEKEIDDALQQIFDTHLDVFPKDDLYATALRFGVLRKQGHVSPD
VFKKFKDEQGNFKAELSTDAKGLLCLYDVAYLSTRGEDILDEAIPFTKEHLRSCISHVDSHMAA
KIEHSLELPLHHRIPRENRHYISVYEGDKERNEVVLELAKLDFNLIQILHQRELDRITWWKEID
LAAKLPFIRDRLVECYWIMGVYFEPIYSRARVFSTKMTILVSVVDDIYDVYATEDELQLFTDAI
YRWDAEDIEQLPQYLKDAFLVLYNTVKDLEEELEPEGNSYRGYYVKDAMKVLARDYFVEHK
WYNRKIVPSVEDYLRISCISVAVHMATVHCCAGMDEIATKEAFEWLKTEPKLVIDASLIGRLLD
DMQSTSFEQQRGHVSSAVQCYMIQYGVSHHEACEKLREMAAIAWKDVNQA CLRPTVFPMPIL
LPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVDPIKV

5 SEQ ID NO: 6
SCH51-3228-9_opt, Secuencia de ADN optimizada en codón de SCH51-3228-9.

ES 2 759 724 T3

ATGGCAAGCACCCCTGCCGCTGCCTGCCTATGGTGATAGCGAAGTTGTTCGTCGTAGCGCAG
GTTTTTCATCCGACCATTGTTGGGGTGATCATTTTCTGAGCTATAAACCGGATCCGACCAAATT
GATGAATGGAATAAACGTGTCTGAAGAAGTGAAGAAGAAGTGAAAAAATCCTGAGCAAT
GCCAAAGGCACCGTTGAGGAACTGAATCTGCTGGATGATCTGGTTCATCTGGGTATCAGCT
ATCACTTTGAGAAAGAAATCGATGATGCACTGCAGCAGATTTTTGATACCCATCTGGATGT
TTTCCCAGAAAGATGATCTGTATGCAACCGCACTGCGTTTTTGGTGTCTGCGTAAACAGGGTC
ATCGTGTTAGTCCGGATGTGTTCAAAAAATTCAAAGATGAACAGGGCAACTTCAAAGCAGA
ACTGAGCACCGATGCAAAAGGTCTGCTGTGTCTGTATGATGTTGCATATCTGAGCACCCGT
GGTGAAGATATTCTGGATGAAGCAATTCCGTTTACCAAAGAACATCTGCGTAGCTGTATTA
GCCATGTTGATAGCCACATGGCAGCGAAAATGAACATAGCCTGGAAGTGCCTCTGCATCA
CCGTATTCGCGTCTGGAAAATCGTCACTATATTAGCGTTTATGAGGGCGATAAAGAACGC
AATGAAGTTGTGCTGGAAGTGGCAAACTGGATTTTAACCTGATTCAGATTCTGCATCAGC
GTGAAGTGCCTGATATTACCACCTGGTGGAAAGAAATTGATCTGGCAGCAAACTGCCGTT
TATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTATGGGCGTGTATTTTGAACCGATT
ATAGCCGTGCACGTGTTTTTAGCACCAAATGACCATTCTGGTTAGCGTGGTGGATGATATC
TATGATGTTTATGCCACCGAAGATGAACTGCAGCTGTTTACCGATGCCATTTATCGTTGGGA
TGCAGAAGATATTGAACAGCTGCCGCAGTATCTGAAAGATGCATTTCTGGTCTGTACAAC
ACCGTGAAAGATCTGGAAAGAAGAACTGGAACCGGAAGGTAATAGCTATCGTGGTTATTAT
GTAAAGATGCCATGAAAGTTCTGGCACGCGATTATTTTGTGAGCACAAATGGTATAACC
GCAAAATTGTCCGAGCGTGGAAAGATTATCTGCGTATTAGCTGCATTAGCGTTGCAGTTCAC
ATGGCAACCGTTCATTGTTGTGCAGGTATGGATGAAATTGCAACCAAAGAAGCATTGAGT
GGCTGAAAACCGAACCGAACTGGTTATTGATGCAAGCCTGATTGGTCGTCTGCTGGACGA
TATGCAGAGCACCGCTTTGAACAGCAGCGTGGTCATGTTAGCAGCGCAGTTCAGTGTTAT
ATGATTCAGTATGGTGTTAGCCATGAAGAAGCATGCGAAAACTGCGCGAAATGGCAGCA
ATTGCATGGAAAGATGTTAATCAGGCATGCTGCGTCCGACCGTTTTTCCGATGCCGATTCT
GCTGCCGAGCATTAACTGGCACGTGTTGCCGAAGTTATCTATCTGCGTGGTGTGTTATA
CCCATGCCGGTGGTGAACCAAAAAACATATTACCGCAATGCTGGTTCGATCCGATTAAGT
TTAA

SEQ ID NO: 7

SCH51-3228-11 de *D. winteri*, Marco de lectura abierto, secuencia de ADN de tipo silvestre.

ES 2 759 724 T3

ATGGCTTCCACCCTCCCTCTCCCAGCTTATGGAGATTCAGAAGTTGTTAGGCGATCTGCCGG
GTTTCATCCGACGATCTGGGGCGATCACTTCTCTCCTACAAGCCTGATCCAACGAAAATA
GATGAATGGAATAAAAGGGTTGAAGAGCTGAAGGAAGAAGTGAAGAAGATATTAAGCAAT
GCAAAAGGGACAGTGGAAGAGCTGAATTTGCTTGATGATCTCGTACACCTTGGGATTAGTT
ATCATTTTGAGAAGGAGATTGATGATGCTTTACAACAAATCTTTGATACCCATCTTGATGTT
TTTCCTAAGGATGATCTATATGCCACCGCTCTCCGATTTGGCGTCTTAAGGAAACAGGGGC
ACCGTGTCTTCCAGATGTATTCAAAAAATTCAAAGATGAGCAGGGGAATTTCAAGGCAGA
GTTGAGCACCGATGCGAAGGGTTTGCTATGTTTATATGATGTGGCTTATCTCAGCACAAGA
GGGAAGATATCTTGGATGAAGCCATTCTTTCACTAAGGAGCACCTTAGGTCTTGTATTA
GCCATGTCGATTCTCATATGGCAGCAAAAATTGAGCATTCTCTAGAGCTTCCCCTTCATCAT
CGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCTATGAAGGAGACAAGGAAAGG
AATGAAGTTGTCCTTGAGCTTGCCAAATTAGATTTCAATCTGATTCAAATCTTGCACCAAAG
AGAGCTGAGGGACATCACAATGTGGTGGAAAGGAGATTGACCTTGCAGCAAAGCTACCTTTT
ATTAGAGATAGGTTGGTGGAGTGCTACTACTGGATCATGGGGGTCTATTTTGAACCAATAT
ACTCCAGGGCTAGGGTTTTTTTCCACTAAAATGACAATCTTGGTCTCAGTTGTGGACGACATA
TATGATGTCTATGCTACGGAGGATGAGCTTCAACTATTCCTGATGCAATCTATAGGTGGG
ATGCTGATGACATTGATCAGCTGCCTCAGTACTTGAAAGATGCTTTTATGGTACTCTATAAC
ACTGTGAAGACTCTAGAAGAAGAACTTGAACCAGAAGGAAACTCTTATCGTGGATACTACG
TAAAAGATGCAATGAAGGTTTTGGCAAGAGATTACTTTGTGGAACACAAATGGTATAACAG
ACAAATTGTGCCATCCGTAGAGGAATACTTGA AAAATTTCTTGCATTAGTGTGGCTGTTCATA
TGGCTACAGTTCATTGTATTGCTGGGATGTATGAAATTGCTACCAAAGAGGCATTCGAATG
GTTGAAGACTGAACCCAAACTTGTTATCGATGCATCTCTGATCGGTCGTCTTCTTGATGACA
TGCAGTCTACCTCGTTTGAGCAACAAAGAGGGCACGTGTCATCAGCAGTACAGTGTTACAT
GGCCCAATATGGAGTAACAGCAGAAGAAGCATGTGAAAAGCTACGAGAAATGGCTGCAAT
TGCTTGAAAGATGTGAATGAAGCATGCCTTAGGCCACGGTATTCCCTATGCCTATCCTCT
TGCCTTCTATCAACTTGGCACGTGTGGCAGAAGTGATCTACCTACGTGGAGATGGATACAC
GCACGCTGGGGGTGAGACCAAAAAACACATCACGGCCATGCTTGTTAAGCCAATTGAAGTC
TGA

SEQ ID NO: 8
SCH51-3228-11 de *D. winteri*, secuencia de aminoácidos.

ES 2 759 724 T3

MASTLPLPAYGDSEVVRRSAGFHPTIWGDHFLSYKPDPTKIDEWNKRVEELKEEVKKILSNAKG
TVEELNLLDDLVLHGLISYHFEKEIDDALQIFDTHLDVFPKDDLYATALRFGVLRKQGHVSPD
VFKKFKDEQGNFKAELSTDAKGLLCLYDVAYLSTRGEDILDEAIPFTKEHLRSCISHVDSHMAA
KIEHSLELPLHHRIPLENRHYISVYEGDKERNEVVLELAKLDFNLIQILHQRELRDITMWWKEI
DLAAKLPFIRDRLVECYWIMGVYFEPIYSRARVFSTKMTILVSVVDDIYDVYATEDELQLFTD
AIYRWDADDIDQLPQYLKDAFMVLYNTVKTLEEELEPEGNSYRGYYVKDAMKVLARDYFVEH
KWYNRQIVPSVEEYLKISCISVAVHMATVHCIAGMYEIATKEAFEWLKTEPKLVIDASLIGRLLD
DMQSTSFEQQRGHVSSAVQCYMAQYGVTAEEACEKLRMAAIAWKDVNEACLRPTVFPMPIL
LPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVKPIEV

SEQ ID NO: 9

SCH51-3228-11_opt de *D. winteri*, Secuencia de ADN optimizada en codón de SCH51-3228-11.

ATGGCATCTACTCTTCCACTGCCGGCTTATGGTGATTCTGAGGTTGTTCGTCGTTCCGCGGG
TTTTCACCCTACCATCTGGGGCGATCACTTCTGTCCTATAAGCCAGACCCGACCAAGATTG
ACGAGTGGAATAAGCGTGTGCGAGGAAGTAAAAGAAGTAAAAAGATCCTGTCCAACG
CAAAGGTACTGTGCGAGGAGCTGAATCTGCTGGATGACCTGGTGCATCTGGGCATCAGCTA
TCACTTCGAAAAGGAAATTGACGACGCTTTCAGCAAATTTTTGATACGCACCTGGACGTC
TTTCCGAAAGATGACCTGTATGCGACCGCGCTGCGCTTTGGTGTGCTGCGTAAACAGGGTC
ATCGCGTGTCTCCTGATGTGTTCAAGAAATTTAAAGATGAACAGGGCAATTTCAAGGCCGA
GTTGAGCACGGACGCCAAAGGTTTGGCTCTGCCTGTACGACGTTGCATATCTGAGCACCCGT
GGTGAAGATATCCTGGACGAAGCGATTCCGTTACCAAGGAACATCTGCGCTCGTGCATTT
CCCATGTAGATAGCCACATGGCGGCCAAGATCGAGCACAGCCTGGAGCTGCCTTTGCACCA
TCGTATTCCGCGCCTGGAGAATCGCCATTACATTAGCGTCTATGAGGGTGACAAAGAGCGC
AACGAAGTCGTGTTAGAGCTGGCGAAGCTGGACTTCAACCTGATTCAAATTTCTGCATCAAC
GCGAGCTGCGCGACATTACCATGTGGTGGAAAGAGATTGATCTGGCAGCGAAGCTGCCGTT

ES 2 759 724 T3

CATCCGCGATCGTCTGGTTGAGTGCTACTACTGGATCATGGGCGTCTACTTCGAGCCGATCT
ACAGCCGCGCTCGTGTGTTTTTCGACGAAGATGACCATCCTGGTTAGCGTTGTTGATGACATT
TATGACGTTTACGCGACCGAAGATGAACTGCAGCTGTTTACGGACGCAATCTACCGTTGGG
ACGCGGATGATATCGACCAGCTGCCGCAATACTTGAAAGATGCGTTCATGGTTTTGTACAA
CACCGTCAAAACGCTGGAAGAAGAAGCTGGAGCCGGAAGGCAACAGCTACCGTGGTTACTA
TGTTAAAGATGCGATGAAAGTTCTGGCGCGGACTACTTCGTCGAGCACAAGTGGTATAAC
CGTCAGATTGTGCCGAGCGTCGAGGAATACCTGAAGATTAGCTGTATCAGCGTTGCCGTTT
ACATGGCAACGGTGCCTGCATCGCCGGTATGTACGAGATTGCGACGAAAGAAGCCTTCGA
ATGGTTGAAAACCGAGCCGAAGCTGGTTATCGACGCCAGCCTGATCGGTCGTTTGTGGAC
GACATGCAAAGCACGAGCTTCGAGCAGCAGCGCGGCCATGTGAGCAGCGCTGTTCAAGTGT
ATATGGCGCAATATGGCGTGACCGCAGAAGAAGCGTGCGAGAAGCTGCGTGAGATGGCAG
CAATTGCGTGGAAGATGTGAATGAAGCCTGTCTGCGTCCGACTGTGTTTCCGATGCCGAT
CCTGCTGCCGAGCATTAACTGGCGCGTGTGGCAGAGGTCATCTATCTGCGTGGTGACGGT
TACACCCACGCGGGTGGCGAAACCAAGAAACATATCACCGCAATGCTGGTTAAGCCGATT
GAAGTGTA

SEQ ID NO: 10

SCH51-998-28 de *D. winteri*, Marco de lectura abierto, secuencia de ADN de tipo silvestre.

ATGGATCTTAGTACTTCACCTGTTCTTTCTTCCCTCCCCCTTCCGGTGGAAGACGGAAAAAA
TCCGGCCGTTCCGCGTTCAGCTGGATTTCACCCAGTATTTGGGGTGATCATTTCCCTCTCCT
ACACTGAAGATCACAAGAAGCTGGATGCATGGAGCGAAAGGACTCAAGTGTGGAAGGAAG
AGGTGAGGAGAATTTAATCAATGCCAAGGGGTCCTAGAAAGAGTTGGATTGTTGGATGC
AATCCAACGCCTTGGGGTGAAATATCACTTTGAGAAAGAGATTGAAGAGGCATTACACCAT
ATTTATGTTGCAGAACTCATGTTTCTACTGATGACTTATATTCCGTTTCTCTCCGGTTTCGA
CTTCTTAGACAACAAGGGTACAATGTATCTGCTGATGTATTTAAAAAGTTCAAAGATGAGA
GGGGCAACTTCAAGGCAAGCTTAAGTACTGATGCCAGGGGGTTGCTAAGCTTGTATGAAGC
TGCATTTCTCAGCATAACGAGGAGATGATATCTTAGATGAAGCCATAACTTTCACAAGAGAG
CAGCTTAAGTCTTCTATGACCCATGTTGATGCCCTCTTGCCAAACAAATAGCCCATGCCTT
AGAGGTACCAGCGCACAAGCGCATACAAAGACTAGAGAACATTCGCTACCTCACAATCTA
CCAAGAAGAGAAAGGAAGGAATGATGTGTTGCTTGAGCTTGCCAAGTTGGATTTCAATATC

ES 2 759 724 T3

TTACAACAATTGCATAAGAAAGAACTGAGAGACCTTACAAAGTGGTGGGAAGGACACAGAC
GTTGCAGGAAAGCTACCTTTCATCAGAGATAGGTTGGTGGAAATGCTATTATTGGATCTTGG
GTGTGTATTATGAGCCAGAATACTCCAGAGCTAGAATTTTTTCTACCAAAATGACAATCAT
GGTCTCAGTTGTTGATGACATATATGACGTATATGCTACTGAAGATGAGCTCCAACTATTCA
CTGATGCAATCTATAGGTGGGATCTGGAGGGCCTAGATCAACTCCCACAGTTCTTGAAAGA
CTGTTTTCTGTACTCTATGACACCGTCAAGGAATTAGAAGACGAACTAGAACCGGAAGGA
AAATCCTATCGTGGATACTATGTAAAGGATGCGATGAAGTTTTGGCTAGAGATTACTTCG
TTGAGCACAATGGTATAACAGAAACATAGTGCCAAGTGTAGAAGAATATCTCCGTGTTTC
TTGCATCAGTGTTCAGTCCATATGGCTAACGTCCATTGCTGTGCTGGGATGGGAGATGTA
ATGAGCAAAGAGGCATTTCGAATGGTTGAAGAGTGAACCAAAGGTTGTAATGGATGCATCA
CTAATTGGCCGACTGCTCGATGACATGCAGTCCACCGAGTTTGAGCAAAAGAGAGGCCATG
TTGCATCGGCTGTCCAATGTTACATGAATGAGTATGGAGTGACTIONTACAAAGAAGCGTGTGA
AAAGCTGCATGAAATGGCTGCCCTTGCATGGAAAGACGTAAACCAGGCTTGCCTTAAACCA
ACTGTTTTCCCTCTCCCTGTATTTATGCCTGCAATCAACCTTGC CGAGTGGCTGAAGTCAT
CTACCTTCGTGGAGATGGGTATACTCATTCAGGAGGAGAGACTAAAGAAAATATCACGTTG
ATGCTTGTCAATCCAATCTCTGTGTGA

SEQ ID NO: 11
SCH51-998-28 de *D. winteri*, secuencia de aminoácidos.

MDLSTSPVLSSSPLPVEDGKNPAVRRSAGFHPSIWGDHFLSYTEDHKKLDAWSERTQVLKEEV
RRILINAKGSLEELDLLDAIQRLGVKYHFEKEIEEALHHIYVAETHVSTDDLYSVSLRFLLRQQ
GYNVSADVFKKFKDERGNFKASLSTDARGLLSLYEAAFLSIRGDDILDEAITFTREQLKSSMTH
VDAPLAKQIAHALEVP AHKRIQRLENIRYLTIYQEEKGRNDVLELAKLDFNILQQLHKKELRD
LTKWWKDTDVAGKLPFIRDRLVECYYWILGVVYYPEYSRARIKSTKMTIMVSVVDDIYDVYAT
EDELQLFTDAIYRWDLEGLDQLPQFLKDCFLVLYDTVKELEDELEPEGKSYRGYYVKDAMKVL
ARDYFVEHKWYNRNIVPSVEEYLRVSCISVAVHMANVHCCAGMGDVMSKEAFEWLKSEPKV
VMDASLIGRLLDDMQSTEFQKRGHVAVQCYMNEYGVTYKEACEKLHEMAALAWKDVN
QACLKPTVFPLPVFMPAINLARVAEVIYLRGDGYTHSGGETKENITLMLVNPISV

5 SEQ ID NO: 12
SCH51-998-28_opt de *D. winteri*, Secuencia de ADN optimizada en codón de SCH51-998-28.

ES 2 759 724 T3

ATGGATCTGAGCACCAGTCCGGTTCTGAGCAGCTCACCGCTGCCGGTTGAAGATGGTAAAA
ATCCGGCAGTTCGTCTAGCGCAGGTTTTCATCCGAGCATTGGGGTGATCATTTTCTGAGC
TATACCGAGGATCACAAAAAACTGGATGCATGGTCAGAACGTACCCAGGTTCTGAAAGAA
GAAGTGCGTTCGTATTCTGATTAATGCAAAAGGTAGCCTGGAAGAAGTGGATCTGCTGGATG
CAATTCAGCGTCTGGGTGTTAAATATCACTTTGAGAAAGAAATCGAAGAAGCCCTGCATCA
TATTTATGTTGCAGAAACCCATGTGTCAACCGATGATCTGTATAGCGTTAGCCTGCGTTTTC
GTCTGCTGCGTCAGCAGGGTTATAATGTTAGCGCAGATGTGTTCAAAAAATTCAAAGATGA
ACGCGGTAACCTCAAAGCAAGCCTGAGCACCGATGCACGTGGTCTGCTGAGCCTGTATGAA
GCAGCATTTCTGAGCATTTCGTGGTGATGATATTCTGGATGAAGCAATTACCTTTACCCGTGA
ACAGCTGAAAAGCAGCATGACCCATGTTGATGCACCGCTGGCAAAACAAATTGCACATGC
ACTGGAAGTTCCGGCACATAAACGTATTCAGCGCCTGGAAAATATTCGCTATCTGACCATT
TACCAAGAAGAGAAAGGTCGTAACGATGTTCTGCTGGAAGTGGCCAAACTGGATTTTAAACA
TTCTGCAGCAGCTGCATAAAAAAGAACTGCGTGATCTGACCAAATGGTGGAAAGATACCG
ATGTTGCAGGTAAACTGCCGTTTATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTCTG
GGCGTTTATTATGAGCCGGAATATAGCCGTGCACGATTTTTAGCACCAAATGACCATTAT
GGTTAGCGTGGTGGATGACATCTATGATGTTTATGCAACCGAAGATGAACTGCAGCTGTTT
ACCGATGCAATTTATCGTTGGGATCTGGAAGGTCTGGATCAGCTGCCCGAGTTCCTGAAAG
ATTGTTTTCTGGTTCTGTATGATACCGTGAAAGAACTGGAAGATGAGCTGGAACCGGAAGG
TAAAAGCTATCGTGGTTATTATGTTAAAGATGCCATGAAAGTTCTGGCACGCGATTATTTTG
TTGAGCACAAATGGTATAACCGCAATATTGTTCCGAGCGTGGAAGAATATCTGCGTGTTAG
CTGTATTAGCGTTGCAGTTCACATGGCAAATGTTTATTGTTGTGCAGGTATGGGTGATGTGA
TGAGCAAAGAAGCATTGGAATGGCTGAAAAGTGAACCGAAAGTTGTTATGGATGCCAGCCT
GATTGGTCGCCTGCTGGACGATATGCAGAGCACCGAATTTGAACAGAAACGTGGTCATGTT
GCAAGCGCAGTTCAGTGTTATATGAATGAATATGGCGTGACCTATAAAGAGGCATGCGAAA
AACTGCATGAAATGGCAGCACTGGCATGGAAAGATGTTAATCAGGCATGTCTGAAACCGA
CCGTTTTTCCGCTGCCTGTTTTTATGCCCTGCAATTAATCTGGCACGTGTTGCCGAAGTTATTT
ACCTGCGTGGGGATGGTTATACCCATAGCGGTGGTGAAACCAAAGAAAACATTACCCTGAT
GCTGGTTAATCCGATTAGCGTTTAA

SEQ ID NO: 13
SCH52-13163-6 de *D. lanceolata*, Secuencia de ADN.

ES 2 759 724 T3

ATGGATGTTCTAATTCCTCCCTGTGGCTTCCACTCTCCCTCTGCCCGAAGATGGAACTT
GGACGTCGTTTCGCAGATCCGCCGGGTTTCATCCGACGGTCTGGGGCGATCACTTCCTCGCTT
ACTCGCCCGATCCAACCAAAATAGATGCTTGGACTAAAAGAGTTGAAGAGCTGAAGCAAG
AAGTGAAGAGGATTCTAAGCAATGTGAAAGGGTCACTGGAAGAGCTGAACTTGCTTGATG
CTATCCAACACCTTGGGATTGGTTATCATTTTGAGAAAGAGATTGATGATGCTTTACAATA
ATCTTTGATTCCCATATTGATGCTTTTCCTACTGATGATCTATATGTGGCTGCCCTCCGATTT
AGCCTACTAAGGCGACAAGGGCACTGTGTTTCTTCAGATGTATTCAAAAAATTCAAAGATG
AGCAGGGGAATTTCAAGGCAGAGCTGAGCACCGATGCGAAAGGTTTGCTGAGTCTCTATGA
CGCGGCGTATCTCAGTGTAAGAGGGGAAGATATATTGGATGAGGCCATTCCTTTCACTAGG
GAGCACCTTAGGACTTGTATTAGCCATGTAGATTCTCATTTGGCAGCAAAAATTGAGCATTG
TCTAGAGCTTCCCTGCATCATCGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTG
TACGAAGGAGAGAAGGAAAGGAATGAAGTTGTACTAGAGCTTGCCAAATTAGATTTCAAT
CTGATTCAAATCTTGCACCAAAGAGAGCTGAGGGACATCACAAACGTGGTGAATGAGATTG
ACCTCGCAGCAAAGCTACCATTTATTAGGGATAGGTTGGTGGAGTGCTACTATTGGATCAT
GGGTGTCTATTTTGAACCAATATTCTCAAGGGCTAGAGTTTTTTTCGACCAAAATGACAATTT
TGGTCTCAGTTGTGACGACATATATGATGTCTACGCTACAGAGGATGAGCTCCAACCTTTTC
ACTGACGCAATCTATAGGTGGGATGCCGAGGACATTGAGCAGCTTCCACAGTACTTGAAAG
ATTCCTTTTCTTGTACTCTATAACACCGTGAAGGACTTAGAAGAGGAGCTGAAACCAGAAGG
AAACTCATATCGTGGAGACTATGTAAAAGATGCGATGAAGGTTTTGGCAAGAGATTACTTT
GTGGAGCACAATGGTATAACAGAAAAATTGTACCGTCAGTAGAGGACTACCTACGAATTT
CTTGCATTAGTGTGGCCGTTTCATATGGCTACAGTTCATTGTTGTGCTGGGATGGATGAAATT
GCAACCAAAGAGGCATTGCAATGGTTGAAGACCGAACCTAAACTTGTTATAGATGCATCAC
TGATTGGGCGTCTCCTCGATGACATGCAGTCCACCTCGTTTGAGCAACAGAGAGGTCATGT
GTCATCGGCGGTACAGTGTTACATGATCCAATATGGCGTATCACACGAAGAAGCGTGTGAG
AAGTTGACAGAAATGGCTGCAATTGCATGGAAAGATGTAAACCAAGCATGCCTTAGGCCC
ACTGTTTTCCCAATGCCTATTCTTCTGCCTTCAATCAACCTTGCACGTGTGGCAGAAGTCAT
CTACCTGCGCGGAGATGGATATACACATGCTGGTGGTGAGACCAAAAAACATATCACGGCC
ATGCTTGTTGAACCAATCCAAGTCTGA

SEQ ID NO: 14

SCH52-13163-6 de *D. lanceolata*, Marco de lectura abierto, secuencia de ADN de tipo silvestre.

ES 2 759 724 T3

MDVLIPSPVASTLPLPEDGNLDVVRRSAGFHPTVWGDHFLAYSPDPTKIDAWTKRVEELKQEV
KRILSNVKGSLEELNLLDAIQHLGIGYHFEKEIDDALQLIFDSHIDAFPTDDLYVAALRFSLLRRQ
GHCVSSDVFKKFKDEQGNFKAELSTDAKGLLSLYDAA YLSVRGEDILDEAIPFTREHLRTCISH
VDSHLAAKIEHSLELPLHHRIPRELENRHYISVYEGERNEVVLELAKLDFNLIQILHQRELRDIT
TWWNEIDLAAKLPFIRDRLVECYWIMGVYFEPIFSRARVFSTKMTILVSVVDDIYDVYATEDE
LQLFTDAIYRWDAEDIEQLPQYLKDSFLVLYNTVKDLEELKPEGNSYRGDYVKDAMKVLAR
DYFVEHKWYNRKIVPSVEDYLRISCISVAVHMATVHCCAGMDEIATKEAFEWLKTEPKLVIDA
SLIGRLLDDMQSTSFEQQRGHVSSAVQCYMIQYGVSHEEACEKLTEMAAIAWKDVNQAACLRP
T VFPMPILLPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVEPIQV

SEQ ID NO: 15

SCH52-13163-6_opt, Secuencia de ADN optimizada en codón de SCH51-13163-6.

ATGGATGTTCTGATTCCGAGTCCGGTTGCAAGCACCTGCCGCTGCCGGAAGATGGTAATC
TGGATGTTGTTTCGTCTAGCGCAGGTTTTTCATCCGACCGTTTGGGGTGATCATTTTCTGGCA
TATAGTCCGGATCCGACCAAAATTGATGCATGGACCAACGTGTTGAGGAACTGAAACAA
GAAGTGAAACGTATTCTGAGCAATGTGAAAGGTAGCCTGGAAGAACTGAATCTGCTGGAT
GCAATTCAGCATCTGGGTATTGGTTATCACTTCGAGAAAGAAATTGATGATGCACTGCAGC
TGATCTTTGATAGCCATATTGATGCCTTTCCGACCGATGATCTGTATGTTGCAGCACTGCGT
TTTAGCCTGCTGCGTCGTCAGGGTCATTGTGTTAGCAGTGATGTTTTCAAAAAATTCAAAGA
CGAGCAGGGCAACTTTAAAGCAGAACTGAGCACCGATGCAAAAAGGTCTGCTGAGCCTGTA
TGATGCCGCATATCTGAGCGTTCGTGGTGAAGATATTCTGGATGAAGCAATTCGTTTACCC
GTGAACATCTGCGTACCTGTATTAGCCATGTGGATAGCCATCTGGCAGCAAAAATTGAACA
TAGTCTGGAACCTGCCTCTGCATCATCGTATCCGCGTCTGGAAAATCGTCACTATATTAGCG
TTTATGAAGGCGAAAAAGAACGCAATGAAGTTGTGCTGGAACCTGGCAAACTGGATTTTAA
CCTGATTCAGATTCTGCATCAGCGTGAACCTGCGTGATATTACCACCTGGTGAATGAAATT
GACCTGGCAGCCAACTGCCGTTTATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTAT

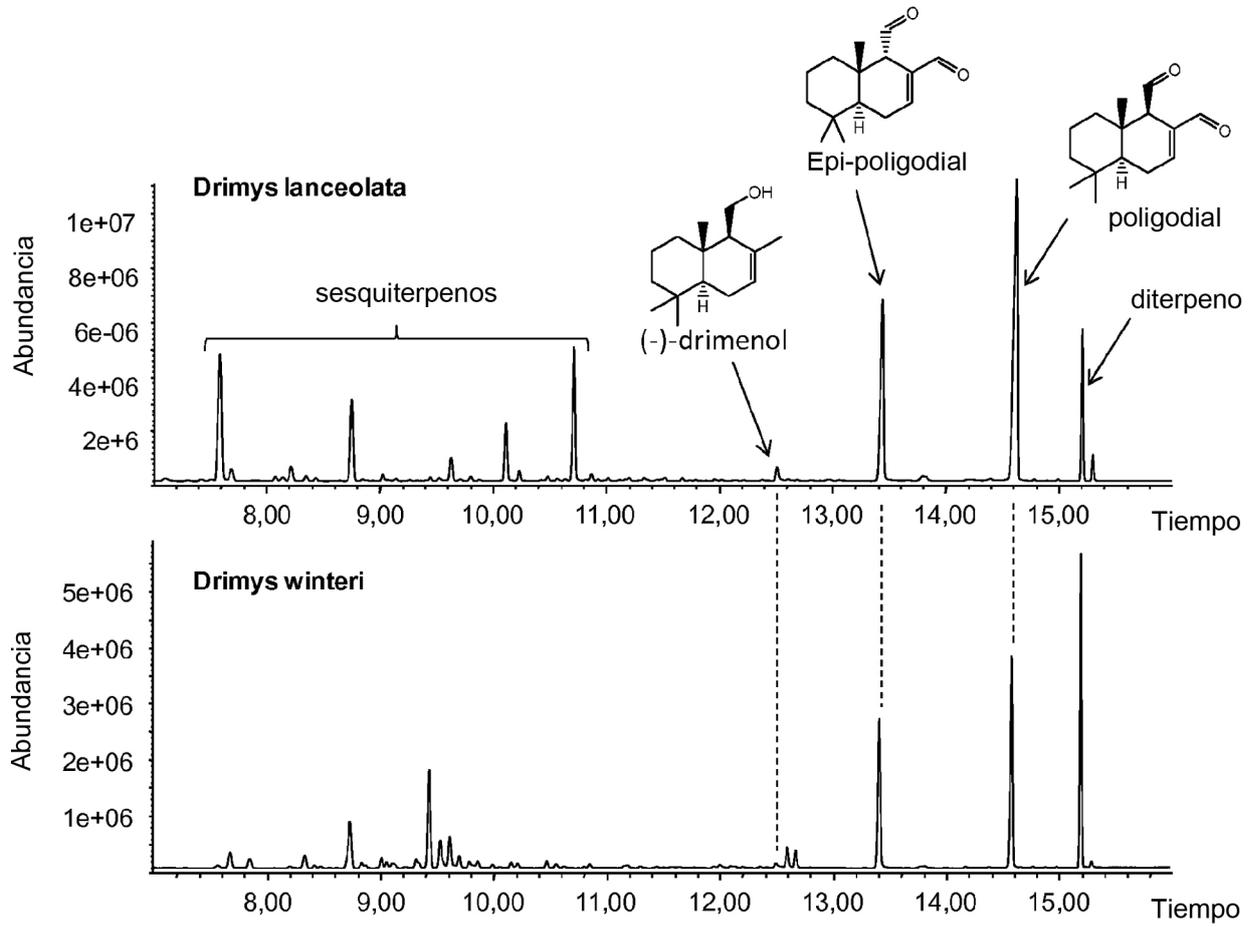
ES 2 759 724 T3

GGGCGTGTATTTTGAACCGATTTTATAGCCGTGCACGTGTGTTTAGCACCAAAATGACCATTC
TGGTTAGCGTGGTGGATGATATCTATGATGTTTATGCAACCGAAGATGAGCTGCAACTGTTT
ACCGATGCCATTTATCGTTGGGATGCAGAAGATATTGAACAGCTGCCTCAGTATCTGAAAG
ATAGCTTTCTGGTTCTGTACAACACCGTGAAAGATCTGGAAGAAGAAGTAAACCGGAAGG
TAATAGCTATCGTGGTGATTATGTTAAAGACGCCATGAAAGTTCTGGCACGCGATTATTTTG
TTGAGCACAAATGGTATAACCGCAAAATTGTTCCGAGCGTGGAAGATTATCTGCGTATTAG
CTGCATTAGCGTTGCAGTTCACATGGCAACCGTTCATTGTTGTGCAGGTATGGATGAAATTG
CAACCAAAGAAGCATTGAGTGGCTGAAAACCGAACCGAAACTGGTTATTGATGCAAGCCT
GATTGGTCGTCTGCTGGACGATATGCAGTCAACCAGCTTTGAACAGCAGCGTGGTTCATGTT
AGCAGCGCAGTTCAGTGTATATGATTCAGTATGGTGTAGCCATGAAGAAGCATGCGAAA
AACTGACCGAAATGGCAGCAATTGCATGGAAAGATGTTAATCAGGCATGTCTGCGTCCGAC
CGTGTTCCTATGCCGATTCTGCTGCCGAGCATTAACTGGCACGTGTTGCCGAAGTTATCT
ATCTGCGTGGTGTATGGTTATACCCATGCCGGTGGTGAAACCAAAAAACATATTACCGCAAT
GCTGGTAGAACCGATTCAGGTTTAA

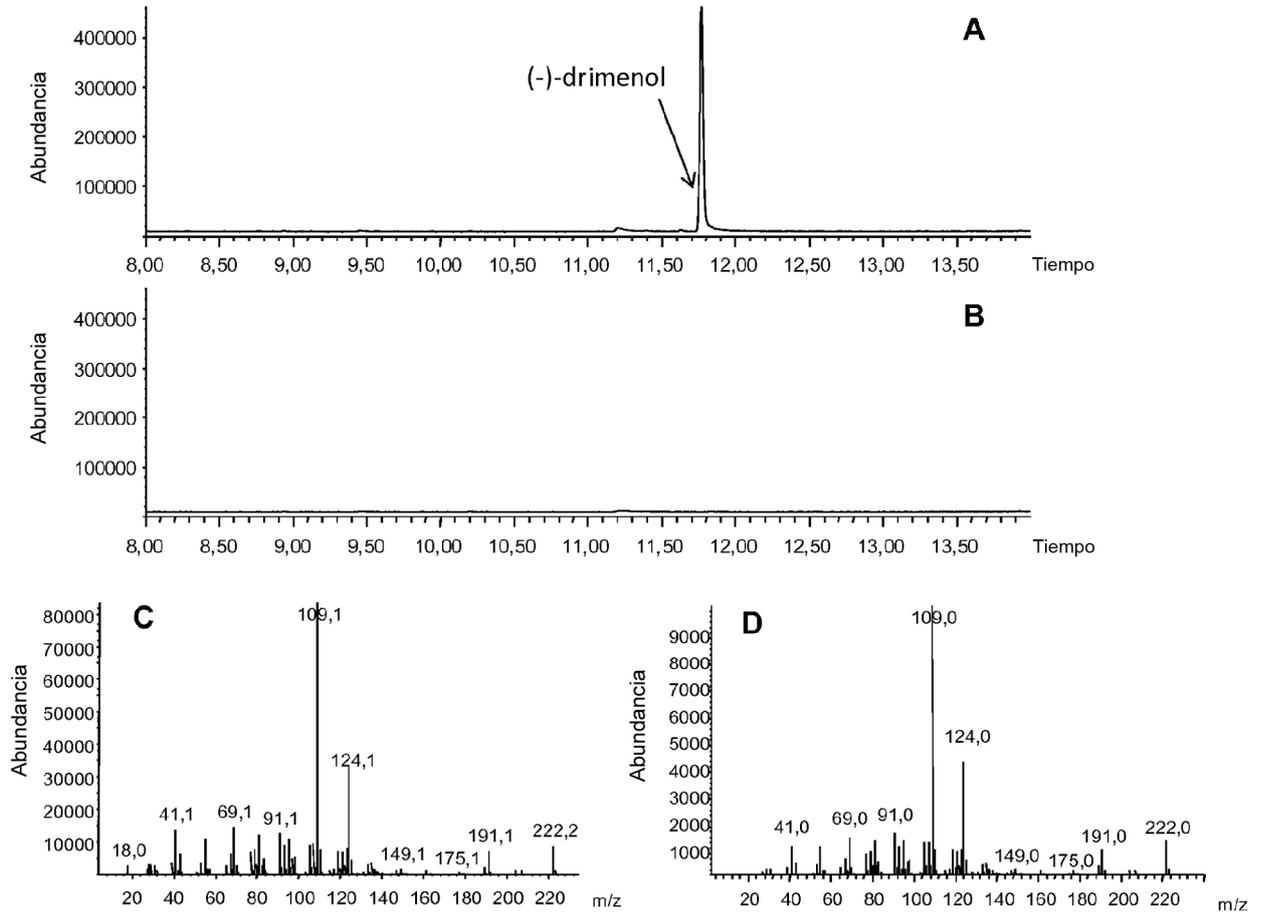
REIVINDICACIONES

1. Un procedimiento para producir drimenol que comprende:
 - i) poner en contacto difosfato de farnesilo (FPP) con un polipéptido que tiene actividad drimenol sintasa y que comprende una secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia a una secuencia de SEQ ID NO: 2 para producir el drimenol; y
 - ii) opcionalmente aislar el drimenol.
2. El procedimiento de conformidad con la reivindicación 1 en el que el drimenol se aísla.
3. El procedimiento de conformidad con la reivindicación 1 o 2 en el que el drimenol se produce con una selectividad mayor que o igual al 30 %.
4. El procedimiento de conformidad con una cualquiera de las reivindicaciones 1 a 3 que comprende poner en contacto el drimenol con al menos una enzima para producir un derivado de drimenol.
5. El procedimiento de conformidad con una cualquiera de las reivindicaciones 1 a 3 que comprende convertir el drimenol en un derivado de drimenol usando una síntesis química.
6. Un polipéptido aislado que tiene actividad drimenol sintasa que comprende una secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia a una secuencia de aminoácidos de SEQ ID NO: 2.
7. Una molécula de ácido nucleico aislada que codifica un polipéptido de conformidad con la reivindicación 6.
8. La molécula de ácido nucleico de la reivindicación 7 que comprende la secuencia de ácido nucleico de SEQ ID NO: 1 o SEQ ID NO: 3.
9. El procedimiento de conformidad con la reivindicación 1 que comprende las etapas de transformar una célula hospedadora no humana o un organismo hospedador no humano con un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 70 % de identidad de secuencia a una secuencia de SEQ ID NO: 2 y cultivar la célula hospedadora o el organismo hospedador en condiciones que permitan la producción del polipéptido.
10. Un vector que comprende la molécula de ácido nucleico de las reivindicaciones 7 u 8.
11. El vector de la reivindicación 10 en el que el vector es un vector procariótico, un vector vírico o un vector eucariótico.
12. El vector de la reivindicación 10 o la reivindicación 11 que es un vector de expresión.
13. El procedimiento de conformidad con la reivindicación 9 en el que la célula es una célula procariota o eucariota.
14. El procedimiento de conformidad con la reivindicación 13 en el que la célula procariota es una célula bacteriana.
15. El procedimiento de conformidad con la reivindicación 13 en el que la célula eucariota es una célula de levadura o una célula vegetal.
16. Un organismo hospedador no humano o una célula hospedadora no humana que comprende la molécula de ácido nucleico de la reivindicación 7 u 8 o el vector de una cualquiera de las reivindicaciones 10 a 12.
17. El organismo hospedador no humano o la célula hospedadora no humana de la reivindicación 16, en los que el organismo o la célula hospedadores no humanos son una bacteria, una levadura, una célula fúngica o una célula vegetal.
18. El organismo hospedador no humano o la célula hospedadora no humana de la reivindicación 17, en los que la bacteria es *E. coli* y la levadura es *Saccharomyces cerevisiae*.

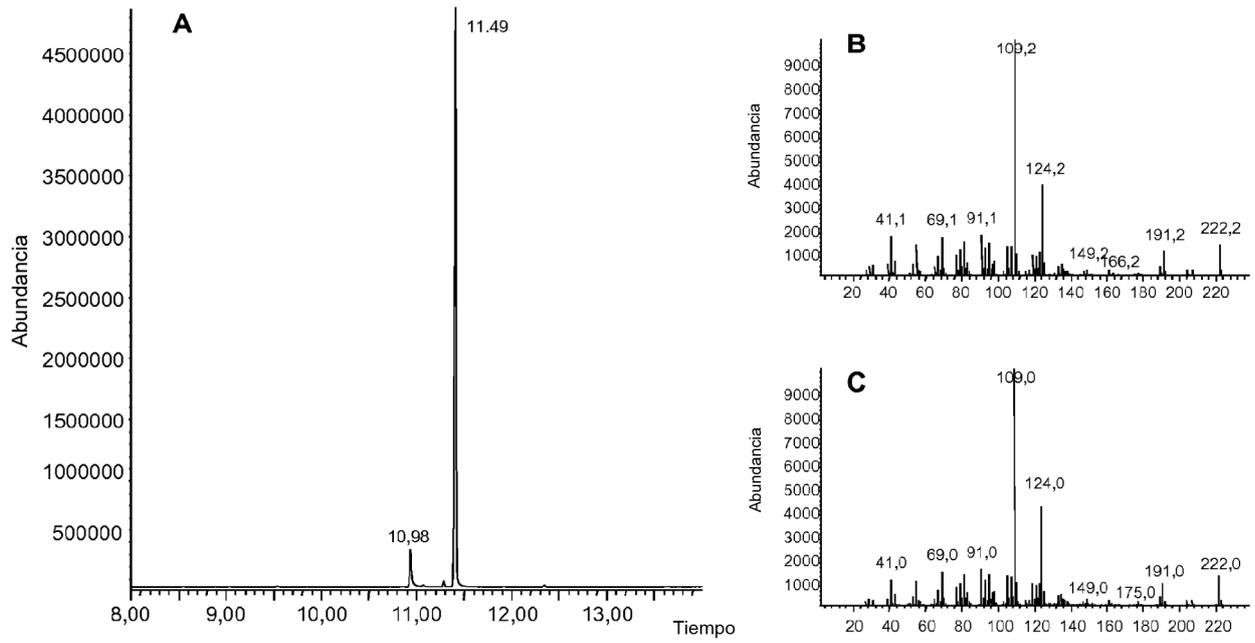
- Figura 1 -



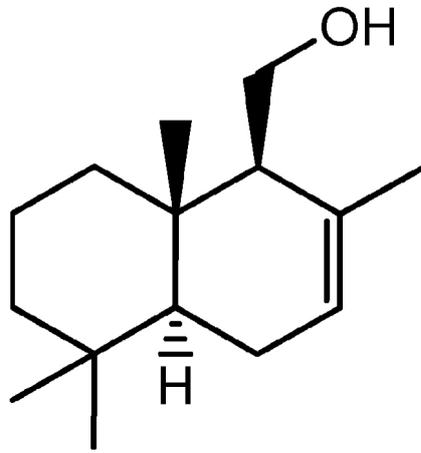
- Figura 2 -



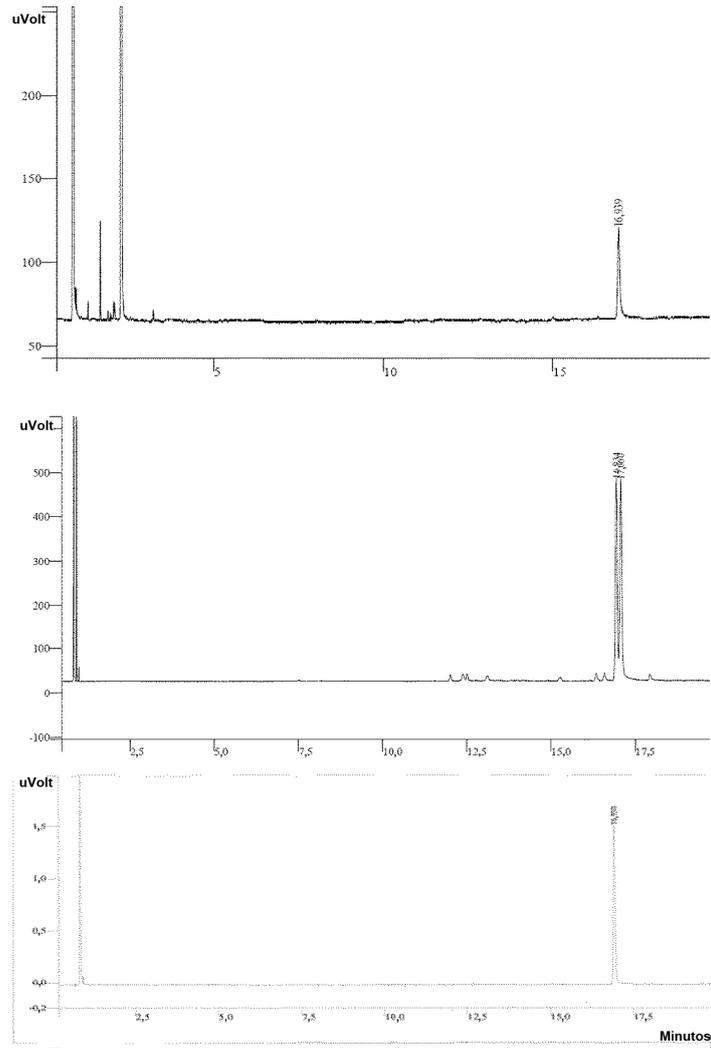
- Figura 3 -



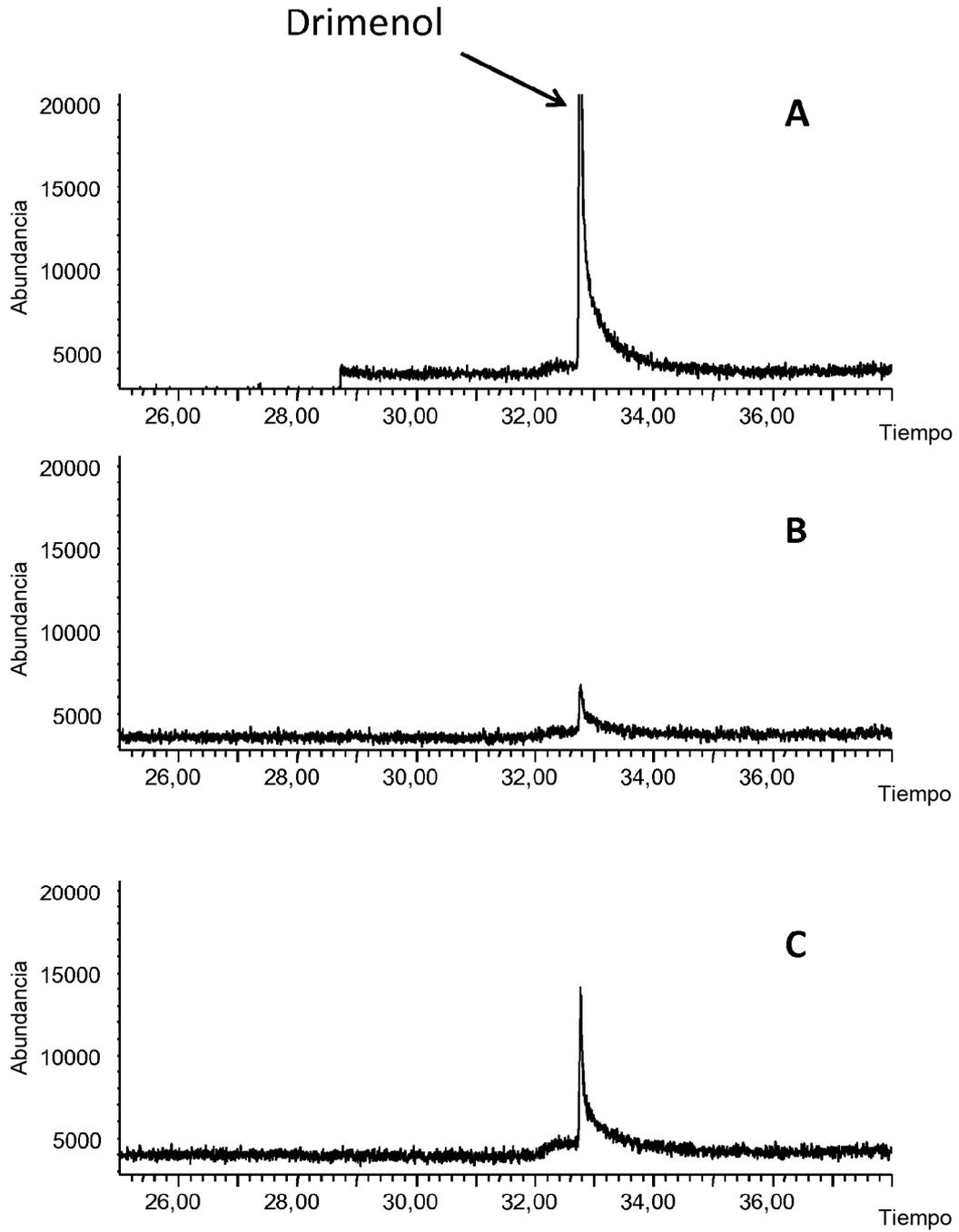
- Figura 4 -



- Figura 5 -



- Figura 6 -



- Figura 7 -

