

DRIMENOL SINTASAS II

Campo Técnico

El campo se relaciona con métodos para producir Drimenol,
5 dicho método comprende poner en contacto por lo menos un
polipéptido con pirofosfato de farnesilo (FPP). En particular,
dicho método se puede llevar a cabo *in vitro* o *in vivo* para
producir Drimenol, un compuesto muy útil en los campos de
perfumería. También se proporciona aquí una secuencia de
10 aminoácidos de un polipéptido útil en los métodos proporcionados
aquí. También se proporcionan aquí un ácido nucleico que
codifica el polipéptido de una realización aquí y un vector de
expresión que contiene dicho ácido nucleico. Se proporciona
adicionalmente aquí un organismo anfitrión no humano o una
15 célula transformada que se va a utilizar en el método para
producir Drimenol.

Antecedentes

Los terpenos se encuentran en la mayoría de los organismos
20 (microorganismos, animales y plantas). Estos compuestos se
componen de cinco unidades de carbono llamados unidades de
isopreno y se clasifican por el número de estas unidades
presentes en su estructura. Por lo tanto los monoterpenos,
sesquiterpenos y diterpenos son terpenos que contienen 10, 15 y
25 20 átomos de carbono respectivamente. Los sesquiterpenos, por
ejemplo, se encuentran ampliamente distribuidos en el reino

vegetal. Muchas moléculas de sesquiterpenos se conocen por sus propiedades de sabor y fragancia y sus efectos cosméticos, medicinales y antimicrobianos. Se han identificado numerosos hidrocarburos de sesquiterpenos y sesquiterpenoides.

5

La producción biosintética de terpenos implica enzimas llamadas sintasas de terpeno. Existe prácticamente una infinidad de sintasas de sesquiterpeno presentes en el reino vegetal, que utilizan el mismo sustrato (pirofosfato de farnesilo, FPP), pero
10 que tienen diferentes perfiles de productos. Se han clonado los genes y cADN que codifican sintasas de sesquiterpeno y las correspondientes enzimas recombinantes caracterizadas.

En la actualidad la principal fuente de Drimenol son las
15 plantas que contienen naturalmente Drimenol y los contenidos de Drimenol en estas fuentes naturales son bajos. Se han desarrollado métodos de síntesis química, pero todavía son complejos y no rentables.

20 **Resumen**

Se proporciona aquí un método para producir Drimenol que comprende:

i) poner en contacto un pirofosfato de terpeno acíclico con un polipéptido que tiene actividad de Drimenol
25 sintasa y que tiene por lo menos, o por lo menos aproximadamente 70% de identidad de secuencia con una

secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 para producir el Drimenol; y

5 ii) opcionalmente aislar el Drimenol.

Adicionalmente se proporciona aquí un polipéptido aislado que tiene actividad de Drimenol que comprende una secuencia de aminoácidos que tiene por lo menos o por lo menos
10 aproximadamente 70%, o más de identidad con la secuencia de aminoácidos de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14.

15 También se proporciona aquí una molécula de ácido nucleico aislada que codifica un polipéptido que tiene por lo menos, o por lo menos aproximadamente 70% de identidad de secuencia con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14

20

Descripción de los dibujos.

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

25

Figura 2. Análisis GCMS de los sesquiterpenos producidos por la D1Tps589 recombinante en ensayos *in vitro*. A. Cromatograma iónico total del perfil de sesquiterpeno de una incubación de la proteína de D1Tps589 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 min. D. Espectro de masa de un estándar auténtico de (-)-drimenol.

Figura 3. Análisis GCMS de los sesquiterpenos producidos *in vivo* por la D1Tps589 recombinante en células bacterianas diseñadas por ingeniería. A. Cromatograma iónico total. B. Espectro de masa del pico a 11.49 min. C. Espectro de masa de un estándar auténtico de (-)-drimenol. El compuesto que se eluye a 10.98 min es farnesol producido por la enzima D1Tps589 o que resulta de la hidrólisis del FPP en exceso producido por las células de *E. coli*.

Figura 4. Estructura de (-)-drimenol producido por la sintasa de D1Tps589 recombinante.

Figura 5. Cromatogramas de GCYFID quirales de (-)-drimenol producido por la enzima recombinante (superior), drimenol racémico obtenido químicamente (medio) y (-)-drimenol auténtico (inferior).

Figura 6. Cromatograma iónico total de Análisis GCMS de los sesquiterpenos producido en ensayos *in vitro* por las proteínas recombinantes SCH51-3228-9 (A), SCH51-998-28 (B) o SCH52-13163-6 (C).

5

Figura 7. Cromatograma iónico total de Análisis GCMS de los sesquiterpenos producidos *in vivo* por células bacterianas modificadas por ingeniería que expresan las proteínas recombinantes diferentes SCH51-3228-9 (A), SCH51 -998- 28 (B) o
 10 SCH52-13163-6 (C). El farnesol detectado resulta de la hidrólisis del exceso de FPP producido por las células de *E. coli* o se puede en parte producir por las proteínas recombinantes.

15 **Descripción detallada**

Para las descripciones de las descripciones y las reivindicaciones adjuntas, el uso de "o" significa "y/o" a menos que se indique lo contrario. Del mismo modo, "comprenden", "comprende", "que comprende" "incluyen", "incluye" y "que
 20 incluye" son intercambiables y no pretenden ser limitantes.

Adicionalmente se debe entender que cuando en la descripción de varias realizaciones se utiliza el término "que comprende", aquellos expertos en la técnica entenderán que, en
 25 algunos casos específicos, una realización se puede describir alternativamente utilizando el lenguaje "que consiste

esencialmente de" o "que consiste de". En un aspecto, se proporciona aquí es un método de producción que comprende Drimenol:

5 i) poner en contacto un pirofosfato de terpeno acíclico, particularmente difosfato de farnesilo (FPP)) con un polipéptido que tiene actividad de Drimenol sintasa y que tiene por lo menos, o por lo menos aproximadamente 70%, particularmente 75%, particularmente 80%, particularmente 85%,
10 particularmente 90%, particularmente 95%, particularmente 96%, particularmente 97%, particularmente 98% o particularmente 99% o más de identidad de secuencia con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 para producir el Drimenol; y

15

ii) opcionalmente aislar el Drimenol.

En un aspecto, se aísla el Drimenol.

20 Adicionalmente se proporciona aquí un polipéptido aislado que tiene actividad de Drimenol que comprende una secuencia de aminoácidos que tiene por lo menos o por lo menos aproximadamente 70%, particularmente 75%, particularmente 80%, particularmente 85%, particularmente 90%, particularmente 95%,
25 particularmente 96%, particularmente 97%, particularmente 98% o más particularmente 99% o más de identidad con secuencia de

aminoácidos de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14.

5 Adicionalmente se proporciona aquí una molécula de ácido nucleico aislada que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene por lo menos o por lo menos aproximadamente 70%, particularmente 75%, particularmente 80%, particularmente 85%, particularmente 90%, particularmente 95%,
10 particularmente 96%, particularmente 97%, particularmente 98% o más particularmente 99% o más de identidad con secuencia de aminoácidos de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO:

15

Adicionalmente se proporciona aquí una molécula de ácido nucleico que comprende la secuencia seleccionada del grupo que consiste de la 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,
20 SEQ ID No: 13 y SEQ ID NO: 15.

Adicionalmente se proporciona aquí un método como se menciona en la Reivindicación 1 que comprende las etapas de transformar una célula anfitriona u organismo no humano con un
25 ácido nucleico que codifica un polipéptido que tiene por lo menos, o por lo menos aproximadamente, 70%, particularmente 75%,

particularmente 80%, particularmente 85%. Particularmente 90%, particularmente 95%, particularmente 96%, particularmente 97%, particularmente 98% o particularmente 99% o más identidad de secuencia de la secuencia de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 y cultivar la célula anfitriona o organismo bajo condiciones que permiten la producción del polipéptido.

Adicionalmente se proporciona por lo menos un vector que comprende las moléculas de ácido nucleico descritas.

Adicionalmente se proporciona aquí un vector seleccionado del grupo de un vector procariótico, vector vírico y un vector eucariótico.

Adicionalmente se proporciona aquí un vector que es un vector de expresión.

Como una "Drimenol sintasa" o como un "polipéptido que tiene una actividad de Drimenol sintasa", entendemos aquí un polipéptido capaz de catalizar la síntesis de Drimenol, en forma de cualquiera de sus estereoisómeros o una mezcla de los mismos, partiendo de un pirofosfato de terpeno acíclico, particularmente FPP. El Drimenol puede ser el solo producto puede ser parte de una mezcla de sesquiterpenos.

La capacidad de un polipéptido para catalizar la síntesis de un sesquiterpeno particular (por ejemplo Drimenol) simplemente se puede confirmar al realizar el ensayo de enzima
5 como se detalla en el Ejemplo 2-5.

También se entiende que los polipéptidos incluyen polipéptidos truncados siempre que mantengan su actividad de Drimenol sintasa.

10

Como se pretende aquí adelante, "una secuencia de nucleótidos obtenida mediante la modificación de la 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
15 NO: 15 o el complemento de la misma abarca cualquier secuencia que se ha obtenido al cambiar la secuencia de la 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 el complemento de la misma utilizando cualquier método conocido en la técnica, por
20 ejemplo al introducir cualquier tipo de mutaciones, tales como mutaciones de supresión, inserción o sustitución. Ejemplos de dichos métodos se citan en la parte de la descripción en relación con los polipéptidos variantes y los métodos para prepararlos.

25

El porcentaje de identidad entre dos secuencias peptídicas o nucleotídicas es una función del número de aminoácidos o residuos de nucleótidos que son idénticos en dos secuencias cuando se ha generado un alineamiento de estas dos secuencias.

5 Los residuos idénticos se definen como residuos que son iguales en las dos secuencias en una posición dada de la alineación. El porcentaje de identidad de secuencia, como se utiliza aquí, se calcula a partir de la alineación óptima, tomando el número de residuos idénticos entre dos secuencias dividiéndolo por el

10 número total de residuos en la secuencia más corta y multiplicando por 100. La alineación óptima es la alineación en la que el porcentaje de identidad es el más alto posible. Se pueden introducir espacios en una o ambas secuencias en una o más posiciones de la alineación para obtener alineación óptima.

15 Estos espacios luego se toman en consideración como residuos no idénticos para el cálculo del porcentaje de identidad de secuencia. La alineación con el propósito de determinar el porcentaje de identidad de aminoácidos o secuencia de ácidos nucleicos se puede conseguir de varias maneras utilizando los

20 programas de ordenador y por ejemplo los programas de ordenador disponibles públicamente en la World Wide Web. Preferiblemente, el programa BLAST (Tatiana et al, FEMS Microbiol Lett, 1999, 174:. 247-250, 1999) ajustado a los parámetros por defecto, disponible en el Centro Nacional de Información Biotecnológica

25 (NCBI) en <http://www.ncbi.nlm.nih.gov/BLAST/bl2seq/wblast2.cgi>, se puede utilizar para obtener una alineación óptima de las

secuencias peptídicas o nucleotídicas y para calcular el porcentaje de identidad de secuencia.

Abreviaturas utilizadas

5	pb	pares bases
	kb	kilobases
	BSA	albúmina de suero bovino
	ADN	ácido desoxirribonucleico
	cADN	ADN complementario
10	TDT	ditiotreitol
	FID	detector de ionización de llama
	FPP	pirofosfato de farnesilo
	GC	cromatógrafo gaseoso
	IPTG	isopropil-D-tiogalacto piranósido
15	LB	caldo de lisogenia
	EM	espectrómetro de masas
	MVA	ácido mevalónico
	PCR	reacción en cadena de polimerasa
	RMCE	intercambio de casete mediado por recombinasa
20	3'-/5'-RACE	amplificación rápida 3'y 5' de extremos de cADN
	ARN	ácido ribonucleico
	mARN	ácido ribonucleico mensajero
	miARN	ARN micro
25	siARN	ARN de interferencia pequeño
	rARN	ARN ribosomal

tARN ARN de transferencia

DEFINICIONES

El término "polipéptido" significa una secuencia de aminoácidos de secuencias de aminoácidos consecutivas, por ejemplo, por lo menos 15 residuos, por lo menos 30 residuos, por lo menos 50 residuos. En algunas realizaciones proporcionadas aquí, un polipéptido comprende una secuencia de aminoácidos que es una enzima, o un fragmento, o una variante de la misma.

10

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

15

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

20

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

25

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

5

Los términos "secuencia de ácidos nucleicos", "ácido nucleico" y "polinucleótido" se utilizan de manera intercambiable significan una secuencia de nucleótidos. Una secuencia de ácidos nucleicos puede ser un desoxirribonucleótido o ribonucleótido de cadena sencilla o de cadena doble de cualquier longitud, e incluyen secuencias codificantes y no codificantes de un gen, exones, intrones, secuencias complementarias codificantes y anticodificantes, ADN genómico, cADN , miARN, siARN, mARN, rARN, tARN, secuencias de ácido nucleico recombinante, secuencias de ADN y/o ARN de origen natural aisladas y purificadas, secuencia de ADN y ARN sintéticas, fragmentos, cebadores y sondas de ácido nucleico. El experto en la técnica es consciente de que las secuencias de ácidos nucleicos de ARN son idénticas a las secuencias de ADN con la diferencia de que la timina (T) se reemplaza por uracilo (U) .

Un "ácido nucleico aislado" o "secuencia de ácidos nucleicos aislados" se define como un ácido nucleico o secuencia de ácidos nucleicos que se encuentra en un entorno diferente de aquel en el que el ácido nucleico o secuencia de ácidos

25

nucleicos se produce de forma natural. El término "de origen natural" como se utiliza aquí que 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 ácidos nucleicos que está presente en un organismo, por ejemplo, en las células de un organismo, que se puede aislar de una fuente en la naturaleza y que no ha sido modificada intencionalmente por un ser humano en el laboratorio es de origen natural.

10 "Secuencia de ácidos nucleicos recombinante" son secuencias de ácidos nucleicos que resultan del uso de métodos de laboratorio (clonación molecular) para reunir el material genético de más de la fuente, crear una secuencia de ácidos nucleicos que no se produce de forma natural y no se encontraría en organismos biológicos.

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

25 El término "gen" significa una secuencia de ADN que comprende una región, que se transcribe en una molécula de ARN,

por ejemplo, un mRNA en una célula, unido operativamente a regiones reguladoras adecuadas, por ejemplo, un promotor. Por lo tanto un gen puede comprender varias secuencias unidas operativamente, tales como un promotor, una secuencia líder 5' que comprende, por ejemplo, las secuencias implicadas en la iniciación de traducción, una región de codificación de cADN o ADN genómico, intrones, exones, y/o secuencia no traducida 3' que comprende la secuencia, por ejemplo, sitios de terminación de transcripción.

10

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 que una o más partes de la secuencia de ácidos nucleicos están presentes que no están asociados entre sí en la naturaleza. Por ejemplo el promotor no está asociado en la naturaleza con parte o la totalidad de la región transcrita o con otra región reguladora. Se entiende que el término "gen quimérico" incluye construcciones de expresión en las que un promotor o secuencia reguladora de transcripción está unida operativamente a una o más secuencias de codificación o a una anticodificante, es decir, complemento de la cadena codificante, o secuencia de repetición invertida (codificante y anticodificante, por lo que la transcripción de ARN forma el ARN de cadena doble luego de transcripción).

25

Una "UTR 3'" o "secuencia 3' no traducida" (también denominada como "región no traducida 3', " o "extremo 3'") se refiere a la secuencia de ácidos nucleicos encontrada en dirección 3' de la secuencia de codificación de un gen, que
5 comprende por ejemplo, un sitio de terminación de transcripción y (en la mayoría, pero no todos imARN eucariotas) una señal de poliadenilación tal como AAUAAA o variantes de la misma. Después de terminación de la transcripción, la transcripción de mARN se puede dividir en dirección 3' de la señal de poliadenilación y
10 se puede agregar una cola poli (A), que está implicada en el transporte de mARN al sitio de traducción, por ejemplo, el citoplasma.

La "Expresión de un gen" implica la transcripción del gen y
15 la traducción del mARN en una proteína. Sobreexpresión se refiere a la producción del producto del gen según se mide por los niveles de mARN, la actividad del polipéptido y/o enzima en las células u organismos transgénicos que excede los niveles de producción en las células u organismos no transformados un
20 antecedente genético similar.

"Vector de expresión" como se utiliza aquí significa una molécula de ácido nucleico diseñada por ingeniería utilizando métodos de biología molecular y tecnología de ADN recombinante
25 para el suministro de ADN extraño o exógeno en una célula anfitriona. El vector de expresión normalmente incluye

secuencias requeridas para la transcripción adecuada de la
secuencia de nucleótidos. La región codificante generalmente
codifica una proteína de interés pero puede también codificar
para un ARN, por ejemplo, un ARN anticodificante, siARN y
5 similares.

Un "vector de expresión" como se utiliza aquí incluye
cualquier vector recombinante lineal o circular, que incluye,
pero no se limita a los vectores virales, bacteriófagos y
10 plásmidos. El experto 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
aquí operativamente unido a por lo menos una secuencia
reguladora, que controla la transcripción, traducción,
15 iniciación y terminación, tal como un promotor de transcripción,
operador o potenciador, o un sitio de unión a mRNA ribosomal y,
opcionalmente, que incluye por lo menos un marcador de
selección. Las secuencias de nucleótidos están "unidas
operativamente" cuando la secuencia reguladora se relaciona
20 funcionalmente con el ácido nucleico de una realización aquí.
"Secuencia reguladora" se refiere a una secuencia de ácidos
nucleicos que determina el nivel de expresión de las secuencias
de ácidos nucleicos de una realización aquí y es capaz de
regular la tasa de transcripción de la secuencia de ácidos
25 nucleicos unida operativamente a la secuencia reguladora. Las

secuencias reguladoras comprenden promotores, potenciadores, factores de transcripción, elementos promotores y similares.

"Promotor" se refiere a una secuencia de ácidos nucleicos que controla la expresión de una secuencia de codificación, al proporcionar un sitio de unión para la polimerasa de ARN y otros factores requeridos para transcripción adecuada que incluyen, sin limitación sitios de unión de factor de transcripción, y sitios de unión de proteína represora y activadora. El significado del término promotor también incluye el término "secuencia reguladora del promotor". Las secuencias reguladoras de promotor pueden incluir elementos en dirección 5' y en dirección 3' que pueden influenciar la transcripción, procesamiento de ARN o estabilidad de la secuencia de ácidos nucleicos de codificación asociada. Los promotores incluyen secuencias de origen natural y sintéticas. Las secuencias de ácidos nucleicos de codificación normalmente se encuentran en dirección 3' del promotor con respecto a la dirección de la transcripción que inicia en el sitio de iniciación de transcripción.

El término "promotor constitutivo" se refiere a un promotor no regulado que permite la transcripción continua de la secuencia de ácidos nucleicos a la que se une operativamente.

Como se utiliza aquí, el término "unido operativamente" se refiere a una unión de elementos de polinucleótidos en una relación funcional. Un ácido nucleico se "une operativamente" cuando se coloca en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, un promotor, o más bien una secuencia reguladora de transcripción, se une operativamente a una secuencia de codificación si esta afecta la transcripción de la secuencia de codificación. Unido operativamente significa que las secuencias de ADN que se unen son normalmente contiguas. La secuencia de nucleótidos asociada a la secuencia promotora puede ser de origen homólogo o heterólogo con respecto a la planta que se va a transformar. La secuencia también puede ser totalmente o parcialmente sintética. Independientemente del origen, la secuencia de ácidos nucleicos asociada con la secuencia del promotor se expresa o silencia de acuerdo con propiedades promotoras a las cuales se une después de unión con el polipéptido de una realización aquí. El ácido nucleico asociado puede codificar una proteína que se desea se exprese o suprima por todo el organismo en todo momento o, alternativamente, en un momento específico o en tejidos, células o compartimentos celulares específicos. Dichas secuencias de nucleótidos particularmente codifican proteínas que confieren rasgos fenotípicos deseables a las células u organismo anfitrión alterado o transformado. Más particularmente, la secuencia de nucleótidos asociada conduce a la producción de Drimenol en el

organismo. En particular, la secuencia de nucleótidos codifica Drimenol sintasa.

"Péptido objetivo" se refiere a una secuencia de aminoácidos que se orienta sobre una proteína o polipéptido a orgánulos intracelulares, es decir, mitocondrias, o plástidos, o al espacio extracelular (péptido de señal de secreción). Una secuencia de ácidos nucleicos que codifica un péptido objetivo se puede fusionar a la secuencia de ácidos nucleicos que codifica el extremo terminal amino, por ejemplo, el extremo N-terminal, de la proteína o polipéptido, o se puede utilizar para reemplazar un polipéptido objetivo nativo.

El término "cebador" se refiere a una secuencia de ácidos nucleicos corta que se hibrida con una secuencia de ácidos nucleicos plantilla y se utiliza para la polimerización de una secuencia de ácidos nucleicos complementaria a la plantilla.

Como se utiliza aquí, el término "célula anfitriona" o "célula transformada" se refiere a una célula (u organismo) modificado para albergar por lo menos una molécula de ácido nucleico, por ejemplo, un gen recombinante que codifica una proteína o secuencia de ácidos nucleicos deseada que, luego de transcripción produce una proteína Drimenol sintasa útil para producir Drimenol. La célula anfitriona es particularmente una célula bacteriana, una célula fúngica o una célula vegetal. La

célula anfitriona puede contener un gen recombinante que se ha integrado en los genomas nucleares o de orgánulos de la célula anfitriona. Alternativamente, el anfitrión puede contener el gen recombinante extra-cromosómicamente. Las secuencias homólogas
5 incluyen secuencias de ortólogos o parálogos. Los métodos para identificar ortólogos o parálogos que incluye métodos filogenéticos, similitud de secuencia e hibridación son conocidos en la técnica y se describen en este documento.

10 Los parálogos son el resultado de la duplicación de genes que da lugar a dos o más genes con secuencias similares y funciones similares. Los parálogos normalmente se agrupan y se forman mediante duplicación de los genes dentro de las especies de plantas relacionadas. Los parálogos se encuentran en los
15 grupos de genes similares utilizando análisis Blast en forma de par o durante el análisis filogenético de familias de genes utilizando programas tales como CLUSTAL. En los parálogos, las secuencias de consenso se pueden identificar características a las secuencias dentro de los genes relacionados y que tienen
20 funciones similares de los genes.

Los ortólogos, o secuencias de ortólogos, son secuencias similares entre sí, ya que se encuentran en especies que descienden de un ancestro común. Por ejemplo, se sabe que las
25 especies de plantas que tienen ancestros comunes contienen muchas enzimas que tienen secuencias y funciones similares. El

experto en la técnica puede identificar secuencias de ortólogos y predecir las funciones de los ortólogos, por ejemplo, mediante la construcción de un árbol poligénico para una familia de genes de una especie utilizando programas CLUSTAL o BLAST. Un método
5 para identificar o confirmar funciones similares entre las secuencias homólogas es mediante la comparación de los perfiles de transcripción en plantas que sobreexpresan o carecen de (en inactivados/atenuados) polipéptidos relacionados. El experto entenderá que los genes que tienen perfiles de transcripción
10 similares, con más de 50% transcripciones reguladas en común, o con más del 70% transcripciones reguladas en común, o mayor de 90% transcripciones reguladas en común tendrán funciones similares. Se espera que los homólogos, parálogos, ortólogos y cualquier otras variantes de las secuencias aquí funcionen de
15 una manera similar al hacer que las plantas produzcan proteínas de Drimenol sintasa.

Una realización de la proporcionada aquí suministra secuencias de aminoácidos de proteínas de Drimenol sintasa que
20 incluyen ortólogos y parálogos, así como métodos para identificar y aislar ortólogos y parálogos de las sintasas Drimenol en otros organismos. En particular, los ortólogos y parálogos de Drimenol sintasa identificados de esta manera retienen actividad de Drimenol sintasa y son capaces de producir
25 Drimenol a partir de precursores de FPP.

El término "marcador seleccionable" se refiere a cualquier gen que luego de expresión se puede utilizar para seleccionar una célula o células que incluyen el marcador seleccionable. Ejemplos de marcadores seleccionables se describen adelante. El experto en la técnica sabe que los diferentes marcadores seleccionables antibióticos, fungicidas, auxotróficos o herbicidas son aplicables a diferentes especies objetivo.

"Drimenol" para los propósitos de esta solicitud se refiere a (-)-drimenol (CAS: 468-68-8).

El término "organismo" se refiere a cualquier organismo multicelular o unicelular no humano, tales como una planta, o un microorganismo. En particular, un microorganismo es una bacteria, una levadura, un alga o un hongo. El término "planta" se utiliza de forma intercambiable para incluir las células vegetales que incluyen protoplastos de planta, tejidos de planta, cultivos de tejidos de células de planta que dan lugar a plantas regeneradas, o partes de plantas u órganos de plantas, tales como raíces, tallos, hojas, flores, polen, óvulos, embriones, frutas y similares. Cualquier planta se puede utilizar para llevar a cabo los métodos de una realización aquí.

El polipéptido que se pone en contacto con un pirofosfato acíclico, por ejemplo, FPP, *in vitro* se puede obtener por extracción a partir de cualquier organismo que lo exprese,

utilizando tecnologías de extracción de proteína o de enzima estándar. Si el organismo anfitrión es un organismo unicelular o célula que libera el polipéptido de una realización aquí en el medio de cultivo, el polipéptido simplemente se puede recolectar
5 del medio de cultivo, por ejemplo mediante centrifugación, seguido opcionalmente por etapas de lavado y resuspensión en soluciones de regulador adecuadas. Si el organismo o célula acumula el polipéptido dentro de sus células, el polipéptido se puede obtener por disrupción o lisis de las células y la
10 extracción adicional del polipéptido a partir del lisado celular.

El polipéptido que tiene una actividad de Drimenol sintasa, ya sea en forma aislada o junto con otras proteínas, por ejemplo
15 en un extracto de proteína cruda obtenido a partir de células o microorganismos cultivados, luego se puede suspender en una solución de regulador a pH óptimo. Si es adecuado, las sales, DTT, cationes inorgánicos y otros tipos de cofactores enzimáticos, se pueden agregar con el fin de optimizar la
20 actividad de enzima. El precursor de FPP se agrega a la suspensión de polipéptido, que luego se incuba a temperatura óptima, por ejemplo entre 15 y 40° C, en particular entre 25 y 35° C, más particularmente a 30° C. Después de incubación, el Drimenol producido se puede aislar de la solución incubada
25 mediante procedimientos de aislamiento estándar, tales como

extracción y destilación del disolvente, opcionalmente después de eliminación de los polipéptidos de la solución.

De acuerdo con otra realización particular, el método de cualquiera de las realizaciones descritas anteriormente se lleva a cabo *in vivo*. En este caso, la etapa a) comprende cultivar un organismo o célula anfitriona no humana capaz de producir FPP y transformada para expresar por lo menos un polipéptido que comprende una secuencia de aminoácidos por lo menos 70% idéntica a una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 y que tiene una actividad de Drimenol sintasa, bajo condiciones que conducen a la producción de Drimenol.

De acuerdo con una realización más particular, el método adicionalmente comprende, antes de la etapa a), transformar un organismo o célula no humana capaz de producir FPP con por lo menos un ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos por lo menos 70% idéntica a una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 y que tiene una actividad de Drimenol sintasa, de tal manera que dicho organismo expresa dicho polipéptido.

Estas realizaciones proporcionadas aquí son particularmente ventajosas ya que es posible llevar a cabo el método *in vivo* sin

aislar previamente el polipéptido. La reacción se produce directamente dentro del organismo o célula transformada para expresar dicho polipéptido.

5 De acuerdo con una realización más particular por lo menos un ácido nucleico utilizado en cualquiera de las realizaciones anteriores comprende una secuencia de nucleótidos que se ha obtenido al modificar la 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
10 ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de la misma. De acuerdo con otra realización, por lo menos un ácido nucleico se aísla a partir de una planta de la familia Winteraceae o la familia Canellaceae, particularmente de *Drimys winteri* o *Drimys lanceolata*.

15

El organismo o célula está destinado a "expresar" un polipéptido, siempre que el organismo o célula se transforme para albergar un ácido nucleico que codifica dicho polipéptido, este ácido nucleico se transcribe a mRNA y el polipéptido se
20 encuentra en el organismo o célula anfitriona. El término "expresa" abarca "expresa heterológamente" y "sobreexpresa", este último se refiere a los niveles de mRNA, actividad de polipéptido y/o enzima sobre y por encima de lo que se mide en un organismo o célula no transformada. Una descripción más
25 detallada de los métodos adecuados para transformar un organismo o célula anfitriona no humana se describirá más adelante en la

parte de la especificación que se dedica a dichos organismos o células anfitriones no humanos transformados.

Un organismo o célula particular está destinado a ser
5 "capaz de producir FPP" cuando se produce FPP naturalmente o cuando no produce FPP naturalmente sino que se transforma para producir FPP, ya sea antes de la transformación con un ácido nucleico como se describe aquí o junto con dicho ácido nucleico. Los organismos o células transformadas para producir una mayor
10 cantidad de FPP que el organismo o célula de origen natural también se abarcan por los "organismos o células capaces de producir FPP". Los métodos para transformar organismos, por ejemplo microorganismos, para que produzcan FPP ya se conocen en la técnica.

15

Para llevar a cabo una realización aquí *in vivo*, el organismo o célula anfitriona se cultiva bajo condiciones que conducen a la producción de Drimenol. De acuerdo con lo anterior, si el anfitrión es una planta transgénica, se
20 proporcionan las condiciones de crecimiento óptimas por ejemplo, tales como luz óptima, agua y condiciones de nutrientes. Si el anfitrión es un organismo unicelular, las condiciones que conducen a la producción de Drimenol pueden comprender adición de cofactores adecuados al medio de cultivo del anfitrión.
25 Adicionalmente, se puede seleccionar un medio de cultivo, con el fin de maximizar la síntesis de Drimenol. Las condiciones de

cultivo óptimas se describen de manera más detallada en los siguientes Ejemplos.

Los organismos anfitriones no humanos adecuados para llevar a cabo el método de una realización aquí *in vivo* pueden ser cualesquier organismos multicelulares o unicelulares no humanos. En una realización particular, el organismo anfitrión no humano utilizado para llevar a cabo una realización aquí *in vivo* es una planta, un procariota o un hongo. Se puede utilizar cualquier planta, procariota u hongo. Particularmente las plantas útiles son aquellas que, naturalmente, producen grandes cantidades de terpenos. En una realización más particular, el organismo anfitrión no humano utilizado para llevar a cabo el método de una realización aquí *in vivo* es un microorganismo. Cualquier microorganismo se puede utilizar pero de acuerdo con una realización aún más particular, dicho microorganismo es una bacteria o levadura. Más particularmente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

Algunos de estos organismos no producen FPP naturalmente. Para ser adecuado para llevar a cabo el método de una realización aquí, estos organismos se tienen que transformar para producir dicho precursor. Se pueden transformar de tal manera, ya sea antes de la modificación con el ácido nucleico descrito de acuerdo con cualquiera de las realizaciones anteriores o simultáneamente, como se explicó anteriormente.

También se pueden utilizar células eucarióticas mayores aisladas, en lugar de organismos completos, como anfitriones para llevar a cabo el método de una realización aquí *in vivo*.

5 Las células eucarióticas adecuadas pueden ser cualesquier células no humanas, pero son particularmente células de planta u hongo.

En otra realización particular, el polipéptido consiste de
10 una secuencia de aminoácidos por lo menos por lo menos 70%, particularmente por lo menos 75%, particularmente por lo menos 80%, particularmente por lo menos 85%, particularmente por lo menos 90%, particularmente por lo menos 95%, particularmente por lo menos 96%, particularmente por lo menos 97%, particularmente
15 por lo menos 98% e incluso más particularmente por lo menos 99% de identidad de secuencia con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 . En una realización incluso más particular, dicho polipéptido consiste de la SEQ ID.

20

De acuerdo con otra realización particular, por lo menos un polipéptido que tiene una actividad de Drimenol sintasa utilizado en cualquiera de las realizaciones descritas anteriormente o codificado por el ácido nucleico utilizado en
25 cualquiera de las realizaciones descritas anteriormente comprende una secuencia de aminoácidos que es una variante de la

SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante ingeniería genética, dado que dicha variante mantiene su actividad de Drimenol sintasa, como se definió anteriormente y tiene el porcentaje requerido de

5 identidad con la 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 codificado por una secuencia de nucleótidos que se ha obtenido al modificar la SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID

10 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 la misma. De acuerdo con una realización más particular, por lo menos un polipéptido que tiene una actividad de Drimenol sintasa utilizado en cualquiera de las realizaciones descritas

15 anteriormente o codificado por el ácido nucleico utilizado en cualquiera de las realizaciones descritas anteriormente consiste de una secuencia de aminoácidos que es una variante de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante ingeniería genética, es decir una secuencia de

20 aminoácidos codificado por una secuencia de nucleótidos que se ha obtenido al modificar la 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 la misma.

De acuerdo con otra realización particular, por lo menos un polipéptido que tiene una actividad de Drimenol sintasa utilizado en cualquiera de las realizaciones descritas anteriormente o codificado por el ácido nucleico utilizado en cualquiera de las realizaciones descritas anteriormente es una variante de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 que se puede encontrar naturalmente en otros organismos, tales como otras especies de planta, dado que mantiene su actividad de Drimenol sintasa como se definió anteriormente y tiene el porcentaje requerido de identidad con la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 .

Como se utiliza aquí, el polipéptido se concibe como un polipéptido o fragmento de péptido que abarca las secuencias de aminoácidos identificadas aquí, así como también polipéptidos truncados o variantes, siempre que mantengan su actividad de Drimenol sintasa como se definió anteriormente y que compartan por lo menos el porcentaje de identidad definido con el fragmento correspondiente de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14.

Ejemplos de polipéptidos variantes son proteínas de origen natural que resultan de eventos de corte y empalme de mRNA alternativos o de división proteolítica de los polipéptidos descritos aquí. Las variaciones atribuibles a proteólisis

incluyen, por ejemplo, diferencias en los terminales N o C luego de expresión en diferentes tipos de células anfitrionas, debido a la eliminación proteolítica de uno o más aminoácidos terminales de los polipéptidos de una realización aquí. Los polipéptidos codificados por un ácido nucleico obtenido por mutación natural o artificial de un ácido nucleico de una realización aquí, como se describe a continuación, también se abarca por una realización aquí.

También se pueden utilizar las variantes de polipéptido que resultan de una fusión de secuencias de péptidos adicionales en los extremos terminales amino y carboxilo en los métodos de una realización aquí. En particular, una dicha 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 entorno o sistema de expresión deseado. Dichas secuencias de péptidos adicionales pueden ser, por ejemplo péptidos de señal. De acuerdo con lo anterior, se abarcan aquí métodos que utilizan polipéptidos variantes, tales como aquellos obtenidos mediante fusión con otros oligo o polipéptidos y/o aquellos que están unidos a los péptidos de señal. Los polipéptidos resultantes de una fusión con otra proteína funcional, tal como otra proteína de la ruta de biosíntesis de terpeno, ventajosamente también se pueden utilizar en los métodos de una realización aquí.

De acuerdo con otra realización, por lo menos un polipéptido que tiene una actividad de Drimenol sintasa utilizado en cualquiera de las realizaciones descritas anteriormente o codificado por el ácido nucleico utilizado en cualquiera de las realizaciones descritas anteriormente se aísla a partir de una planta de la familia Winteraceae o la familia Canellaceae, particularmente de *Drimys winteri* o *Drimys lanceolata*.

Una herramienta importante para llevar a cabo el método de una realización aquí es el polipéptido en sí mismo. Por lo tanto se proporciona aquí un polipéptido que tiene una actividad de Drimenol sintasa y que comprende una secuencia de aminoácidos por lo menos 70% idéntica a la SEQ ID NO:2.

De acuerdo con una realización particular, el polipéptido es capaz de producir una mezcla de sesquiterpenos en los que el Drimenol representa por lo menos 20%, particularmente por lo menos 30%, particularmente por lo menos 35%, particularmente por lo menos 90%, particularmente por lo menos 95%, más particularmente por lo menos 98% de los sesquiterpenos producidos. En otro aspecto proporcionado aquí, el Drimenol se produce con más de o igual a 95%, más particularmente 98% de selectividad.

De acuerdo con una realización particular, el polipéptido comprende una secuencia de aminoácidos por lo menos 70%, particularmente por lo menos 75%, particularmente por lo menos 80%, particularmente por lo menos 85%, particularmente por lo menos 90%, particularmente por lo menos 95%, particularmente por lo menos 96%, particularmente por lo menos 97%, particularmente por lo menos 98% e incluso más particularmente por lo menos 99% idéntica a una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14 . De acuerdo con una realización más particular, el polipéptido comprende una secuencia de aminoácidos seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 y SEQ ID NO: 14.

De acuerdo con otra realización particular, el polipéptido consiste de una secuencia de aminoácidos por lo menos 70%, particularmente por lo menos 75%, particularmente por lo menos 80%, particularmente por lo menos 85%, particularmente por lo menos 90%, particularmente por lo menos 95%, particularmente por lo menos 96%, particularmente por lo menos 97%, particularmente por lo menos 98% e incluso más particularmente por lo menos 99% idéntica a una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 . De acuerdo con una realización más particular, el polipéptido consiste de un aminoácido seleccionado del grupo que

consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14.

Por lo menos un polipéptido comprende una secuencia de aminoácidos que es una variante de la SEQ ID NO:2, ya sea obtenida mediante ingeniería genética o encontrada naturalmente en plantas de *Drimys* o en otras especies de planta. En otros términos, cuando el variante de polipéptido se obtiene mediante ingeniería genética, dicho polipéptido comprende una secuencia de aminoácidos codificado por una secuencia de nucleótidos que se ha obtenido al modificar la 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 la misma. De acuerdo con una realización más particular, por lo menos un polipéptido que tiene una actividad de Drimenol sintasa consiste de una secuencia de aminoácidos que es una variante de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 obtenida mediante ingeniería genética, es decir una secuencia de aminoácidos codificado por una secuencia de nucleótidos que se ha obtenido al modificar la 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 la misma.

De acuerdo con otra realización, el polipéptido se aísla a partir de una planta de la familia Winteraceae o la familia

Canellaceae, particularmente de *Drimys winteri* o *Drimys lanceolata*. Como se utiliza aquí, el polipéptido se concibe como un polipéptido o fragmento de péptido que abarca la secuencia de aminoácidos identificada aquí, así como también polipéptidos truncados o variantes, siempre que mantengan su actividad como se definió anteriormente y que compartan por lo menos el porcentaje de identidad definido con el fragmento correspondiente de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14.

10

Como se mencionó anteriormente, el ácido nucleico que codifica el polipéptido de una realización aquí es una herramienta útil para modificar organismos o células anfitriones no humanos destinados a ser utilizados cuando el método se lleva a cabo *in vivo*.

15

Por lo tanto también se proporciona aquí un ácido nucleico que codifica un polipéptido de acuerdo con cualquiera de las realizaciones anteriormente.

20

De acuerdo con una realización particular, el ácido nucleico comprende una secuencia de nucleótidos por lo menos 50%, particularmente por lo menos 55%, particularmente por lo menos 60%, particularmente por lo menos 65%, particularmente por lo menos 70%, particularmente por lo menos 75%, particularmente por lo menos 80%, particularmente por lo menos 85%,

25

particularmente por lo menos 90%, más particularmente por lo menos 95% particularmente por lo menos 96%, particularmente por lo menos 97%, particularmente por lo menos 98%, e incluso más particularmente por lo menos 99% idéntica a una secuencia

5 seleccionada del grupo que consiste de una secuencia seleccionada del grupo que consiste de 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 la misma. De acuerdo con una realización más particular, el

10 ácido nucleico comprende la secuencia de nucleótidos seleccionada del grupo que consiste de 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 de la misma.

15 De acuerdo con otra realización particular, el ácido nucleico consiste de una secuencia de nucleótidos por lo menos 70%, particularmente por lo menos 75%, particularmente por lo menos 80%, particularmente por lo menos 85%, particularmente por lo menos 90%, particularmente por lo menos 95%, particularmente

20 por lo menos 96%, particularmente por lo menos 97%, particularmente por lo menos 98% e incluso más particularmente por lo menos 99% o más identidad con una secuencia seleccionada del grupo que consiste de la 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

25 SEQ ID NO: 12, SEQ ID NO: 13 o SEQ ID NO: 15 o el complemento de la misma.

De acuerdo con una realización incluso más particular, el ácido nucleico consiste de una secuencia seleccionada del grupo que consiste de la 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 la misma.

El ácido nucleico de una realización aquí se puede definir como que incluye polímeros de desoxirribonucleótidos o ribonucleótidos ya sea en forma o cadena sencilla o doble (ADN y/o ARN). El término "secuencia de nucleótidos" también se debe entender como que comprende una molécula de polinucleótido o una molécula de oligonucleótido en forma de un fragmento separado o como un componente de un ácido nucleico más grande. Los ácidos nucleicos de una realización aquí también abarcan ciertas secuencias de nucleótidos aisladas que incluyen aquellas que están sustancialmente libres de material contaminante endógeno. El ácido nucleico de una realización aquí se puede truncar, siempre que codifique un polipéptido abarcado aquí, como se describió anteriormente.

En una realización, el ácido nucleico de una realización aquí puede estar presente ya sea en naturalmente en plantas de la especie *Drimys* u otras especies, o se puede obtener al modificar la 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 la misma. Particularmente dicho ácido nucleico consiste de una secuencia de nucleótidos que se ha obtenido al modificar la 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 la misma.

Los ácidos nucleicos que comprenden una secuencia obtenida mediante mutación 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 la misma se abarcan mediante una realización aquí, siempre que las secuencias comprendan compartir por lo menos el porcentaje de identidad definido con los fragmentos correspondientes de la 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 los mismos y siempre que codifiquen un polipéptido que tiene una actividad de Drimenol sintasa, como define en cualquiera de las realizaciones anteriores. Las mutaciones pueden ser cualquier clase de mutaciones de estos ácidos nucleicos, tales como mutaciones puntuales, mutaciones de supresión, mutaciones de inserción y/o mutaciones de desplazamiento de marco. Se puede preparar un ácido nucleico variante con el fin de adaptar su secuencia de nucleótidos a un sistema de expresión específico. Por ejemplo, se conocen sistemas de expresión bacterianos que expresan de

manera más eficiente polipéptidos si se codifican aminoácidos por codones particulares.

Debido a la degeneración del código genético, más de un
5 codón puede codificar la misma secuencia de aminoácidos, múltiples secuencias de ácidos nucleicos pueden codificar la misma proteína o polipéptido, todas estas secuencias de ADN que se abarcan por una realización aquí. Cuando sea apropiado, las secuencias de ácido nucleico que codifican la Drimenol sintasa
10 se pueden optimizar para aumentar la expresión en la célula anfitriona. Por ejemplo, los nucleótidos de una realización aquí se pueden sintetizar utilizando codones particulares por un anfitrión para mejor expresión.

15 Otra herramienta importante para transformar organismos o células anfitrionas adecuadas para llevar a cabo el método de una realización aquí *in vivo* es un vector de expresión que comprende un ácido nucleico de acuerdo con cualquier modalidad de una realización aquí. Por lo tanto también se proporciona
20 aquí dicho vector.

Se pueden utilizar los vectores de expresión proporcionados aquí en los métodos para preparar un organismo y/o célula anfitrión transformada genéticamente, en organismos y/o células
25 anfitriones que albergan los ácidos nucleicos de una realización aquí y en los métodos para elaborar polipéptidos que tienen una

actividad de Drimenol sintasa, como se describe adicionalmente adelante.

Los organismos y células anfitriones no humanos
5 recombinantes transformados para albergar por lo menos un ácido nucleico de una realización aquí de tal manera que exprese o sobre-exprese de forma heteróloga por lo menos un polipéptido de una realización aquí también son herramientas muy útiles para llevar a cabo el método de una realización aquí. Por lo tanto
10 también se proporcionan aquí organismos y células anfitriones no humanos.

Se puede utilizar un ácido nucleico de acuerdo con cualquiera de las realizaciones anteriores para transformar los
15 organismos y células anfitriones no humanos y el polipéptido expresado puede ser cualquiera de los polipéptidos descritos anteriormente.

Los organismos anfitriones no humanos de una realización
20 aquí pueden ser cualquiera de los organismos multicelulares o unicelulares no humanos. En una realización particular, el organismo anfitrión no humano es una planta, un procariota o un hongo. Cualquier planta, procariota o hongo es adecuado para ser transformado de acuerdo con los métodos proporcionados aquí. Las
25 plantas particularmente útiles son aquellas que naturalmente producen altas cantidades de terpenos.

En una realización más particular, el organismo anfitrión no humano es un microorganismo. Cualquier microorganismo es adecuado para ser utilizado aquí, pero de acuerdo con una
5 realización incluso más particular, dicho microorganismo es una bacteria o levadura. Más particularmente, dicha bacteria es *E. coli* y dicha levadura es *Saccharomyces cerevisiae*.

También se pueden transformar células eucarióticas mayores
10 aisladas, en lugar de organismos completos. Como células eucariotas mayores, entendemos aquí cualquier célula eucariota no humana excepto células de levadura. Las células eucariotas mayores particulares son células vegetales o células fúngicas.

15 Una variante también puede diferir del polipéptido de una realización aquí mediante adhesión de grupos de modificación que se unen de forma covalente o no covalente a la estructura principal del polipéptido. La variante también incluye un polipéptido que difiere del polipéptido descrito aquí por sitios
20 de glicosilación unidos a N o unidos a O introducidos, y/o una adición de residuos de cisteína. El experto reconocerá cómo modificar una secuencia de aminoácidos y conservar la actividad biológica.

25 La funcionalidad o actividad de cualquier proteína de Drimenol sintasa, variante o fragmento, se pueden determinar

utilizando diversos métodos. Por ejemplo, se puede utilizar sobreexpresión transitoria o estable en células vegetales, bacterianas o de levadura para probar si la proteína tiene actividad, es decir, produce Drimenol a partir de precursores de FPP. Se puede evaluar la actividad de Drimenol sintasa en un sistema de expresión microbiano, tal como el ensayo descrito en el Ejemplo 2 o 3 aquí en la producción de Drimenol, lo que indica funcionalidad. Una variante o derivado de un polipéptido de Drimenol sintasa de una realización aquí conserva la capacidad de producir Drimenol a partir de precursores FPP. Las variantes de la secuencia de aminoácidos de las Drimenol sintasas proporcionadas aquí pueden tener funciones biológicas deseables adicionales, que incluyen, por ejemplo, utilización de sustrato alterado, cinéticas de reacción, distribución del producto u otras alteraciones.

Una realización aquí proporciona polipéptidos de una realización aquí que se van a utilizar en un método para producir Drimenol al poner en contacto un precursor de FPP con los polipéptidos de una realización aquí ya sea *in vitro* o *in vivo*.

También se proporciona aquí un polinucleótido aislado, recombinante o sintético que codifica un polipéptido o variante de polipéptido proporcionado aquí.

Una modalidad de una realización aquí proporciona una secuencia de ácidos nucleicos aislada, recombinante o sintética de la 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 una variante de la misma que codifica una Drimenol sintasa que tiene la secuencia de aminoácidos que es por lo menos 70%, 75%, 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% o 99% idéntica a una secuencia de aminoácidos seleccionada del grupo que consiste de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 o SEQ ID NO: 14 o fragmentos de la misma que cataliza la producción de Drimenol en una célula a partir de un precursor de FPP. Las realizaciones proporcionadas aquí incluyen, pero no se limitan a cADN, ADN genómico y secuencias de ARN. Cualesquier secuencia de ácidos nucleicos que codifica la Drimenol sintasa o variantes de la misma por lo tanto se menciona aquí como una secuencia de codificación de Drimenol sintasa.

De acuerdo con una realización particular, el ácido nucleico de la 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 SEQ es la secuencia de codificación de un gen de Drimenol sintasa que codifica la Drimenol sintasa obtenida como se describe en los ejemplos.

Un fragmento de un polinucleótido de la 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 por lo menos 15 bp, por lo menos 30 bp, por lo menos 40 bp, por lo menos 50 bp y/o por lo menos 60 bp en longitud del polinucleótido de una realización aquí. Particularmente el fragmento de un polinucleótido comprende por lo menos 25, más particularmente por lo menos 50, más particularmente por lo menos 75, más particularmente por lo menos 100, más particularmente por lo menos 150, más particularmente por lo menos 200, más particularmente por lo menos 300, más particularmente por lo menos 400, más particularmente por lo menos 500, más particularmente por lo menos 600, más particularmente por lo menos 700, más particularmente por lo menos 800, más particularmente por lo menos 900, más particularmente por lo menos 1000 nucleótidos contiguos del polinucleótido de una realización aquí. Sin estar limitado, el fragmento de los polinucleótidos aquí se puede utilizar como una sonda PCR, y/o como una sonda, o silenciamiento de gen anticodificante o iARN.

Es claro para el experto en la técnica que los genes, que incluyen los polinucleótidos de una realización aquí, se pueden clonar en base de la información de la secuencia de nucleótidos disponible, tale como se encuentra en la lista de secuencias

adjunta, por métodos conocidos en la técnica. Estos incluyen, por ejemplo, el diseño de cebadores de ADN que representan las secuencias flanqueantes de dicho en el cual uno se genera en orientaciones codificantes y que inicia la síntesis de la cadena codificante y el otro se crea de forma complementaria inversa y genera la cadena anticodificante. Las polimerasas de ADN termoestables, tales como aquellas utilizadas en la reacción en cadena de la polimerasa se utilizan comúnmente para llevar a cabo dichos experimentos.

10

Alternativamente, las secuencias de ADN que representan genes se pueden sintetizar químicamente y se introducen posteriormente en las moléculas de vector de ADN que se pueden multiplicar por ejemplo bacterias compatibles tales como, por ejemplo, *E. coli*.

15

En una realización relacionada de una realización aquí, se proporcionan cebadores y/o sondas PCR para detectar secuencias de ácidos nucleicos que codifican una Drimenol sintasa. El experto en la técnica será consciente de los métodos para sintetizar pares de cebadores de PCR degenerados o específicos para amplificar una secuencia de ácidos nucleicos que codifica la Drimenol sintasa o fragmentos de la misma, con base en la 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. Un equipo de detección de secuencias de ácidos

20

25

nucleicos que codifica la Drimenol sintasa puede incluir cebadores y/o sondas específicas para secuencias de ácidos nucleicos que codifican la Drimenol sintasa, y un protocolo asociado para utilizar los cebadores y/o sondas para detectar

5 secuencias de ácidos nucleicos que codifican la Drimenol sintasa en una muestra. Dichos equipos de detección se pueden utilizar para determinar si se ha modificado una planta, es decir, transformado con una secuencia que codifica la Drimenol sintasa.

10 Se proporcionan aquí secuencias de ácidos nucleicos obtenidos por mutaciones de la 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 pueden hacer de forma rutinaria dichas mutaciones. Es claro para el experto

15 en la técnica que las mutaciones, supresiones, inserciones y/o sustituciones de uno o más nucleótidos se pueden introducir en la secuencia de ADN 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. En general, una mutación

20 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 secuencias de ADN variante de acuerdo con una realización aquí, la secuencia de interés se une

25 operativamente a un gen marcador seleccionable o detectable y la expresión del gen indicador se prueba en ensayos de expresión

transitoria con protoplastos o en las plantas transformadas de forma estable. El experto en la técnica reconocerá que las secuencias de ADN capaces de dirigir la expresión se construyen como módulos. De acuerdo con lo anterior, los niveles de

5 expresión de fragmentos de ADN más cortos puede ser diferentes de los del fragmento más largo y pueden ser diferentes unos de otros. Adicionalmente también se proporciona aquí equivalentes funcionales de la secuencia de ácidos nucleicos que codifica las proteínas de Drimenol sintasa, es decir, secuencias de

10 nucleótidos que se hibridan bajo condiciones rigurosas con la secuencia de ácidos nucleicos de la 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. El experto en la técnica será consciente de los métodos para identificar

15 secuencias homólogas en otros organismos y métodos (identificados en la sección de Definición aquí) para determinar el porcentaje de identidad de secuencia entre secuencias homólogas. Dichas moléculas de ADN recién identificadas, luego se pueden secuenciar y la secuencia se puede comparar con la

20 secuencia de ácidos nucleicos de la 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 y se prueba para equivalencia funcional. Se proporcionan aquí moléculas de ADN que tiene por lo menos 70% particularmente 75%,

25 particularmente 80%, particularmente 85%, particularmente 90%, particularmente 95%, particularmente 96% particularmente 97%

particularmente 98%, o más particularmente 99% o más de identidad de secuencia con la secuencia de nucleótidos de la 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
5 SEQ ID NO: 15

Una realización relacionada proporciona una secuencia de ácidos nucleicos que es complementaria a la secuencia de ácidos nucleicos de acuerdo con SEQ ID NO: 1 o la SEQ ID NO: 3, tal
10 como ARN inhibidores, o secuencia de ácidos nucleicos que se hibrida bajo condiciones rigurosas con por lo menos parte de la secuencia de nucleótidos de acuerdo con la 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

15

Una realización alternativa proporcionada en este documento proporciona un método para alterar la expresión de genes en una célula anfitriona. Por ejemplo, el polinucleótido de una realización aquí se puede mejorar o sobreexpresar o inducir en
20 ciertos contextos (por ejemplo después de picaduras o agujones de insectos o por exposición a una cierta temperatura) en una célula anfitriona u organismo anfitrión.

La alteración de la expresión de un polinucleótido
25 proporcionado aquí también resulta en "expresión ectópica", que es un patrón de expresión diferente en un organismo alterado y

en un control u organismo tipo silvestre. La alteración de la expresión se produce a partir de interacciones de polipéptido de una realización aquí con moduladores exógenos o endógenos, o como resultado de la modificación química del polipéptido. El
5 término también se refiere a un patrón de expresión alterada del polinucleótido de una realización aquí que se altera por debajo del nivel de detección o actividad completamente suprimida.

En una realización, varias secuencias de ácidos nucleicos
10 que codifican Drimenol sintasa se co-expresan en un único anfitrión, particularmente bajo el control de diferentes promotores. Alternativamente, pueden estar presentes varias proteínas de sintasa Drimenol que codifican secuencias de ácidos nucleicos sobre un único vector de transformación o se co-
15 transforma al mismo tiempo utilizando vectores separados y que seleccionan transformantes que comprenden tanto los genes quiméricos. Del mismo modo, una o más Drimenol sintasa que codifican genes se pueden expresar en una sola planta junto con otros genes quiméricos, por ejemplo que codifican otras
20 proteínas que mejoran la resistencia a las plagas de insectos, u otros.

Las secuencias de ácidos nucleicos de una realización aquí que codifican proteínas de Drimenol sintasa se pueden insertar
25 en vectores de expresión y/o estar contenidas en genes quiméricos insertados en vectores de expresión, para producir

proteínas de Drimenol sintasa en una célula anfitriona u organismo anfitrión. Los vectores para la inserción de transgenes en el genoma de células anfitrionas son bien conocidos en la técnica e incluyen plásmidos, virus, cósmidos y cromosomas artificiales. Los vectores binarios o de co-integración en los que se inserta un gen quimérico también se utilizan para la transformación de células anfitrionas.

Una realización de la proporcionada en el presente documento proporciona vectores de expresión recombinantes que comprenden una secuencia de ácidos nucleicos de un gen de Drimenol sintasa, o un gen quimérico que comprende una secuencia de ácidos nucleicos de un gen de Drimenol sintasa, unido operativamente a secuencias de ácidos nucleicos asociadas, tales como, por ejemplo, secuencias de promotor. Por ejemplo, un gen quimérico que comprende una secuencia de ácidos nucleicos de la SEQ ID NO: 1 o SEQ ID NO: 3 se pueden unir operativamente a una secuencia de promotor adecuada para la expresión en células vegetales, células bacterianas o células fúngicas, opcionalmente unidas a una secuencia de ácidos nucleicos no traducida 3'.

Alternativamente, la secuencia de promotor puede estar ya presente en un vector de tal manera que la secuencia de ácidos nucleicos que se van a transcribir se inserta en el vector en dirección 3' de la secuencia de promotor. Los vectores

normalmente se diseñan para tener un origen de replicación, un sitio de clonación múltiple, y un marcador seleccionable.

Los siguientes ejemplos son sólo ilustrativos y no pretenden limitar el alcance de las reivindicaciones o realizaciones proporcionadas aquí.

Ejemplos

10 Ejemplo 1

Material de planta de *Drimys lanceolata* y *Drimys winteri* y secuenciación del transcriptoma de hoja,

Se obtuvieron plantas de *Drimys winteri* y de Bluebell Nursery (Leicestershire, Reino Unido). Para análisis de la composición en moléculas de terpeno, se recolectaron las hojas y el disolvente se extrajo por medio de MTBE (metil terc-butil éter). El extracto se analizó por GCMS utilizando un sistema Agilent 6890 Series GC conectado a un detector de masas Agilent 5975. El GC estaba equipado con un diámetro interno de 0.25 mm en una columna capilar DB-LMS de 30 m (Agilent). El gas portador fue He en un flujo constante de 1 mL/min. La temperatura de horno inicial fue de 50° C (se mantuvo 1 min) seguido de un gradiente de 10° C/min a 300° C. La inyección se elaboró en un inyector de división/sin división establecido en 260° C y utilizado en modo sin división. La identificación de los

productos se basó en la comparación de los índices de espectros de masa y retención con estándares auténticos y bases de datos de espectros de masa interna. Las hojas de las dos plantas contenían cantidades significativas de compuestos de drimane sesquiterpeno que incluyen (-)-drimenol, poligodial y epipoligodial (Figura 1).

De esta manera se tomaron hojas pequeñas de *D. winteri* y *D. lanceolata* para análisis del transcriptoma. Se extrajo ARN total por medio del Reactivo de ARN de Planta Concert™ Reactivo ARN (Invitrogen). Este ARN total se procesó utilizando la técnica de ARN-Sec total Illumina y el secuenciador 2000 Illumina HiSeq. Un total de 101 y 105 millones de lecturas pares de 2x100 bp se generaron para *D. winteri* y *D. lanceolata*, respectivamente. Las lecturas se ensamblaron con el ensamblador genómico Velvet de novo (<http://www.ebi.ac.uk/~zerbino/velvet/>) y el software Oasis (<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 obtuvieron 28'255 cóntigos con un tamaño promedio de 1'179 pb. Los cóntigos se buscaron utilizando el algoritmo tBlastn (Altschul et al, J. Mol. Biol. 215, 403-410, 1990) y utilizando como consulta las secuencias de aminoácidos de sintasas de sesquiterpeno conocidas. Este método proporciona las secuencias de 37 nuevas sintasas de sesquiterpenos putativos. La actividad enzimática de estas sintasas se evaluó

como se describe en los ejemplos siguientes para las sintasas que muestran actividad de Drimenol sintasa.

Ejemplo 2. Expresión funcional y caracterización de D1Tps589 de
D. lanceolata.

Las secuencias de ADN de una de las sintasas D1Tps589 de sesquiterpeno seleccionadas se optimizaron con codones, 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 D1Tps589 se realizó en células de *E. coli* KRX (Promega). Las colonias individuales de células transformadas con el plásmido de expresión pJ444SR-D1Tps589 se utilizaron 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° y se dejaron 1 hora para el equilibrio. La expresión de la proteína luego se indujo por la adición de IPTG 1 mM y 0.2% de L-ramnosa y el cultivo se incubó durante la noche a 20° C. El día siguiente, las células se recolectaron mediante centrifugación, se resuspendieron en 0.1 volumen de MOPSO 50 mM pH 7, 10% de glicerol y se lisaron mediante sonicación. Los extractos se limpiaron mediante centrifugación (30 min a 20,000 g) y los sobrenadantes que contienen las proteínas solubles se utilizaron para experimentos adicionales.

Los extractos de proteína *E. coli* cruda que contienen la proteína recombinante se utilizaron para la caracterización de 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 de 0.1 a 0.5 mg de proteína cruda. Los tubos se incubaron 12 a 24 horas a 30° C y se extrajeron dos veces con un volumen de pentano. Después de concentración bajo un flujo de nitrógeno, los extractos se analizaron mediante GC y GC-MS y se compararon con extractos de ensayos con proteínas de control. El análisis de los productos formados por las enzimas fue hecho por GCMS como se describe en el ejemplo 1. Un control negativo se realizó en las mismas condiciones utilizando células de *E. coli* transformadas con un plásmido pJ444 vacío. En estas condiciones, la enzima D1Tps589 recombinante producida (-)-drimenol como producto principal con una selectividad mayor de 98% (Figura 2). La identidad del (-)-drimenol se confirmó por coincidencia del espectro de masas y el tiempo de retención de un estándar auténtico de Drimenol aislado de aceite de sándalo del oeste (*Amyris balsamifera*).

Ejemplo 3. Producción in vivo de (-)-drimenol en células de *E. coli* utilizando D1Tps589.

Para evaluar la producción *in vivo* de (-)-drimenol en células heterólogas, las células de *E. coli* se transformaron con el plásmido de expresión pJ444SR-D1Tps589 y se evaluó la

producción de sesquiterpenos del grupo FPP endógeno. Para aumentar la productividad de las células, una sintasa de FPP heteróloga y una de las enzimas de una ruta de mevalonato heterólogo completo (MVA) también se expresaron en las mismas

5 células. La construcción del plásmido de expresión que contiene un gen de sintasa FPP y el gen para una ruta de MVA completa se describe en la patente WO2013064411 o en Schalk et al (2013) J. Am. Chem. Soc. 134, 18900-18903. En resumen, se preparó un plásmido de expresión que contiene dos operones compuestos de

10 los genes que codifican las enzimas para una ruta de mevalonato completa. Un primer operón sintético que consiste en genes de acetoacetyl CoA tiolasa-(atoB) de *E. coli*, HMG-CoA de *Staphylococcus aureus* (mvas), reductasa HMG-CoA de *Staphylococcus aureus* (MvaA) y sintasa FPP de *Saccharomyces*

15 *cerevisiae* (ERG20) se sintetizan *in vitro* (DNA2.0, Menlo Park, CA, EE.UU.) y se une en el vector *NcoI-BamHI* digerido pACYCDuet-1 (Invitrogen) dando pACYC-29258. Un segundo operón que contiene una quinasa mevalonato (MvaK1), una fosfomevalonato quinasa (MvaK2), una descarboxilasa mevalonato de difosfato (MVAD), y

20 una isomerasa de difosfato de isopentenilo (idi) se amplificó a partir de ADN genómico de *Streptococcus pneumoniae* (ATCC BAA-334) y se ligó en el segundo sitio de clonación múltiple de pACYC-29258 proporcionando el plásmido pACYC-29258-4506. Por lo tanto este plásmido contiene los genes que codifican todas las

25 enzimas de la ruta biosintética que conduce de acetyl-coenzima A a FPP.

Las células *E. coli* KRX (Promega) se co-transformaron con el plásmido pACYC-29258-4506 y el plásmido pJ444SR-DlTps589. Las células transformadas se seleccionaron sobre placas de LB-
5 agarosa con carbenicilina (50 µg/ml) y cloranfenicol (34 µg/ml). Las colonias individuales se utilizaron para inocular 5 mL de medio líquido LB complementado con los mismos antibióticos. El cultivo se incubó durante la noche a 37° C. Al día siguiente, 2 mL de medio TB complementado con los mismos antibióticos se
10 inocularon con 0.2 mL del cultivo durante la noche. Después de 6 horas de incubación a 37° C, el cultivo se enfrió a 28° C y se agregaron IPTG 0.1 mM y 0.2% de ramnosa a cada tubo. Los cultivos se incubaron durante 48 horas a 28° C. Los cultivos luego se extrajeron dos veces con 2 volúmenes de MTBE, la fase
15 orgánica se concentró a 500 µL y se analizó mediante GC-MS como se describió anteriormente en el Ejemplo 1.

En estas condiciones *in vivo* la enzima recombinante DlTps589 produce (-)-drimenol como producto importante con la
20 misma selectividad aparente como en el ensayo *in vitro* descrito en el ejemplo 2 (Figura 3). Utilizando estas células de *E. coli* más grandes modificadas por ingeniería (1 L), los cultivos se utilizaron para purificar el sesquiterpeno producido por la enzima en cantidad suficiente para confirmar la estructura
25 mediante análisis de RMN y medición rotación específica cuando la estructura de (-)-drimenol se muestra en la figura 4. La

pureza enantiomérica se confirmó mediante análisis GC quirál en un sistema GC Varian CP-3800 equipado con una columna Chirasil (Agilent) y utilizando temperatura de gradiente de horno de 3.0° C/min desde 125 hasta 180° C (Figura 5).

5

Ejemplo 4. Expresión funcional y caracterización de SCH51-3228-9 y SCH51-3228-11 a partir de *D. winteri*.

La SCH51-3228-9 y SCH51-3228-11 son otras dos secuencias de ADN que supuestamente codifican sintasas de sesquiterpeno y se
10 aíslan de las secuencias del transcriptoma de *Drymis winteri*. Las secuencias de aminoácidos deducidas comparten 92.6 y 95.1% de identidad, respectivamente con la secuencia de aminoácidos D1Tps589. Las dos secuencias se sintetizaron con codones, sintetizaron *in vitro* (Invitrogen) y se clonaron entre los
15 sitios de reconocimiento de enzimas de restricción *NdeI* y *KpnI* 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)
20 (Invitrogen). Las colonias individuales de células transformadas con plásmidos de expresión pETDuet-SCH51-3228-9 o pEDTDuet-SCH51-3228-11 se utilizaron para producir las enzimas recombinantes como se describe en el Ejemplo 2. Los extractos de proteína de *E. coli* crudos que contienen las proteínas
25 recombinantes se utilizaron para la caracterización de las actividades enzimáticas como se describe en el ejemplo 2 a

excepción de las de las condiciones de análisis GCMS que se realizaron como sigue. El análisis GC-MS se realizó en un sistema Agilent 6890 Serie GC conectado a un detector de masas Agilent 5975. El GC se equipó con diámetro interno de 0.25 mm por columna capilar DB-lms 30 m (Agilent). El gas portador fue He en un flujo constante de 1 mL/min. La temperatura de horno inicial fue de 50° C (se mantuvo 5 min) seguido de un gradiente de 5° C/min a 300° C. La inyección se hizo en modo de división a 250° C con una relación de división de 5: 1.

10

Las dos enzimas recombinantes producen (-)-drimenol como producto importante con alta selectividad. La identidad del (-)-drimenol se confirmó mediante coincidencia del espectro de masas y el tiempo de retención de un estándar de Drimenol auténtico aislado a partir de aceite de sándalo del oeste (*Amyris balsamifera*) (Figura 6).

15

Utilizando sistema de células de *E. Coli* completo y el método descrito en el ejemplo 3 (a excepción de las condiciones de análisis GCMS que fueron como se describió anteriormente) el Drimenol también se podría producir *in vivo* en cultivos de bacterias utilizando proteínas recombinantes SCH51-3228-9 y SCH51-3228-11 (Figura 7).

20

Ejemplo 5. Expresión funcional y caracterización de SCH51-998-28 a partir de *D. winteri* y SCH52-13163-6 a partir de *D. lanceolata*.

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

Las proteínas recombinantes se produjeron en células de *E. coli* BL21 (DE3) (Invitrogen) y los ensayos in vitro que utilizan FPP como sustrato se realizaron como se describe en el ejemplo 2 y 4. Estos ensayos mostraron actividad de Drimenol sintasa para SCH51-998-28 y SCH52-13163-6 (figura 6). Utilizando células de *E. coli* se sobreproduce FPP desde una ruta de mevalonato recombinante (ejemplo 2 y 4), el Drimenol también se podría producir in vivo utilizando las proteínas SCH51-998-28 y SCH52-13163-6 (Figura 7).

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

Las secuencias de aminoácidos de las Drimenol sintasas de *Drimys winteri* y *Drimys lanceolata* se alinearon utilizando el programa de alineación múltiple ClustalW (Thompson et al, 1994, Nucleic Acid Res. 22 (22), 4673-4680 y las identidades de secuencia se calcularon basándose en esta alineación.

Porcentaje de identidad (%) entre las diferentes Drimenol sintasas de especies *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

5 - Listado de secuencias -

SEQ ID NO.: 1

SCH52-ctg589, Marco de lectura abierto de C1Tps589 de *D. lanceolata*, secuencia de ADN tipo silvestre.

ATGGATCTTATTAATCCCTCCCCAGCGGCTTCCACCCTCCCTCTCCCAGTTGATGGAGATTC
 AGAAGTTGTTAGGCGATCTGCCGGGTTTCATCCGACTATCTGGGGCGATCACTTCCTCTCCT
 ACAAGCCCGATCCAAAGAAAATAGATGCATGGAATAAAAGGGTTGAAGAGCTGAAGGAAG
 AAGTGAAGAAGATATTAAGCAATGCAAAGGGACGGTGGAAGAGCTGAATTTGATTGATG
 ATCTCGTACACCTTGGGATTAGTTATCATTTTGAGAAGGAGATTGATGATGCTCTACAACAC
 ATCTTTGATACCCATCTTGATGATTTTCCTAAGGATGATCTATATGTCGCCGCTCTCCGATTT
 GGCGTCTTAAGGAAACAGGGGCACCGTGTTTCTCCAGATGTATTCAAAAATTCAAAGATG
 AGCAGGGGAATTTCAAGGCAGAGTTGAGCACCGATGCGAAAGGTTTGCTATGTTTAAATGA
 TGTGGCTTATCTCAGCACAAGAGGGGAAGATATCTTGGATGAAGCCATTCTTTCACTGAG
 GAGCACCTTAGGTCTTGTATTAGCCATGTAGATTCTCATATGGCAGCAAAAATTGAACATTC
 TCTCGAGCTTCCCCTTCATCATCGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCT
 ATGAAGGAGACAAGGAAAGGAACGAAGTTGTCCTTGAGCTTGCCAATTTAGATTTCAATCT
 GATTCAAATCTTGCACCAAAGAGAGCTGAGAGACATCACAATGTGGTGGAAGGAGATTGA
 CCTTGCAGCAAAGCTGCCTTTTATTAGGGATAGGTTGGTGGAGTGCTACTACTGGATCATG
 GGGGTCTATTTTGAACCAATATACTCGAGGGCTAGGGTTTTTCCACCAAATGACAATGTT
 GGTCTCAGTTGTGGACGACATATATGATGTGTATGCTACCGAGGATGAGCTTCAACTATTC
 ACTGATGCCATCTATAGGTGGGATGCTGATGACATTGATCAGCTGCCTCAGTACTTGAAAG
 ATGCTTTTATGGTACTCTACAACACTGTGAAGACTCTAGAAGAAGAAGTGAACCAGAAGG
 AACTCTTATCGTGGATTCTATGTAAAAGATGCAATGAAGGTTTTGGCAAGGGATTACTTT
 GTGGAGCACAAATGGTATAACAGAAAAATTGTGCCATCCGTAGAGGAATACTTGAAAATTT
 CTTGCATCAGTGTGGCCGTTTCATATGGCTACAGTTCACTGTATTGCTGGGATGTATGAAATT
 GCAACCAAAGAGGCATTGCAATGGTTGATGACTGAGCCCAAAGTGTATTGATGCATCTC
 TGATTGGTCGTCTCCTTGATGACATGCAGTCCACCTCGTTTGAGCAACAGAGAGGCCACGT
 GTCATCAGCAGTACAGTGTTACATGGCTGAATATGGTGTAACAGCGGAAGAAGCATGTGAA
 AAGCTCCGAGATATGGCTGCAATTGCTTGGAAGATGTGAACGAGGCATGCCTTAGGCCCA
 CGGTTTTCCCTATGCCTATCCTTTTGCTTCTATCAACTTGGCACGTGTGGCAGAAGTCATCT
 ACCTACGTGGAGATGGATACACGCACGCTGGGGGTGAGACCAAGAAACACATCACGGCCA
 TGCTTGTTAAGCCAATTGAAGTCTGA

SEQ ID NO.: 2

5 ClTps589 de *D. lanceolata*, secuencia de aminoácidos.

MDLINPSAASTLPLPVDGDSEVVRRSAGFHPTIWGDHFLSYKPDPKKIDAWNKRVEELKEEVK
 KILSNAKGTVEELNLIDDLVHLGISYHFEKEIDDALQHIFDTHLDDFPKDDLYVAALRFGVLRKQ
 GHRVSPDVFKFKDEQGNFKAELSTDAKGLLCLNDVAYLSTRGEDILDEAIPFTEEHLRSCISHV
 DSHMAAKIEHSLELPLHHRIPRLNRHYISVYEGDKERNEVVLELANLDFNLIQHQLHRELDRIT
 MWWKEIDLAAKLPFIRDRLVECYWIMGVYFEPIYSRARVFSTKMTMLVSVVDDIYDVYATE
 DELQLFTDAIYRWDADDIDQLPQYLKDAFMVLYNTVKTLEEELEPEGNSYRGFYVKDAMKVL
 ARDYFVEHKWYNRKIVPSVEEYLKISCISVAVHMATVHCIAAGMYEIATKEAFEWLMTPEKLVID
 ASLIGRLLDDMQSTSFEQQRGHVSSAVQCYMAEYGVTAEEACEKLRDMAAIAWKDVNEACLR
 PTVFPMILLPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVKPIEV

SEQ ID NO.: 3

- 5 D1Tps589_opt, secuencia de ADN optimizada con codón de D1Tps589 de *D. lanceolata*.

ATGGACCTGATTAACCCGAGCCCTGCTGCATCCACCCTGCCACTGCCAGTCGATGGTGATA
 GCGAAGTTGTGCGCCGTAGCGCGGGTTTCCATCCGACCATCTGGGGTGACCACTTTCTGTCT
 TATAAGCCGGACCCGAAAAAGATTGATGCGTGGAACAAGCGTGTTGAGGAACTGAAAGAA
 GAGGTCAAAAAGATTTTGAGCAATGCGAAAGGCACGGTTGAGGAACTGAATTTGATTGAC
 GACCTGGTACACCTGGGTATTAGCTATCACTTTGAGAAAAGAAATCGACGACGCGCTGCAGC
 ATATCTTCGATACGCACCTGGATGATTTCCCGAAAGATGACCTCTACGTGGCTGCGCTGCGT
 TTTGGCGTCCTGCGTAAGCAAGGCCATCGTGTCAGCCCGGACGTCTTTAAGAAATTCAAAG
 ACGAGCAAGGCAACTTCAAAGCGGAGCTGTCAACCGATGCAAAGGGCCTGTTGTGCCTGA
 ACGATGTGGCGTACCTGAGCACCCGTGGTGAGGATATCCTGGACGAAGCGATCCCGTTCAC
 GGAAGAACATTTGCGCTCGTGCAATTAGCCACGTTGATAGCCACATGGCAGCGAAGATTGAG
 CACTCTCTGGAGCTGCCGCTGCACCATCGCATTCCGCGTTTAGAGAATCGCCATTACATCTC
 CGTGTACGAGGGTGACAAAGAGCGTAATGAAGTCGTTCTGGAGTTGGCTAACTTGGACTTT
 AATCTTATCCAGATCCTGCACCAGCGCGAGCTGCGCGACATCACGATGTGGTGGAAGAAA
 TTGATCTGGCCGCAAAGCTGCCGTTTATTCGTGACCGTCTGGTGGAGTGTTACTATTGGATT

ATGGGCGTGTACTTCGAGCCGATCTACAGCCGTGCGCGCGTGTTTAGCACCAAGATGACCA
 TGCTGGTTAGCGTGGTGGATGACATCTATGATGTCTACGCTACGGAAGATGAGTTGCAGCT
 GTTTACCGACGCCATTTACAGATGGGACGCCGATGACATTGATCAACTGCCGCAATATCTG
 AAAGACGCCTTTATGGTTCTGTACAACACCGTCAAAACCCTGGAAGAAGAACTGGAGCCGG
 AAGGTAACCTTTATCGTGGTTTCTACGTTAAAGATGCGATGAAAGTTCTGGCGCGTGACTAT
 TTCGTTGAGCATAAGTGGTACAATCGTAAGATCGTCCCGTCCGTTGAAGAGTACTTGAAGA
 TTAGCTGTATCAGCGTCGCAGTCCACATGGCGACCGTGCACTGTATCGCCGGCATGTATGA
 GATCGCCACGAAAGAAGCATTGAGTGGCTGATGACCGAGCCGAAACTGGTGATTGACGC
 AAGCCTGATTGGTCGCCTGCTGGACGATATGCAGAGCACGAGCTTTGAGCAGCAGCGCGGT
 CATGTTAGCTCCGCAGTTCAATGCTACATGGCTGAGTACGGTGTGACTGCCGAAGAAGCAT
 GCGAGAAGCTGCGTGATATGGCGGCCATTGCGTGGAAGATGTGAATGAAGCATGCCTGC
 GCCCGACCGTTTTTCCCGATGCCGATTTTACTGCCTAGCATCAACCTGGCACGTGTGGCGGA
 AGTTATCTATCTGCGTGCGGACGGTTATACGCACGCGGGTGGTGAGACTAAGAAGCACATC
 ACCGCGATGCTGGTCAAGCCGATCGAAGTGTA

SEQ ID NO.: 4

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

ATGGCTTCCACCCTCCCTCTCCCAGCTTATGGAGATTGAGAAGTTGTTAGGCGATCTGCCGG
 GTTTCATCCGACGATCTGGGGCGATCACTTCCTCTCCTACAAGCCTGATCCAACGAAAATA
 GATGAATGGAATAAAAGGGTTGAAGAGCTGAAGGAAGAAGTGAAGAAGATATTAAGCAAT
 GCAAAAGGGACAGTGGAAGAGCTGAATTTGCTTGATGATCTCGTACACCTTGGGATTAGTT
 ATCATTTTGAGAAGGAGATTGATGATGCTTTACAACAAATCTTTGATACCCATCTTGATGTT
 TTTCTAAGGATGATCTATATGCCACCGCTCTCCGATTTGGCGTCTTAAGGAAACAGGGGC
 ACCGTGTTTCTCCAGATGTATTCAAAAATTCAAAGATGAGCAGGGGAATTTCAAGGCAGA
 GTTGAGCACCGATGCGAAGGGTTTGCTATGTTTATATGATGTGGCTTATCTCAGCACAAGA
 GGGGAAGATATCTTGATGAAGCCATTCCTTTCACTAAGGAGCACCTTAGGTCTTGATTA
 GCCATGTCGATTCTCATATGGCAGCAAAAATTGAGCATTCTCTAGAGCTTCCCCTTCATCAT
 CGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCTATGAAGGAGACAAGGAAAGG
 AATGAAGTTGTCCTTGAGCTTGCCAAATTAGATTTCAATCTGATTCAAATCTTGACCAAAG
 AGAGCTGAGGGACATCACACGTGGTGGAAGGAGATTGACCTTGACGCAAGCTACCTTTT

ATTAGGGATAGGTTGGTGGAGTGCTACTATTGGATCATGGGAGTCTATTTTGAACCAATAT
 ACTCAAGGGCTAGAGTTTTTTCGACCAAAATGACAATCTTGGTCTCAGTTGTGGACGACAT
 ATATGATGTATATGCTACAGAGGATGAGCTCCAACTTTTCACTGATGCAATCTATAGGTGG
 GATGCTGAGGACATTGAGCAGCTTCCACAGTACTTGAAAGATGCTTTTCTTGACTCTATAA
 CACTGTGAAGGACCTAGAAGAGGAATTGGAACCAGAAGGAAACTCTTATCGTGGATACTA
 TGTAAGATGCGATGAAGGTTTTGGCAAGGGATTACTTTGTGGAGCACAATGGTATAAC
 AGAAAAATTGTGCCATCAGTAGAGGACTACCTGCGAATTTCTTGCAATTAGTGTTGCCGTTCA
 TATGGCCACAGTTCATTGTATTGCTGGGATGTATGAAATTGCAACCAAAGAGGCATTGCGAA
 TGGTTGAAGACGGAACCTAACTTGTTATAGATGCATCACTGATTGGGCGTCTCCTCGATG
 ACATGCAGTCCACCTCGTTTGAGCAACAGAGAGGTCATGTGTCATCAGCGGTACAGTGTTA
 CATGATCCAATATGGGGTATCACACGAAGAAGCGTGTGAGAAGTTGCGAGAAATGGCTGC
 AATTGCGTGGAAGATGTAAACCAAGCATGCCTTAGGCCCACTGTTTTCCCTATGCCTATTC
 TTCTGCCCTCCATCAACCTTGCACGTGTGGCAGAAGTGATTACCTACGCGGAGATGGATAT
 ACACATGCGGGTGGTGAGACCAAAAAACATATCACGGCCATGCTTGTTGATCCAATCAAAG
 TCTGA

SEQ ID NO.: 5

- 5 SCH51-3228-9 de *D. winteri*, secuencia de aminoácidos.

MASTLPLPAYGDSEVVRRSAGFHPTIWGDHFLSYKPDPTKIDEWNKRVEELKEEVKKILSNAKC
 TVEELNLLDDLVLHGLISYHFEKEIDDALQQIFDTHLDVFPKDDLYATALRFGVLRKQGHRVSPD
 VFKKFKDEQGNFKAELSTDAKGLLCLYDVAYLSTRGEDILDEAIPFTKEHLRSCISHVDSHMAA
 KIEHSLELPLHHRIPRLNRHYISVYEGDKERNEVVLELAKLDFNLIQILHQRELDRDITWWKEID
 LAAKLPIRDLVECYWIMGVYFEPIYSRARVFSTKMTILVSVDIYDVYATEDELQLFTDAI
 YRWDAEDIEQLPQYLKDAFLVLYNTVKDLEEELEPEGNSYRGYYVKDAMKVLARDYFVEHK
 WYNRKIVPSVEDYLRISCISVAVHMATVHCCAGMDEIATKEAFEWLKTEPKLVIDASLIGRLLD
 DMQSTSFEQQRGHVSSAVQCYMIQYGVSHHEACEKLREMAAIAWKDVNQACLRPTVFPMPIIL
 LPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVDPIKV

SEQ ID NO.: 6

SCH51-3228-9_opt, secuencia de ADN optimizada con codón de SCH51-3228-9.

ATGGCAAGCACCCCTGCCGCTGCCTATGGTGATAGCGAAGTTGTTTCGTCTAGCGCAG
GTTTTTCATCCGACCATTGTTGGGGTGATCATTTTCTGAGCTATAAACCGGATCCGACCAAAATT
GATGAATGGAATAAACGTGTCTGAAGAACTGAAAGAAGAAGTGAAAAAATCCTGAGCAAT
GCCAAAGGCACCGTTGAGGAACTGAATCTGCTGGATGATCTGGTTCATCTGGGTATCAGCT
ATCACTTTGAGAAAGAAATCGATGATGCACTGCAGCAGATTTTTGATACCCATCTGGATGT
TTTCCCGAAAGATGATCTGTATGCAACCGCACTGCGTTTTGGTGTTCTGCGTAAACAGGGTC
ATCGTGTTAGTCCGGATGTGTTCAAAAAATTCAAAGATGAACAGGGCAACTTCAAAGCAGA
ACTGAGCACCGATGCAAAAGGTCTGCTGTGTCTGTATGATGTTGCATATCTGAGCACCCGT
GGTGAAGATATTCTGGATGAAGCAATTCCGTTTACCAAAGAACATCTGCGTAGCTGTATTA
GCCATGTTGATAGCCACATGGCAGCGAAAAATTGAACATAGCCTGGAAGTGCCTCTGCATCA
CCGTATTCCGCGTCTGGAAAATCGTCACTATATTAGCGTTTATGAGGGCGATAAAGAACGC
AATGAAGTTGTGCTGGAAGTGGCAAACTGGATTTTAACCTGATTCAGATTCTGCATCAGC
GTGAAGTGCCTGATATTACCACCTGGTGGAAAGAAATTGATCTGGCAGCAAACTGCCGTT
TATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTATGGGCGTGTATTTTGAACCGATTT
ATAGCCGTGCACGTGTTTTTAGCACCAAAATGACCATTCTGGTTAGCGTGGTGGATGATATC
TATGATGTTTATGCCACCGAAGATGAACTGCAGCTGTTTACCGATGCCATTTATCGTTGGGA
TGCAGAAGATATTGAACAGCTGCCGCAGTATCTGAAAGATGCATTTCTGGTCTGTACAAC
ACCGTGAAAGATCTGGAAGAAGAACTGGAACCGGAAGGTAATAGCTATCGTGGTTATTAT
GTAAAGATGCCATGAAAGTTCTGGCACGCGATTATTTTGTGAGCACAAATGGTATAACC
GCAAAATTGTTCCGAGCGTGGAAGATTATCTGCGTATTAGCTGCATTAGCGTTGCAGTTCAC
ATGGCAACCGTTTATTGTTGTGCAGGTATGGATGAAATTGCAACCAAGAAGCATTGAGT
GGCTGAAAACCGAACCGAACTGGTTATTGATGCAAGCCTGATTGGTCGTCTGCTGGACGA
TATGCAGAGCACCAAGCTTTGAACAGCAGCGTGGTCATGTTAGCAGCGCAGTTCAGTGTAT
ATGATTCAGTATGGTGTAGCCATGAAGAAGCATGCGAAAACTGCGCGAAATGGCAGCA
ATTGCATGGAAAGATGTTAATCAGGCATGTCTGCGTCCGACCGTTTTTCCGATGCCGATTCT
GCTGCCGAGCATTAACTGGCACGTGTTGCCGAAGTTATCTATCTGCGTGGTGATGGTTATA
CCCATGCCGGTGGTGAAACCAAAAAACATATTACCGCAATGCTGGTCGATCCGATTAAAGT
TTAA

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

ATGGCTTCCACCCTCCCTCTCCCAGCTTATGGAGATTCAGAAGTTGTTAGGCGATCTGCCGG
 GTTTCATCCGACGATCTGGGGCGATCACTTCCTCTCCTACAAGCCTGATCCAACGAAAATA
 GATGAATGGAATAAAAGGGTTGAAGAGCTGAAGGAAGAAGTGAAGAAGATATTAAGCAAT
 GCAAAAGGGACAGTGGAAGAGCTGAATTTGCTTGATGATCTCGTACACCTTGGGATTAGTT
 ATCATTTTGAAGAAGGAGATTGATGATGCTTTACAACAAATCTTTGATACCCATCTTGATGTT
 TTTCTAAGGATGATCTATATGCCACCGCTCTCCGATTTGGCGTCTTAAGGAAACAGGGGC
 ACCGTGTTTCTCCAGATGTATTCAAAAAATTCAAAGATGAGCAGGGGAATTTCAAGGCAGA
 GTTGAGCACCGATGCGAAGGGTTTGCTATGTTTATATGATGTGGCTTATCTCAGCACAAGA
 GGGGAAGATATCTTGGATGAAGCCATTCTTTTCACTAAGGAGCACCTTAGGTCTTGTTATTA
 GCCATGTTCGATTCTCATATGGCAGCAAAAATTGAGCATTCTCTAGAGCTTCCCCTTCATCAT
 CGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTCTATGAAGGAGACAAGGAAAGG
 AATGAAGTTGTCCTTGAGCTTGCCAAATTAGATTTCAATCTGATTCAAATCTTGCACCAAAG
 AGAGCTGAGGGACATCACAATGTGGTGGAAGGAGATTGACCTTGCAGCAAAGCTACCTTTT
 ATTAGAGATAGGTTGGTGGAGTGCTACTACTGGATCATGGGGGTCTATTTTGAACCAATAT
 ACTCCAGGGCTAGGGTTTTTCCACTAAAATGACAATCTTGGTCTCAGTTGTGGACGACATA
 TATGATGTCTATGCTACGGAGGATGAGCTTCAACTATTCACTGATGCAATCTATAGGTGGG
 ATGCTGATGACATTGATCAGCTGCCTCAGTACTTGAAAGATGCTTTTATGGTACTCTATAAC
 ACTGTGAAGACTCTAGAAGAAGAACTTGAACCAGAAGGAAACTCTTATCGTGGATACTACG
 TAAAAGATGCAATGAAGGTTTTTGGCAAGAGATTACTTTGTGGAACACAAATGGTATAACAG
 ACAAATTGTGCCATCCGTAGAGGAATACTTGAAAATTTCTTGCATTAGTGTGGCTGTTTATA
 TGGCTACAGTTCATTGTATTGCTGGGATGTATGAAATTGCTACCAAAGAGGCATTGGAATG
 GTTGAAGACTGAACCCAACTTGTTATCGATGCATCTCTGATCGGTCGTCTTCTTGATGACA
 TGCAGTCTACCTCGTTTGAACAACAAAGAGGGGCACGTGTCATCAGCAGTACAGTGTTACAT
 GGCCCAATATGGAGTAACAGCAGAAGAAGCATGTGAAAAGCTACGAGAAATGGCTGCAAT
 TGCTTGGAAAGATGTGAATGAAGCATGCCTTAGGCCCACGGTATTCCCTATGCCTATCCTCT
 TGCCTTCTATCAACTTGGCACGTGTGGCAGAAGTGATCTACCTACGTGGAGATGGATACAC
 GCACGCTGGGGGTGAGACCAAAAAACACATCACGGCCATGCTTGTTAAGCCAATTGAAGTC
 TGA

5 SEQ ID NO.: 8

SCH51-3228-11 de *D. winteri*, secuencia de aminoácidos.

MASTLPLPAYGDSEVVRRSAGFHPTIWGDHFLSYKPDPTKIDEWNKRVEELKEEVKKILSNAKG
TVEELNLLDDLVLHGLISYHFEKEIDDALQQIFDTHLDVFPKDDLYATALRFGVLRKQGHRVSPD
VFKKFKDEQGNFKAELSTDAKGLLCLYDVAYLSTRGEDILDEAIPFTKEHLRSCISHVDSHMAA
KIEHSLELPLHHRIPRLNRHYISVYEGDKERNEVVLELAKLDFNLIQILHQRELRDITMWWKEI
DLAAKLPIFIRDRLVECYWIMGVYFEPISRARVFSTKMTILVSVDDIYDVYATEDELQLFTD
AIYRWDADDIDQLPQYLKDAFMVLYNTVKTLLEELEPEGNSYRGYYVKDAMKVLARDYFVEH
KWYNRQIVPSVEEYLKISCISVAVHMATVHCIAGMYEIATKEAFEWLKTEPKLVIDASLIGRLLD
DMQSTSFEQQRGHVSSAVQCYMAQYGVTAEEACEKLREMAAIAWKDVNEACLRPTVFPMPIL
LPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVKPIEV

SEQ ID NO.: 9

- 5 SCH51-3228-11_opt de *D. winteri*, secuencia de ADN optimizada con
codón de SCH51-3228-11.

ATGGCATCTACTCTTCCACTGCCGGCTTATGGTGATTCTGAGGTTGTTCGTCGTTCCGCGGG
TTTTACCCTACCATCTGGGGCGATCACTTTCTGTCCTATAAGCCAGACCCGACCAAGATTG
ACGAGTGGAATAAGCGTGTCTGAGGAAGTAAAGAAAGTAAAAAGATCCTGTCCAACG
CAAAAGGTACTGTCTGAGGAGCTGAATCTGCTGGATGACCTGGTGCATCTGGGCATCAGCTA
TCACTTCGAAAAGGAAATTGACGACGCTTTGCAGCAAATTTTGTATACGCACCTGGACGTC
TTCCGAAAGATGACCTGTATGCGACCGCGCTGCGCTTTGGTGTGCTGCGTAAACAGGGTC
ATCGCGTGTCTCCTGATGTGTTCAAGAAATTTAAAGATGAACAGGGCAATTTCAAGGCCGA
GTTGAGCACGGACGCCAAAGGTTTGCTCTGCCTGTACGACGTTGCATATCTGAGCACCCGT
GGTGAAGATATCCTGGACGAAGCGATTCCGTTACCAAGGAACATCTGCGCTCGTGCAATTT
CCCATGTAGATAGCCACATGGCGGCCAAGATCGAGCACAGCCTGGAGCTGCCTTTGCACCA
TCGTATTCCGCGCCTGGAGAATCGCCATTACATTAGCGTCTATGAGGGTGACAAAGAGCGC
AACGAAGTCGTGTTAGAGCTGGCGAAGCTGGACTTCAACCTGATTCAAATTCTGCATCAAC
GCGAGCTGCGCGACATTACCATGTGGTGGAAAGAGATTGATCTGGCAGCGAAGCTGCCGTT

CATCCGCGATCGTCTGGTTGAGTGCTACTACTGGATCATGGGCGTCTACTTCGAGCCGATCT
 ACAGCCGCGCTCGTGTGTTTTTCGACGAAGATGACCATCCTGGTTAGCGTTGTTGATGACATT
 TATGACGTTTACGCGACCGAAGATGAACTGCAGCTGTTTACGGACGCAATCTACCGTTGGG
 ACGCGGATGATATCGACCAGCTGCCGCAATACTTGAAAGATGCGTTCATGGTTTTGTACAA
 CACCGTCAAAACGCTGGAAGAAGAACTGGAGCCGGAAGGCAACAGCTACCGTGGTTACTA
 TGTTAAAGATGCGATGAAAGTTCTGGCGCGCGACTACTTCGTCGAGCACAAGTGGTATAAC
 CGTCAGATTGTGCCGAGCGTCGAGGAATACCTGAAGATTAGCTGTATCAGCGTTGCCGTTT
 ACATGGCAACGGTGCACTGCATCGCCGGTATGTACGAGATTGCGACGAAAGAAGCCTTCGA
 ATGGTTGAAAACCGAGCCGAAGCTGGTTATCGACGCCAGCCTGATCGGTCGTTTGCTGGAC
 GACATGCAAAGCACGAGCTTCGAGCAGCAGCGCGGCCATGTGAGCAGCGCTGTTCAGTGT
 ATATGGCGCAATATGGCGTGACCGCAGAAGAAGCGTGCGAGAAGCTGCGTGAGATGGCAG
 CAATTGCGTGGAAGATGTGAATGAAGCCTGTCTGCGTCCGACTGTGTTTCCGATGCCGAT
 CCTGCTGCCGAGCATTAACTGGCGCGTGTGGCAGAGGTCATCTATCTGCGTGGTGACGGT
 TACACCCACGCGGGTGGCGAAACCAAGAAACATATCACCGCAATGCTGGTTAAGCCGATT
 GAAGTGTA

SEQ ID NO.: 10

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

ATGGATCTTAGTACTTCACCTGTTCTTTCTTCTCCCCCTTCCGGTGGAAGACGGAAAAA
 TCCGGCCGTTTCGCCGTTGAGCTGGATTTCACCCAGTATTTGGGGTGATCATTTCTCTCCT
 AACTGAAGATCACAAGAAGCTGGATGCATGGAGCGAAAGGACTCAAGTGTTGAAGGAAG
 AGGTGAGGAGAATTTTAATCAATGCCAAGGGGTCAGTAGAAGAGTTGGATTTGTTGGATGC
 AATCCAACGCCTTGGGGTGAAATATCACTTTGAGAAAGAGATTGAAGAGGCATTACACCAT
 ATTTATGTTGCAGAACTCATGTTTCTACTGATGACTTATATTCCGTTTCTCTCCGGTTTCGA
 CTTCTTAGACAACAAGGGTACAATGTATCTGCTGATGTATTTAAAAAGTTCAAAGATGAGA
 GGGGCAACTTCAAGGCAAGCTTAAGTACTGATGCCAGGGGGTTGCTAAGCTTGTATGAAGC
 TGCATTTCTCAGCATACGAGGAGATGATATCTTAGATGAAGCCATAACTTTCACAAGAGAG
 CAGCTTAAGTCTTCTATGACCCATGTTGATGCCCTCTTGCCAAACAAATAGCCCATGCCTT
 AGAGGTACCAGCGCACAAGCGCATACAAAGACTAGAGAACATTCGCTACCTCACAATCTA
 CCAAGAAGAGAAAGGAAGGAATGATGTGTTGCTTGAGCTTGCCAAGTTGGATTTCAATATC

TTACAACAATTGCATAAGAAAGAACTGAGAGACCTTACAAAGTGGTGGGAAGGACACAGAC
 GTTGACAGGAAAGCTACCTTTCATCAGAGATAGGTTGGTGGGAATGCTATTATTGGATCTTGG
 GTGTGTATTATGAGCCAGAATACTCCAGAGCTAGAATTTTTCTACCAAAATGACAATCAT
 GGTCTCAGTTGTTGATGACATATATGACGTATATGCTACTGAAGATGAGCTCCAATATTCA
 CTGATGCAATCTATAGGTGGGATCTGGAGGGCCTAGATCAACTCCCACAGTTCTTGAAAGA
 CTGTTTTCTTGTACTCTATGACACCGTCAAGGAATTAGAAGACGAACTAGAACCGGAAGGA
 AAATCCTATCGTGGATACTATGTAAAGGATGCGATGAAGGTTTTGGCTAGAGATTACTTCG
 TTGAGCACAAATGGTATAACAGAAACATAGTGCCAAGTGTAGAAGAATATCTCCGTGTTTC
 TTGCATCAGTGTTGCAGTCCATATGGCTAACGTCCATTGCTGTGCTGGGATGGGAGATGTA
 ATGAGCAAAGAGGGCATTCTGAATGGTTGAAGAGTGAACCAAAGGTTGTAATGGATGCATCA
 CTAATTGGCCGACTGCTCGATGACATGCAGTCCACCGAGTTTGAGCAAAAGAGAGGGCCATG
 TTGCATCGGCTGTCCAATGTTACATGAATGAGTATGGAGTGAAGTACAAAGAAGCGTGTGA
 AAAGCTGCATGAAATGGCTGCCCTTGCATGGAAAGACGTAAACCAGGCTTGCCCTAAACCA
 ACTGTTTTCCCTCTCCCTGTATTTATGCCTGCAATCAACCTTGCGCGAGTGGCTGAAGTCAT
 CTACCTTCGTGGAGATGGGTATACTCATTGAGGAGGAGAGACTAAAGAAAATATCACGTTG
 ATGCTTGTCAATCCAATCTCTGTGTGA

SEQ ID NO.: 11

- 5 SCH51-998-28 de *D. winteri*, secuencia de aminoácidos.

MDLSTSPVLSSSPLPVEDGKNPAVRRSAGFHPSIWGDHFLSYTEDHKKLDAWSERTQVLKEEV
 RRILINAKGSLEELDLLDAIQRLGVKYHFEKEIEEALHHIYVAETHVSTDDLYSVSLRFLLRQQ
 GYNVSADVFKFKDERGNFKASLSTDARGLLSLYEAFLSIRGDDILDEAITFTREQLKSSMTH
 VDAPLAKQIAHALEVPAHKRIQRLNIRYLTIYQEEKGRNDVLLELAKLDFNILQQLHKKELRD
 LTKWWKDTDVAGKLPIRDLVECYWILGVYYPEYSRARIKSTKMTIMVSVDIYDVYAT
 EDELQLFTDAIYRWDLEGLDQLPQFLKDCFLVLYDTVKELEDELEPEGKSYRGYYVKDAMKVL
 ARDYFVEHKWYNRNIVPSVEEYLRVSCISVAVHMANVHCCAGMGDVMSKEAFEWLKSEPKV
 VMDASLIGRLLDDMQSTEFQKRGHVASAVQCYMNEYGVITYKEACEKLHEMAALAWKDVN
 QACLKPTVFPLPVFMPAINLARVAEVIYLRGDGYTHSGGETKENITLMLVNPISV

SEQ ID NO.: 12

SCH51-998-28_opt de *D. winteri*, secuencia de ADN optimizada con codón de SCH51-998-28.

ATGGATCTGAGCACCAGTCCGGTTCTGAGCAGCTCACCGCTGCCGGTTGAAGATGGTAAAA
 ATCCGGCAGTTCGTCTGATAGCGCAGGTTTTATCCGAGCATTGTTGGGGTGATCATTTTCTGAGC
 TATACCGAGGATCACAAAAAAGTGGATGCATGGTCAGAACGTACCCAGGTTCTGAAAGAA
 GAAGTGCCTCGTATTCTGATTAATGCAAAAGGTAGCCTGGAAGAACTGGATCTGCTGGATG
 CAATTCAGCGTCTGGGTGTTAAATATCACTTTGAGAAAGAAATCGAAGAAGCCCTGCATCA
 TATTTATGTTGCAGAAACCCATGTGTCAACCGATGATCTGTATAGCGTTAGCCTGCGTTTTCT
 GTCTGCTGCGTCAGCAGGGTTATAATGTTAGCGCAGATGTGTTCAAAAAATTCAAAGATGA
 ACGCGGTAACCTCAAAGCAAGCCTGAGCACCAGATGCACGTGGTCTGCTGAGCCTGTATGAA
 GCAGCATTTCTGAGCATTCGTGGTGATGATATTCTGGATGAAGCAATTACCTTTACCCGTGA
 ACAGCTGAAAAGCAGCATGACCCATGTTGATGCACCGCTGGCAAACAAATTGCACATGC
 ACTGGAAGTTCCGGCACATAAACGTATTCAGCGCCTGGAAAATATTCGCTATCTGACCATT
 TACCAAGAAGAGAAAGGTTCGTAACGATGTTCTGCTGGAAGTGGCCAACTGGATTTTAAACA
 TTCTGCAGCAGCTGCATAAAAAAGAACTGCGTGATCTGACCAAATGGTGGAAAGATACCG
 ATGTTGCAGGTAAACTGCCGTTTATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTCTG
 GCGTTTTATTATGAGCCGGAATATAGCCGTGCACGTATTTTATGACCAAAAATGACCATTAT
 GGTTAGCGTGGTGGATGACATCTATGATGTTTATGCAACCGAAGATGAACTGCAGCTGTTT
 ACCGATGCAATTTATCGTTGGGATCTGGAAGGTCTGGATCAGCTGCCGCAGTTCCTGAAAG
 ATTGTTTTCTGGTTCTGTATGATACCGTGAAAGAACTGGAAGATGAGCTGGAACCGGAAGG
 TAAAAGCTATCGTGGTTATTATGTTAAAGATGCCATGAAAGTTCTGGCACGCGATTATTTTG
 TTGAGCACAAATGGTATAACCGCAATATTGTTCCGAGCGTGGAAGAATATCTGCGTGTTAG
 CTGTATTAGCGTTGCAGTTCACATGGCAAATGTTTATTGTTGTGCAGGTATGGGTGATGTGA
 TGAGCAAAGAAGCATTTGAATGGCTGAAAAGTGAACCGAAAGTTGTTATGGATGCCAGCCT
 GATTGGTGCCTGCTGGACGATATGCAGAGCACCGAATTTGAACAGAAACGTGGTCATGTT
 GCAAGCGCAGTTCAGTGTTATATGAATGAATATGGCGTGACCTATAAAGAGGCATGCGAAA
 AACTGCATGAAATGGCAGCACTGGCATGGAAAGATGTTAATCAGGCATGTCTGAAACCGA
 CCGTTTTTCCGCTGCCTGTTTTTATGCCTGCAATTAATCTGGCACGTGTTGCCGAAGTTATTT
 ACCTGCGTGGGGATGGTTATACCCATAGCGGTGGTGAAACCAAAGAAAACATTACCCTGAT
 GCTGGTTAATCCGATTAGCGTTTAA

SCH52-13163-6 de *D. lanceolata*, secuencia de ADN.

ATGGATGTTCTAATTCCTCCCCTGTGGCTTCCACTCTCCCTCTGCCCCGAAGATGGAACTT
GGACGTCGTTTCGCAGATCCGCCGGGTTTCATCCGACGGTCTGGGGCGATCACTTCCTCGCTT
ACTCGCCCGATCCAACCAAAATAGATGCTTGGACTAAAAGAGTTGAAGAGCTGAAGCAAG
AAGTGAAGAGGATTCTAAGCAATGTGAAAGGGTCACTGGAAGAGCTGAACTTGCTTGATG
CTATCCAACACCTTGGGATTGGTTATCATTTTGAGAAAGAGATTGATGATGCTTTACAATA
ATCTTTGATTCCCATATTGATGCTTTTCTACTGATGATCTATATGTGGCTGCCCTCCGATTT
AGCCTACTAAGGCGACAAGGGCACTGTGTTTCTTCAGATGTATTCAAAAAATTCAAAGATG
AGCAGGGGAATTTCAAGGCAGAGCTGAGCACCGATGCGAAAGGTTTGCTGAGTCTCTATGA
CGCGGCGTATCTCAGTGTAAGAGGGGAAGATATATTGGATGAGGCCATTCCTTTCACTAGG
GAGCACCTTAGGACTTGTATTAGCCATGTAGATTCTCATTTGGCAGCAAAAATTGAGCATTC
TCTAGAGCTTCCCCTGCATCATCGCATACCAAGGCTAGAGAACAGGCACTACATCTCAGTG
TACGAAGGAGAGAAGGAAAGGAATGAAGTTGTACTAGAGCTTGCCAAATTAGATTTCAAT
CTGATTCAAATCTTGACCAAAGAGAGCTGAGGGACATCACAACGTGGTGGAATGAGATTG
ACCTCGCAGCAAAGCTACCATTTATTAGGGATAGGTTGGTGGAGTGCTACTATTGGATCAT
GGGTGTCTATTTTGAACCAATATTCTCAAGGGCTAGAGTTTTTTTCGACCAAAATGACAATTT
TGGTCTCAGTTGTCGACGACATATATGATGTCTACGCTACAGAGGATGAGCTCCAACTTTTCT
ACTGACGCAATCTATAGGTGGGATGCCGAGGACATTGAGCAGCTTCCACAGTACTTGAAAG
ATTCTTTTCTTGTAATCTATAACACCGTGAAGGACTTAGAAGAGGAGCTGAAACCAGAAGG
AAACTCATATCGTGGAGACTATGTAAAAGATGCGATGAAGGTTTTGGCAAGAGATTACTTT
GTGGAGCACAAATGGTATAACAGAAAAATTGTACCGTCAGTAGAGGACTACCTACGAATTT
CTTGCAATTAGTGTTGCCGTTTCATATGGCTACAGTTCATTGTTGTGCTGGGATGGATGAAATT
GCAACCAAAGAGGCATTTCGAATGGTTGAAGACCGAACCTAACTTGTTATAGATGCATCAC
TGATTGGGCGTCTCCTCGATGACATGCAGTCCACCTCGTTTGAGCAACAGAGAGGTCATGT
GTCATCGGCGGTACAGTGTTACATGATCCAATATGGCGTATCACACGAAGAAGCGTGTGAG
AAGTTGACAGAAATGGCTGCAATTGCATGGAAAGATGTAAACCAAGCATGCCTTAGGCCC
ACTGTTTTCCCAATGCCTATTCTTCTGCCTTCAATCAACCTTGACGTGTGGCAGAAGTCAT
CTACCTGCGCGGAGATGGATATACACATGCTGGTGGTGAGACCAAAAAACATATCACGGCC
ATGCTTGTTGAACCAATCCAAGTCTGA

SEQ ID NO.: 14

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

MDVLIPSPVASTLPLPEDGNLDVRRSAGFHPTVWGDHFLAYSPDPTKIDAWTKRVEELKQEV
KRILSNVKGSLLEENLLDAIQHLGIGYHFEKEIDDALQLIFDSHIDAFPTDDLYVAALRFSLLRRQ
GHCVSSDVFKKFKDEQGNFKAELSTDAKGLLSLYDAAYLSVRGEDILDEAIPFTREHLRTCISH
VDSHLAAKIEHSLELPLHHRIPRLNRHYISVYEGEKERNEVVLELAKLDFNLIQILHQRELRDIT
TWWNEIDLAAKLPFIRDRLVECYWIMGVYFEPIFSRARVFSTKMTILVSVVDDIYDVYATEDE
LQLFTDAIYRWDAEDIEQLPQYLKDSFLVLYNTVKDLEEELKPEGNSYRGDYVKDAMKVLAR
DYFVEHKWYNRKIVPSVEDYLRISCISVAVHMATVHCCAGMDEIATKEAFEWLKTEPKLVIDA
SLIGRLLDDMQSTSFEQQRGHVSSAVQCYMIQYGVSHEEACEKLTEMAAIAWKDVNQACLRPT
VFPMPIILLPSINLARVAEVIYLRGDGYTHAGGETKKHITAMLVEPIQV

5 SEQ ID NO.: 15

SCH52-13163-6_opt, secuencia de ADN optimizada con codón de SCH51-13163-6.

ATGGATGTTCTGATTCCGAGTCCGGTTGCAAGCACCTGCCGCTGCCGGAAGATGGTAATC
TGGATGTTGTTTCGTCTAGCGCAGGTTTTTCATCCGACCGTTTGGGGTGATCATTTTTCTGGCA
TATAGTCCGGATCCGACCAAAATTGATGCATGGACCAAACGTGTTGAGGAACTGAAACAA
GAAGTGAAACGTATTCTGAGCAATGTGAAAGGTAGCCTGGAAGAACTGAATCTGCTGGAT
GCAATTCAGCATCTGGGTATTGGTTATCACTTCGAGAAAGAAATTGATGATGCACTGCAGC
TGATCTTTGATAGCCATATTGATGCCTTTCCGACCGATGATCTGTATGTTGCAGCACTGCGT
TTTAGCCTGCTGCGTCGTCAGGGTCATTGTGTTAGCAGTGATGTTTTCAAAAAATTCAAAGA
CGAGCAGGGCAACTTTAAAGCAGAACTGAGCACCGATGCAAAAGGTCTGCTGAGCCTGTA
TGATGCCGCATATCTGAGCGTTCGTGGTGAAGATATTCTGGATGAAGCAATTCCGTTTACCC
GTGAACATCTGCGTACCTGTATTAGCCATGTGGATAGCCATCTGGCAGCAAAAATTGAACA
TAGTCTGGAAGTGCCTCTGCATCATCGTATTCCGCGTCTGGAAAATCGTCACTATATTAGCG
TTTATGAAGGCGAAAAAGAACGCAATGAAGTTGTGCTGGAAGTGGCAAACTGGATTTTAA
CCTGATTCAGATTCTGCATCAGCGTGAAGTGCCTGATATTACCACCTGGTGGAATGAAATT
GACCTGGCAGCCAACTGCCGTTTATTCGTGATCGTCTGGTTGAATGCTATTATTGGATTAT
GGGCGTGTATTTTGAACCGATTTTATAGCCGTGCACGTGTGTTTAGCACCAAAATGACCATTC
TGGTTAGCGTGGTGGATGATATCTATGATGTTTATGCAACCGAAGATGAGCTGCAACTGTTT
ACCGATGCCATTTATCGTTGGGATGCAGAAGATATTGAACAGCTGCCTCAGTATCTGAAAG
ATAGCTTTCTGGTTCTGTACAACACCGTGAAAGATCTGGAAGAAGAACTGAAACCGGAAGG
TAATAGCTATCGTGGTGATTATGTTAAAGACGCCATGAAAGTTCTGGCACGCGATTATTTTG
TTGAGCACAAATGGTATAACCGCAAAATTGTTCCGAGCGTGGAAAGATTATCTGCGTATTAG
CTGCATTAGCGTTGCAGTTCACATGGCAACCGTTCATTGTTGTGCAGGTATGGATGAAATTG
CAACCAAAGAAGCATTTGAGTGGCTGAAAACCGAACCGAACTGGTTATTGATGCAAGCCT
GATTGGTCTGCTGCTGGACGATATGCAGTCAACCAGCTTTGAACAGCAGCGTGGTCATGTT
AGCAGCGCAGTTCAGTGTTATATGATTCAGTATGGTGTTAGCCATGAAGAAGCATGCGAAA
AACTGACCGAAATGGCAGCAATTGCATGGAAAGATGTAAATCAGGCATGTCTGCGTCCGAC
CGTGTTCCTATGCCGATTCTGCTGCCGAGCATTAACTCTGGCACGTGTTGCCGAAGTTATCT
ATCTGCGTGGTGATGGTTATACCCATGCCGGTGGTGAAACCAAAAAACATATTACCGCAAT
GCTGGTAGAACCGATTACAGGTTTAA