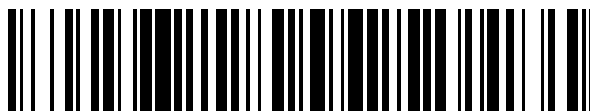


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 747 844**

51 Int. Cl.:

C12N 9/02 (2006.01)

C12N 15/52 (2006.01)

C12P 7/18 (2006.01)

C12P 7/62 (2006.01)

C12P 15/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.01.2014 PCT/EP2014/050246**

87 Fecha y número de publicación internacional: **16.07.2015 WO15104051**

96 Fecha de presentación y número de la solicitud europea: **08.01.2014 E 14700343 (8)**

97 Fecha y número de publicación de la concesión europea: **03.07.2019 EP 3092307**

54 Título: **Agrupación de genes para biosíntesis de meleólidos y su aplicación**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.03.2020

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80686 München , DE**

72 Inventor/es:

**JENNEWEIN, STEFAN y
ENGELS, BENEDIKT**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 747 844 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Agrupación de genes para biosíntesis de meleólidos y su aplicación

5 La presente invención se refiere a una agrupación de genes recién identificada, en particular a polinucleótidos recién identificados dentro de dicha agrupación de genes, así como a los polipéptidos codificados por los polinucleótidos, y a su producción y usos, así como a sus variantes y a sus usos. Además, la presente invención se refiere a un método para producir protoiludeno hidroxilado y/o ésteres arílicos de tipo protoiludeno sesquiterpenoide usando los polinucleótidos/polipéptidos recién identificados.

Antecedentes de la invención

10 Debido al hecho de que la prevalencia de patógenos microbianos con resistencia a la terapia antibiótica convencional aumenta constantemente, la investigación ha comenzado a centrarse en la búsqueda y el desarrollo de nuevos fármacos antimicrobianos con nuevos modos de acción. Los productos naturales son tradicionalmente una fuente rica de compuestos bioactivos, pero aunque los terpenoides son, con diferencia, la clase más grande de productos naturales, sólo hay tres fármacos antimicrobianos aprobados en la familia de los terpenoides y los tres son derivados semisintéticos del producto natural fúngico pleuromutilina (tiamulina, valnemulina y retapamulina).

15 Los basidiomicetos son una fuente particularmente rica de sesquiterpenoides complejos, estructuralmente diversos y bioactivos, y determinados ésteres arílicos sesquiterpenoides de tipo protoiludeno del género *Armillaria* son pistas prometedoras para el desarrollo de nuevos fármacos antimicrobianos. El género *Armillaria*, también conocido comúnmente como hongo de miel, actualmente se clasifica en la familia *Physalacriaceae* y comprende más de 600 especies en todo el mundo. Las especies de *Armillaria* se consideran no sólo hongos comestibles, sino también patógenos de la raíz notorios que atacan los árboles de corteza dura y las coníferas, así como los árboles frutales y la vid.

20 Se conocen miembros del género *Armillaria* para la producción de dos tipos de ésteres arílicos de tipo protoiludeno sesquiterpenoide, que se dividen según la posición del doble enlace en armililorselinatos y meleólidos. Hoy en día se conocen más de 50 estructuras de armililorselinatos y meleólidos de *Armillaria* spp., lo que los convierte en uno de los grupos más diversos de productos naturales conocidos de hongos.

25 Si bien la biosíntesis de algunos terpenos, en particular de origen vegetal, como el mentol, la artemisinina o el taxol, se ha estudiado ampliamente en el pasado, se sabe poco sobre la síntesis de la gran mayoría, en particular de los terpenos producidos por hongos.

30 Los ésteres arílicos de tipo protoiludeno se construyen a partir de un alcohol protoiludanol sesquiterpenoide y un resto aromático derivado de ácido orselínico. Se cree que la biosíntesis de todos los sesquiterpenoide de tipo meleólido y armililorselinato se deriva de la ciclación del precursor de sesquiterpeno universal difosfato de farnesilo a protoiludeno. El sistema de anillo de protoiludeno se somete luego a varias reacciones de hidroxilación, y los alcoholes de protoiludanol resultantes se someten luego a esterificación en la posición C5-hidroxilo con derivados del ácido orselínico policétido.

35 Mientras que el ácido orselínico y sus derivados son biosintetizados por numerosos microorganismos, y los protoiludenos están muy difundidos en los homobasidiomicetos, el acoplamiento de ambos parece ser único en el género *Armillaria*

El documento WO 2012/016912 A1 da a conocer microorganismos huésped y métodos para la producción de meleólidos, en el que un gen de protoiludeno sintasa de *Armillaria* se expresa en el microorganismo huésped.

40 Engels *et al.* ("Cloning and characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides", *Journal of Biological Chemistry* 286(9):6871-6878 (2011) también da a conocer la expresión heteróloga de una protoiludeno sintasa de *A. gallica* en *E. coli*.

45 Dado que, como ya se mencionó anteriormente, los meleólidos y armililoreselinatos, en medicina, son de gran interés debido a su potencial actividad antimicrobiana y citotóxica, sería deseable producirlos específicamente o enriquecerlos de forma controlada.

Por tanto, un objeto de la presente invención es proporcionar herramientas nuevas y mejoradas mediante las cuales pueda lograrse una producción dirigida, o bien homóloga o bien heteróloga.

Sumario de la invención

50 Según la invención, este y otros objetos se logran proporcionando un microorganismo huésped según la reivindicación 1. Según la invención, el microorganismo huésped se transforma para comprender y expresar un polinucleótido que codifica un polipéptido con actividad citocromo P450 monooxigenasa y que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en: (i) a) SEQ ID NO 5 a 8 de la lista de

secuencias adjunta; b) una secuencia de ácido nucleico complementaria a SEQ ID NO 5 a 8; c) secuencias de ácido nucleico que se hibridan en condiciones rigurosas con las secuencias de ácido nucleico definidas en a) y b) o sus hebras complementarias, (ii) un vector que contiene una secuencia de ácido nucleico de (i), que codifica un polipéptido con actividad citocromo P450 monooxigenasa, estando la secuencia de ácido nucleico unida operativamente a secuencias de control reconocidas por un microorganismo huésped transformado con el vector, preferiblemente el vector es un vector de expresión, (iii) SEQ ID NO 9 de la lista de secuencias adjunta, en el que el microorganismo huésped se transforma adicionalmente para expresar una 2-hidroxi-3-metilglutaril-CoA (HMG-CoA) reductasa.

Se da a conocer en el presente documento una agrupación de genes que comprende o consiste en los siguientes genes: citocromo P450 monooxigenasa 1 de *Armillaria gallica*, citocromo P450 monooxigenasa 2 de *Armillaria gallica*, P450 monooxigenasa 3 de *Armillaria gallica*, protoiludeno sintasa de *Armillaria gallica*, citocromo P450 monooxigenasa 4 de *Armillaria gallica*, en la que los genes están dispuestos en el orden mencionado.

Los objetos se logran adicionalmente mediante un método respectivo para producir un protoiludeno hidroxilado y/o un éster arílico de tipo protoiludeno sesquiterpenoide tal como se reivindica en las reivindicaciones adjuntas.

Además, la invención proporciona el uso de la agrupación de genes mencionada anteriormente para producir un protoiludeno hidroxilado o un éster arílico de tipo protoiludeno sesquiterpenoide, y mediante un método respectivo que usa la agrupación de genes para producir un protoiludeno hidroxilado y/o un éster arílico de tipo protoiludeno sesquiterpenoide.

Los objetos se logran completamente de esa manera.

Los polinucleótidos mencionados anteriormente, codifican cada uno para un polipéptido de *Armillaria gallica* con actividad monooxigenasa que, según el conocimiento de los inventores, se ha identificado, purificado y caracterizado enzimáticamente por primera vez. Estos polinucleótidos recién identificados catalizan reacciones de hidroxilación que convierten 6-protoiludeno en 6-protoiludeno mono o multi hidroxilado, cuya reacción representa una etapa crucial en la síntesis de protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide. Por tanto, al haber identificado el gen que codifica las monooxigenasas, se ha encontrado y generado una herramienta valiosa y eficaz para influir en la producción de protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide: los genes pueden, o bien solos o bien en una agrupación que también comprende protoiludeno sintasa, expresarse o sobreexpresarse de manera heteróloga con el fin de generar múltiples copias de dichos genes, por medio de los cuales la tasa de ciclación y la generación de protoiludenos hidroxilados en la síntesis de ésteres arílicos de tipo protoiludeno sesquiterpenoide se eleva, y, por tanto, aumenta la producción de ésteres arílicos de tipo protoiludeno sesquiterpenoide. De esa manera, pueden obtenerse protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide altamente enriquecidos, que pueden o bien usarse como potentes sustancias antimicrobianas y citotóxicas y pueden usarse como herramientas terapéuticas en todos los diferentes campos de tratamiento y medicina, o bien modificarse adicionalmente para dar otras sustancias antimicrobianas derivadas de dichos protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide.

Según la presente invención, el término "polinucleótido(s)" se refiere generalmente a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN sin modificar o ARN o ADN modificado. El/los "polinucleótido(s)" incluye(n), sin limitación, ADN mono y bicatenario, ADN que es una mezcla de regiones mono y bicatenarias o regiones mono, bi y tricatenarias, ARN mono y bicatenario, y ARN que es una mezcla de regiones mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser regiones monocatenarias o, más normalmente, regiones bicatenarias o tricatenarias, o una mezcla de regiones mono y bicatenarias. Además, "polinucleótido" tal como se usa en el presente documento se refiere a regiones tricatenarias que comprenden ARN o ADN o tanto ARN como ADN. Las hebras en tales regiones pueden ser de la misma molécula o de moléculas diferentes. Las regiones pueden incluir todas de una o más de las moléculas, pero más normalmente implican sólo una región de algunas de las moléculas. Una de las moléculas de una región de triple hélice a menudo es un oligonucleótido. Tal como se usa en el presente documento, el término "polinucleótido(s)" también incluye ADN o ARN tal como se describió anteriormente que contienen una o más bases modificadas. Por tanto, ADN o ARN con estructuras principales modificadas por estabilidad o por otros motivos son "polinucleótido(s)" ya que ese término está previsto en el presente documento. Además, ADN o ARN que comprende bases inusuales, tales como inosina, o bases modificadas, tales como bases tritiladas, por nombrar sólo dos ejemplos, son polinucleótidos ya que el término se usa en el presente documento. Se apreciará que se ha realizado una gran variedad de modificaciones al ADN y al ARN que satisfacen muchos fines útiles conocidos para los expertos en la técnica. El término "polinucleótido(s)" tal como se emplea en el presente documento abarca tales formas química, enzimática o metabólicamente modificadas de polinucleótidos, así como las formas químicas de ADN y ARN características de virus y células, que incluyen, por ejemplo, células simples y complejas. Además, "polinucleótido(s)" también abarca polinucleótidos cortos a menudo denominados oligonucleótido(s).

"Polipéptido(s)" se refiere a cualquier péptido o proteína que comprende dos o más aminoácidos unidos entre sí por enlaces peptídicos o enlaces peptídicos modificados. "Polipéptido(s)" se refiere tanto a cadenas cortas, denominadas comúnmente péptidos, oligopéptidos y oligómeros como a cadenas más largas denominadas generalmente proteínas. Los polipéptidos pueden contener aminoácidos distintos de los 20 aminoácidos codificados

por el gen. El/los “polipéptido(s)” incluye(n) aquellos modificados o bien por procesos naturales, tales como procesamiento y otras modificaciones postraduccionales, pero también por técnicas de modificación química. Tales modificaciones están bien descritas en textos básicos y en monografías más detalladas, así como en una bibliografía de investigación voluminosa, y los conocen bien los expertos en la técnica. Se apreciará que el mismo tipo de modificación puede estar presente en el mismo grado o grado variable en varios sitios en un polipéptido dado. Además, un polipéptido dado puede contener muchos tipos de modificaciones. Las modificaciones pueden producirse en cualquier sitio en un polipéptido, incluyendo la estructura principal peptídica, las cadenas laterales de aminoácido, y los extremos terminales amino o carboxilo. Las modificaciones incluyen, por ejemplo, acetilación, acilación, ADP-ribosilación, amidación, unión covalente de flavina, unión covalente de un resto hemo, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un lípido o derivado de lípido, unión covalente de fosfatidilinositol, reticulación, ciclación, formación de enlaces disulfuro, desmetilación, formación de reticulaciones covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicosilación, formación de anclaje de GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, procesamiento proteolítico, fosforilación, prenilación, racemización, glicosilación, unión de lípidos, sulfatación, gamma-carboxilación de residuos de ácido glutámico, hidroxilación y ADP-ribosilación, selenoilación, sulfatación, adición de aminoácidos mediada por transferencia de ARN a proteínas, tales como arginilación y ubiquitinación. Los polipéptidos pueden ser ramificados o cíclicos, con o sin ramificación. Los polipéptidos cíclicos, ramificados y ramificados circulares pueden resultar de procesos naturales postraduccionales y pueden producirse también por métodos completamente sintéticos.

“Aislado” significa alterado “por la mano del hombre” de su estado natural, es decir, si se produce en la naturaleza, se ha cambiado o retirado de su entorno original, o ambos. Por ejemplo, un polinucleótido o un polipéptido naturalmente presente en un organismo vivo no está “aislado”, pero el mismo polinucleótido o polipéptido separado de los materiales coexistentes de su estado natural está “aislado”, tal como se emplea el término en el presente documento. De manera similar, una secuencia “sintética”, tal como se usa el término en el presente documento, significa cualquier secuencia que se haya generado sintéticamente y no directamente aislada de una fuente natural. “Recombinante” significa ADN modificado por ingeniería genética preparado mediante el trasplante o el corte y empalme de genes de una especie en las células de un organismo huésped de una especie diferente. Dicho ADN se convierte en parte de la composición genética del huésped y se replica.

El término “polinucleótido que codifica un polipéptido” tal como se usa en el presente documento abarca polinucleótidos que incluyen una secuencia que codifica un polipéptido de la invención, particularmente monooxigenasas de *Armillaria gallica*, que tienen las secuencias de aminoácidos tal como se expone en SEQ ID NO 1 a 4. El término también abarca polinucleótidos que incluyen una región continua individual o regiones discontinuas que codifican el polipéptido (por ejemplo, interrumpidas por fago integrado o secuencia de inserción o edición) junto con regiones adicionales que también pueden contener secuencias codificantes y/o no codificantes.

“Variante(s)” tal como se usa el término en el presente documento, es un polinucleótido o polipéptido que difiere de un polinucleótido o polipéptido de referencia respectivamente, pero conserva propiedades esenciales. Una variante típica de un polinucleótido difiere en secuencia de nucleótidos de otro polinucleótido de referencia. Cambios en la secuencia de nucleótidos de la variante pueden o no alterar la secuencia de aminoácidos de un polipéptido codificado por el polinucleótido de referencia. Los cambios de nucleótido pueden dar como resultado sustituciones, adiciones, deleciones, fusiones y truncaciones de aminoácidos en el polipéptido codificado por la secuencia de referencia, tal como se comenta a continuación. Una variante típica de un polipéptido difiere en secuencia de aminoácidos de otro polipéptido de referencia. Generalmente, las diferencias están limitadas de modo que las secuencias del polipéptido de referencia y la variante son estrechamente similares en general y, en muchas regiones, idénticas. Una variante y polipéptido de referencia pueden diferir en la secuencia de aminoácidos en una o más sustituciones, adiciones, deleciones en cualquier combinación. Un residuo de aminoácido sustituido o insertado puede o no ser uno codificado por el código genético. Una variante de un polinucleótido o polipéptido puede ser una que se produce de manera natural tal como una variante alélica, o puede ser una variante que no se conoce que se produzca de manera natural. Pueden producirse variantes de polinucleótidos y polipéptidos que no se producen de manera natural mediante técnicas de mutagénesis, mediante síntesis directa, y mediante otros métodos recombinantes conocidos para los expertos en la técnica.

También se incluyen dentro de la presente invención y/o en la definición de polipéptido variante polipéptidos “ortólogos” (ortólogos), que son péptidos codificados por genes en diferentes especies que evolucionaron a partir de un gen ancestral común por especiación. Normalmente, los ortólogos conservan la misma función en el curso de la evolución.

Además, el término “célula huésped” o “microorganismo huésped” se define por la presente como una célula o microorganismo o célula de microorganismo, que se ha transformado o transfectado, o es capaz de transformación o transfección mediante una secuencia de polinucleótidos exógena, expresando así de manera heteróloga el polinucleótido introducido.

Por la presente y tal como se entiende generalmente, un “éster arílico de tipo protoiludeno sesquiterpenoide” ha de entenderse que representa o bien un armililoreselinato o un meleólido. Como tal, se construyen a partir de un protoiludeno sesquiterpenoide oxigenado y ácido orselínico.

Según el microorganismo huésped y el método según la invención, el polinucleótido consiste en uno de SEQ ID NO 5 a 9 de la lista de secuencias adjunta y codifica un polipéptido con actividad monooxigenasa (SEQ ID NO 5 a 8) o representa una agrupación de genes que comprende 4 monooxigenasas y protoiludeno sintasa. SEQ ID NO 5 a 8 representa las secuencias de cuatro monooxigenasas de *Armillaria gallica* diferentes contenidas en la agrupación de genes para síntesis de meleólidos recién identificada de *Armillaria gallica*, agrupación que también comprende el gen de protoiludeno sintasa. SEQ ID NO 9 muestra la secuencia de la agrupación.

Las SEQ ID NO 5 a 8 tal como se da a conocer en la lista de secuencias adjunta son las cuatro secuencias de ADNc, respectivamente, de las monooxigenasas 1 a 4 tal como se identifica y caracteriza de *Armillaria gallica*. Debe entenderse que otras variantes de las mismas, que tienen al menos una identidad de secuencia del 90%, y que incluso pueden encontrarse en otras especies de *Armillaria*, también son adecuadas y parte de la invención, puesto que con las monooxigenasas recién identificadas se proporcionan herramientas valiosas mediante las cuales pueden identificarse monooxigenasas similares, es decir monooxigenasas que difieren ligeramente de SEQ ID NO 5 a 8, mediante comparación de secuencias y posterior prueba enzimática.

La invención también se refiere a vectores comprendidos en el microorganismo huésped, que comprenden un polinucleótido o polinucleótidos de la invención, microorganismos huésped y/o células huésped que se modifican por ingeniería genética con vectores de la invención y la producción de polipéptidos de la invención mediante técnicas recombinantes.

También pueden emplearse sistemas de traducción libres de células para producir tales polipéptidos usando ARN derivado de los constructos de ADN de la invención.

El vector tal como se emplea en el microorganismo huésped según la invención, contiene una secuencia de ácido nucleico tal como se definió anteriormente, que codifica un polipéptido con actividad monooxigenasa, en el que la secuencia de ácido nucleico se une operativamente a secuencias de control reconocidas por una célula/microorganismo huésped transformado o transfectado con el vector. Según un aspecto de la invención, el vector es un vector de expresión, y, según otro aspecto, el vector puede estar presente en forma de un plásmido, cósmido, fago, liposoma o virus.

Para la producción recombinante, las células huésped pueden modificarse por ingeniería genética para incorporar sistemas de expresión o porciones del mismo o polinucleótidos de la invención. La introducción de un polinucleótido en la célula huésped puede efectuarse mediante métodos descritos en muchos manuales de laboratorio convencionales, tales como Davis *et al.*, Basic Methods in Molecular Biology, (1986), y Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Por tanto, el polinucleótido según la invención puede, por ejemplo, estar comprendido en un vector que va a transformarse/transfectarse de manera estable en células de microorganismo huésped. En el vector, el polinucleótido de la invención está bajo el control de un promotor inducible, de modo que la expresión del gen/polinucleótido puede dirigirse específicamente y, si se desea, el gen o genes pueden sobreexpresarse de esa manera.

Puede usarse una gran variedad de sistemas de expresión para producir los polipéptidos de la invención. Tales vectores incluyen, entre otros, vectores cromosómicos, episómicos y derivados de virus, por ejemplo, vectores derivados de plásmidos bacterianos, de bacteriófago, de transposones, de episomas de levadura, de elementos de inserción, de elementos cromosómicos de levadura, de virus, y vectores derivados de combinaciones de los mismos, tales como los derivados de elementos genéticos de plásmidos y bacteriófagos, tales como cósmidos y fagómidos. Los constructos del sistema de expresión pueden contener regiones de control que regulan así como generan la expresión. Generalmente, puede usarse cualquier sistema o vector adecuado para mantener, propagar o expresar polinucleótidos y/o para expresar un polipéptido en un huésped para la expresión en este aspecto. La secuencia de ADN apropiada puede insertarse en el sistema de expresión mediante cualquiera de una variedad de técnicas bien conocidas y de rutina, tales como, por ejemplo, las expuestas en Sambrook *et al.*, véase lo anterior.

En el presente documento también se da a conocer un péptido aislado que consiste en una secuencia de aminoácidos seleccionada del grupo que consiste en:

(a) una secuencia de aminoácidos mostrada en una de SEQ ID NO: 1 a 4;

b) una secuencia de aminoácidos de una variante alélica de una secuencia de aminoácidos mostrada en SEQ ID NO 1 a 4, en la que dicha variante alélica está codificada por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8, respectivamente;

c) una secuencia de aminoácidos de un ortólogo de una secuencia de aminoácidos mostrada en SEQ ID NO 1 a 4, en la que dicho ortólogo está codificado por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8; y

(d) un fragmento de una secuencia de aminoácidos mostrada en SEQ ID NO 1 a 4, en el que dicho fragmento

comprende al menos 10 aminoácidos contiguos y presenta actividad monooxigenasa.

La invención se refiere a una célula/microorganismo huésped según la reivindicación 1, que se ha transformado/transfectado con el polinucleótido según la invención o que contiene un vector tal como se definió anteriormente y en particular un microorganismo o célula huésped que se selecciona del grupo que consiste en hongos incluyendo levadura y bacterias.

Debe entenderse que el microorganismo huésped expresa de manera heteróloga el polinucleótido recién identificado, es decir al menos una de las monooxigenasas identificadas en el presente documento, preferiblemente en combinación con una protoiludeno sintasa, o expresa de manera heteróloga la agrupación de genes tal como se expone además a continuación.

Según otro aspecto de la invención, se usa un microorganismo huésped, con el ácido nucleico que codifica el polipéptido con actividad monooxigenasa que se adapta al uso de codón del respectivo microorganismo huésped.

Según otra realización de la invención, la célula huésped es una *Saccharomyces spp.*, en particular *Saccharomyces cerevisiae*, una *Aspergillus spp.*, en particular *Aspergillus nidulans*, un homobasidiomiceto, en particular un homobasidiomiceto del género *Armillaria*, en particular *Armillaria gallica*, *Armillaria mellea*, *Armillaria ostoyae*, y otros miembros del género *Armillaria*.

Aún otro aspecto de la invención se refiere a un método para producir meleólidos tal como se reivindica en las reivindicaciones adjuntas y que comprende las etapas de:

a) hacer crecer, en condiciones de nutrientes adecuadas que permiten la producción de un protoiludeno hidroxilado o éster arílico de tipo protoiludeno sesquiterpenoide, un microorganismo tal como se definió anteriormente; y

b) aislar dicho protoiludeno hidroxilado o dicho éster arílico de tipo protoiludeno sesquiterpenoide del microorganismo huésped o del medio de crecimiento.

Según un aspecto de la invención, el protoiludeno hidroxilado se selecciona de 8-alfa-hidroxi-6-protoiludeno; 8-alfa,13-hidroxi-6-protoiludeno; y el éster arílico de tipo protoiludeno sesquiterpenoide se selecciona de meleólido I, armilaridina, meleólido A, meleólido F, meleólido B, meleólido K, evernitato de armililo, armilarina, arnamiol, meleólido J, armilarivina, 10-alfa-hidroxi-meleólido, armililorselinato, meleólido E, 1-O-trifluoroacetil-meleólido E, meleólido H, 5'-O-metil-meledonal, meledonal, arnamial, meleólido C, meleólido D, meledonal A, meledonal C, meledonol, deshidroarmililoreslinato, armilano, armilaridina.

Según aún otro aspecto, el método que comprende las etapas mencionadas anteriormente consiste en las siguientes etapas:

a) hacer crecer, en condiciones de nutrientes adecuadas, un microorganismo huésped transformado o transfectado para expresar una 2-hidroxi-3-metilglutaril-CoA (HMG-CoA) reductasa y para comprender una secuencia de ácido nucleico seleccionada de a) SEQ ID NO 5 a 8 del protocolo de secuencia abarcado, b) una secuencia de ácido nucleico complementaria a SEQ ID NO 5 a 8, y c) secuencias de ácido nucleico que se hibridan en condiciones rigurosas con las secuencias de ácido nucleico definidas en a) y b) o sus hebras complementarias;

b) sobreexpresar la secuencia de ácido nucleico;

c) potenciar así la producción de protoiludeno hidroxilado o éster arílico de tipo protoiludeno sesquiterpenoide en la célula del microorganismo, y

d) aislar dicho protoiludeno hidroxilado o éster arílico de tipo protoiludeno sesquiterpenoide del microorganismo o del medio de crecimiento.

Según una realización de la presente invención, la invención también se refiere a un método para producir un protoiludeno hidroxilado y/o un éster arílico de tipo protoiludeno sesquiterpenoide, comprendiendo el método las etapas de

a) proporcionar un microorganismo huésped que se ha transformado para expresar de manera heteróloga al menos todo lo siguiente: i) una protoiludeno sintasa, y ii) al menos una citocromo P450 monooxigenasa, en la que al menos una de (i) o (ii) es ajena a dicho microorganismo huésped,

b) hacer crecer, en condiciones de nutrientes adecuadas que permiten la producción de los ésteres arílicos de tipo protoiludeno sesquiterpenoide, produciendo así el microorganismo huésped proporcionado en la etapa a) el protoiludeno hidroxilado y, según sea el caso, el éster arílico de tipo protoiludeno sesquiterpenoide.

El microorganismo huésped empleado en el método según la invención comprende por tanto al menos dos genes, es decir una protoiludeno sintasa y al menos una citocromo P450 monooxigenasa, preferiblemente CYP-Arm1 a 4 de *Armillaria gallica* tal como se identifica en la presente invención.

Por consiguiente, el método puede comprender además la etapa c):

c) aislar dicho protoiludeno hidroxilado y/o dicho éster arílico de tipo protoiludeno sesquiterpenoide del microorganismo huésped o del medio de crecimiento.

5 En el método según la invención, la protoiludeno sintasa es una protoiludeno sintasa de uno de los siguientes: *Armillaria spp.*, en particular de *Armillaria gallica*, y/o la citocromo P450 monooxigenasa es CYPArm 1, 2, 3 ó 4 o CYP-Arm 1 a 4 monooxigenasas tal como se identifica en el presente documento.

Por tanto, en el método de la invención, la citocromo P450 monooxigenasa tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:

a) una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO 1 a 4 de la lista de secuencias adjunta;

10 b) una secuencia de aminoácidos de una variante alélica de una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO 1 a 4, en la que dicha variante alélica está codificada por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8;

15 c) una secuencia de aminoácidos de un ortólogo de una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO 1 a 4, en la que dicho ortólogo está codificado por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8.

d) un fragmento funcional de una secuencia de aminoácidos mostrada en SEQ ID NO 2, en el que dicho fragmento comprende al menos 10 aminoácidos contiguos, satisfaciendo la función de una citocromo P450 monooxigenasa.

20 Además, en un refinamiento del método de la invención, en la etapa a) se proporciona un microorganismo que se ha transformado adicionalmente para expresar una NADPH-citocromo P450 reductasa, preferiblemente una NADPH:citocromo reductasa de *Taxus chinensis*.

Esta enzima se introduce con el fin de mejorar la reducción de la enzima heteróloga del citocromo P450 en levadura.

Según el microorganismo huésped y/o el método según la invención, el microorganismo huésped proporcionado en la etapa a) se transforma adicionalmente para expresar una 2-hidroxi-3-metilglutaril-CoA (HMG-CoA) reductasa.

25 Esta enzima se introduce con el fin de aumentar el flujo a través de la ruta de ácido mevalónico para proporcionar difosfato de farnesilo adicional para reforzar la biosíntesis de protoiludeno.

Debe entenderse que preferiblemente el éster arílico de tipo protoiludeno sesquiterpenoide se selecciona de meleólidos y armililorselinatos, en particular de protoiludeno monohidroxilado, en particular 8-hidroxi-6-protoiludeno, protoiludeno dihidroxilado, en particular 8,13-hidroxi-6-protoiludeno.

30 En el presente documento se dan a conocer protoiludeno hidroxilado o el éster arílico de tipo protoiludeno sesquiterpenoide obtenidos de un método según la invención.

Además, la presente invención también se refiere al uso de una agrupación de genes que comprende al menos (i) una protoiludeno sintasa y (ii) al menos una citocromo P450 monooxigenasa para la producción de un protoiludeno hidroxilado y/o un éster arílico de tipo protoiludeno sesquiterpenoide.

35 En una realización preferida del uso y del método según la invención, se usa una agrupación de genes que tiene una secuencia de nucleótidos mostrada en SEQ ID NO 9 de la lista de secuencias adjunta.

Con los métodos dados a conocer y la expresión dirigida de la protoiludeno sintasa y citocromo P450 monooxigenasas es posible producir y acumular protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide en el microorganismo huésped o en el medio en el que está contenido el microorganismo huésped.

40 Debe entenderse que se haya dentro de la destreza y conocimiento de un experto en la técnica, complementar la(s) célula(s) del microorganismo huésped o cultivo de células del microorganismo huésped con precursores adicionales y/o esenciales u otras sustancias, nutrientes o metabolitos que pueden ser necesarios para completar o ayudar en la producción potenciada de protoiludeno hidroxilado o éster arílico de tipo protoiludeno sesquiterpenoide.

45 Se dan a conocer en el presente documento protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide obtenidos mediante el método tal como se definió anteriormente. Los protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide así generados pueden usarse de manera eficaz como medicamentos, por ejemplo, para tratar cualquier infección o enfermedad provocada por bacterias.

50 Debe entenderse que la producción de protoiludenos hidroxilados o ésteres arílicos de tipo protoiludeno sesquiterpenoide según la invención puede realizarse por medio de una sobreexpresión heteróloga u homóloga del polinucleótido que codifica la protoiludeno sintasa. Los sistemas y métodos, así como las respectivas células huésped adecuadas, serán evidentes para los expertos en la técnica tras la lectura de las enseñanzas de esta

invención.

Se derivan ventajas adicionales de la descripción de las realizaciones y los dibujos adjuntos.

5 No hace falta decir que las características mencionadas anteriormente y las características que aún deben explicarse a continuación pueden usarse no sólo en las combinaciones especificadas respectivamente, sino también en otras combinaciones o por sí mismas, sin apartarse del alcance de la presente invención.

Breve descripción de las figuras

Varias realizaciones de la invención se ilustran en las figuras y se explican con más detalle en la siguiente descripción. En las figuras:

10 la figura 1 muestra un esquema de la síntesis de meleólido I que implica la parte de la síntesis de terpeno que comienza con la ciclación de difosfato de farnesilo a 6-protoiludeno, seguida por la oxigenación por medio de citocromo P450 monooxigenasas; la etapa final es la unión de ácido orselínico obtenido de la biosíntesis de policétido mediante una aciltransferasa;

15 la figura 2 muestra un esquema de la parte de biosíntesis de terpeno de la agrupación de genes de meleólido; la agrupación contiene la protoiludeno sintasa, cuatro citocromo P450 monooxigenasas y dos proteínas hipotéticas (figura 2A); las figuras 2B a 2E presentan las secuencias de las cuatro citocromo P450 monooxigenasas identificadas;

la figura 3 muestra resultados de inmunotransferencia de tipo Western de diferentes muestras de preparación de proteínas microsomales de citocromo P450 monooxigenasa CYP-Arm 1 a 4 que expresan clones de *Saccharomyces cerevisiae* y el control negativo;

20 la figura 4 muestra un ensayo no radiactivo y análisis de CG/EM de extracto de disolvente obtenido de alimentación *in vivo* con 6-protoiludeno: el extracto de la alimentación *in vivo* del clon que expresa CYP-Arm3 (B) revela un pico adicional en comparación con el control negativo (A) a un tiempo de retención de 10,6 min;

la figura 5 muestra un esquema de la levadura de panadería de gemación *Saccharomyces cerevisiae* para la producción de un supuesto hidroxiprotoiludeno;

25 la figura 6 muestra un esquema que presenta la estructura de hidroxiprotoiludeno obtenido de fermentación que se elucidó mediante espectroscopía de RMN como 8-hidroxi-protoiludeno.

Descripción detallada de realizaciones preferidas

30 La figura 1 muestra un dibujo esquemático de algunos representantes de ésteres arílicos de tipo protoiludeno sesquiterpenoide, que se producen por miembros del género *Armillaria*. Los dos tipos de ésteres arílicos de tipo protoiludeno sesquiterpenoide se dividen según la posición del doble enlace en armililorselinatos y meleólidos (figura 1).

Material y métodos

35 *Cepas y condiciones de crecimiento:* se cultivó la cepa de *Armillaria gallica* FU02472 tal como se describió previamente (Engels *et al.*, "Cloning and Characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides, J. Biol. Chem. 286:6871-6878 (2011)). Se obtuvo la cepa de levadura *S. cerevisiae* CEN.PK2-1C (MATa, ura3-52, trp1-289, Leu2-3_112, his3 Δ1, MAL2-8c, SUC2) de EUROSCARF (EUROpean Saccharomyces Cerevisiae Archive for Functional analysis) Universität Frankfurt (Frankfurt am Main, Alemania) y se mantuvo en agar de medio mínimo SC (Engels *et al.*, "Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards TAXol production", Metab. Eng. 10, 201-206 (2008)). Se usó medio YP tamponado (triptona al 2% p/v, extracto de levadura al 1% p/v y MES 50 mM, pH5,5) con glucosa al 2% p/v en cultivo discontinuo y galactosa al 2% p/v en fermentación semicontinua para inducción de la expresión impulsada por el promotor Gal1. Se realizó la transformación de *S. cerevisiae* usando el método de acetato de litio con selección previamente descrito (Engels *et al.*, 2008).

45 *Aislamiento de ADN genómico y construcción de bibliotecas:* se aisló ADN genómico de *Armillaria gallica* usando el método de bromuro de cetiltrimetilamonio (Murray y Thompson, 1980). Para la construcción de bibliotecas genómicas se usaron el kit de vector Lambda DASH II/BamHI en combinación con los extractos de empaquetamiento Gigapack III Gold según las pautas del fabricante (Stratagene, Heidelberg, Alemania). Para la construcción de una biblioteca de "paseo" (*walking*) de genoma de *A. gallica*, se usó el kit Clontech (Saint-Germain-en-Laye, Francia) GenomeWalker™ Universal según el protocolo del fabricante.

50 *Examen de bibliotecas, secuenciación y paseo genómico:* se transfirieron plagas de fago a membranas de nailon cargadas (GE Healthcare) y se examinaron usando sondas de ácido nucleico marcadas con ³²P. Para la prehibridación e hibridación Roti®Hybrid Quick (Carl Roth, Karlsruhe, Alemania) complementada con ADN de esperma de salmón 0,1 mg/ml. Para el marcaje de las sondas, se usó el kit de marcaje de ADN DecaLabel™

(Fermentats, St. Leon-Rot, Alemania; desde 2010 Thermo Fisher Scientific). Las membranas se hibridaron durante 15 h a 65°C y luego se lavaron cuatro veces con tampón de hibridación Roti®Hybrid Quick diluido (Roti®Hybrid Quick:H₂O 1:2 durante 30 min, 1:5 durante 30 min y 1:10 dos veces durante 10 min). Las membranas se expusieron entonces a película de rayos X a -80°C durante 4 días. Entonces se secuenciaron cósmidos positivos aislados mediante paseo con cebadores. Los fragmentos obtenidos de la amplificación por PCR del paseo genómico se subclonaron para secuenciación de ADN en el kit de clonación Zero Blunt® (Invitrogen, Karlsruhe, Alemania).

Expresión heteróloga de citocromo P450 en levadura y pruebas funcionales por medio de alimentación in vivo: se clonaron las cuatro monooxigenasas dependientes del citocromo P450 de *A. gallica* por medio de clonación Gateway en el vector de expresión de levadura inducible de galactosa pYES-DEST52 (Invitrogen). Para la expresión heteróloga de las citocromo P450 monooxigenasas, se usó la cepa de levadura haploide CEN.PK2-1C.

Para mejorar la reducción de la enzima de citocromo P450 heterólogo en levadura, se clonó NADPH:citocromo P450 reductasa de *Taxus chinensis* (Jennewein *et al.*, "Coexpression in yeast of Taxus cytochrome P450 reductase with cytochrome P450 oxigenases involved in Taxol biosynthesis" Biotechnol. Bioeng. 89:588-598 (2005)) en el vector de expresión de levadura pCM183 para coexpresión heteróloga. Para la preparación de la fracción microsómica de levadura se usaron aproximadamente 5 g (peso húmedo) de biomasa de levadura recombinante inducida con galactosa para la preparación de esferoblastos. Después de la alteración de los esferoblastos usando perlas de vidrio, se limpió el extracto bruto resultante mediante centrifugación y se recogieron membranas ER del sobrenadante por medio de una ultracentrifugación 105.000xg durante tres horas. Para las pruebas funcionales por medio de alimentación *in vivo*, se coexpresaron de manera heteróloga las citocromo P450 monooxigenasas de *A. gallica* con la citocromo P450 monooxigenasa de *T. chinensis*. Para las cepas inducidas con galactosa se usó o bien 6-protoiludeno frío o bien [³H]-6-protoiludeno marcado con tritio. Después de incubación durante la noche, se añadió 1 ml de salmuera a los cultivos de células de levadura inducidos y se extrajo el cultivo tres veces con 5 ml de n-pentano. Se agruparon y se concentraron las fracciones de disolvente orgánico para radiocromatografía en capa fina (CCF) o análisis de CG/EM.

Fermentación de S. cerevisiae recombinante para la síntesis total de hidroxiprotoiludeno: para la biosíntesis total de cantidades suficientes de hidroxiprotoiludeno para elucidación de la estructura por medio de análisis de RMN, se transformaron las cepas de levadura CEN.PK2-1C que coexpresan la citocromo P450 monooxigenasa CYP-Arm3 de *A. gallica* con la NADPH:citocromo P450 reductasa de *T. chinensis* con una versión truncada de la HMG-CoA reductasa 1 de levadura (Dimster-Denk *et al.*, "Feedback regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Saccharomyces cerevisiae*", Mol.Biol.Cell 5:655-665 (1994)) clonadas en el vector pRS315 que contiene un promotor PGK y terminador de la transcripción *cyc1*. Se clonó protoiludeno sintasa de *A. gallica* en el vector pRS423 (que complementa la auxotrofia de histidina de levadura CEN.PK2-1C) que contiene un promotor Gal1 y un terminador de la transcripción *cyc1* (ambos obtenidos del vector pYES-DEST52).

Espectrometría por RMN: se registraron todos los espectros de RMN en un espectrómetro de RMN de 500 MHz DRX de Bruker usando una sonda BBO de 5 mm con un gradiente en z. Se disolvieron muestras de hidroxiprotoiludeno en diclorometano deuterado hasta una concentración final de aproximadamente 50 mM y se calibraron con respecto a las señales de disolvente a 5,32 ppm y 54,00 ppm, respectivamente. Se adquirió espectro de ¹³C-RMN mediante 1000 tránsitos con un pulso de 30° y un retardo de 2 segundos. Se adquirió el ¹³C-DEPT-135 mediante 400 tránsitos. Se adquirieron todos los espectros bidimensionales usando programas de pulso convencional de BRUKER. Se midieron los espectros DQF-COSY y gs-NOESY con 512 incrementos con dos o cuatro tránsitos por incremento, respectivamente. Se usó un tiempo de mezclado de 600 ms para el espectro NOESY. Se midió la correlación heteronuclear mediante espectros gs-HSQC y gs-HMBC, se adquirió con 512 incrementos con cuatro y ocho tránsitos por incremento, respectivamente. Ambos espectros se optimizaron para acoplamiento ¹J y ⁿJ de 140 Hz y 8 Hz, respectivamente. Se procesaron los datos usando software TopSpin 1.3 (Bruker Biospin) y MestReNova 8.0 (MestreLab Research).

Resultados

Aislamiento de agrupación de genes biosintética usando examen de biblioteca de fagos y paseo genómico

Para el aislamiento de la agrupación de genes de tipo protoiludeno sesquiterpenoide se construyó una biblioteca de fagos de ADN genómico de *Armillaria gallica*; se realizó el examen usando protoiludeno sintasa de *A. gallica* como sonda. Se usaron dos de las placas positivas obtenidas de la segunda ronda de examen para aislamiento de ADN de cósmido. La digestión de restricción del ADN de cósmido obtenido mostró patrones de fragmento idénticos que condujeron a la conclusión de que ambos cósmidos son idénticos. El inserto de uno de los cósmidos (cósmido 5.1) se secuenció entonces mediante paseo con cebadores que reveló un fragmento de 22,902 kb.

El análisis de los datos de secuencia obtenidos identificó la secuencia genómica de protoiludeno sintasa aislada previamente y tres secuencias con alta homología con monooxigenasas dependientes del citocromo P450 de las cuales una parece ser una secuencia parcial. Además de la protoiludeno sintasa y las citocromo P450 monooxigenasas, se identificaron dos secuencias con funciones desconocidas (marcadas como proteínas hipotéticas) y una secuencia con alta homología con una GMC oxidorreductasa en el fragmento de ~ 23 kb.

Con los ésteres arílicos de tipo protoiludeno oxigenándose en hasta ocho posiciones, la identificación de tres citocromo P450 monooxigenasas similares asociada estrechamente con la protoiludeno sintasa ya proporcionó una primera indicación de un posible agrupamiento de la ruta biosintética en *A. gallica*.

5 A continuación se construyó una biblioteca de paseo genómico ligando fragmentos adaptadores de ADN con ADN genómico de *A. gallica* digerido por restricción y mediante secuenciación de los fragmentos de PCR obtenidos, pudieron completarse las citocromo P450 monooxigenasas parciales encontradas del inserto de cósmido de la secuencia. Entonces se usó parte de la secuencia recién identificada para examinar $1,7 \times 10^4$ unidades formadoras de placa adicionales de la biblioteca de fagos genómica. El examen dio como resultado un clon que también demostró ser positivo en una posterior segunda ronda de examen. La secuenciación del inserto de cósmido reveló un fragmento genómico de *A. gallica* de 16,27 kb que contenía una citocromo P450 monooxigenasa adicional (denominada CYP-Arm1) además de tres supuestos genes con homología con una endonucleasa de reparación de ADN, un activador de la transcripción y una transcriptasa inversa. Un intento de paseo con cebadores 3' al fragmento de cósmido 5.1 no reveló más genes con posible función en la biosíntesis de terpenoides o policétido. La figura 2 muestra el esquema de la parte de biosíntesis de terpeno de la agrupación de genes de meleóido obtenida del examen de la biblioteca de ADN genómico con sondas radiactivas (protoiludeno sintasa); la agrupación contiene la protoiludeno sintasa ("Pro 1"), cuatro citocromo P450 monooxigenasas, denominadas "CYP-Arm1", "CYP-Arm2", "CYP-Arm3" y "CYP-Arm4", y dos proteínas hipotéticas.

Amplificación de las citocromo P450 monooxigenasas identificadas de la biblioteca de ADNc de A. gallica

20 La secuenciación del ADN genómico clonado reveló un alto grado de fragmentaciones de intrones/exones de las cuatro secuencias de citocromo P450 monooxigenasa genómicas encontradas. La secuencia genómica de citocromo-P450 CYP-Arm1 que se extiende más de 2158 pb contiene 10 intrones, con exones que varían de tamaño desde 38 pb hasta 537 pb. Las secuencias genómicas de CYP-Arm2 y CYP-Arm3 muestran alta similitud, conteniendo ambas 11 intrones y presentando una identidad de secuencia global del 73,8%. El análisis de CYP-Arm4 identificó un marco de lectura abierto que codifica una proteína de 509 aminoácidos, se interrumpe el ORF deducido en 14 intrones dando exones de 4 a 190 pb de longitud. El análisis adicional de las citocromo P450 monooxigenasas aisladas en los niveles de secuencia de aminoácidos presenta una similitud de secuencia de aminoácidos de CYP-Arm1, CYP-Arm2 y CYP-Arm3 en comparación con CYP-Arm4 del 32,5%, el 34,0% y el 34,0%, respectivamente.

30 Basándose en estas secuencias predichas, se amplificaron los correspondientes ADNc mediante reacción en cadena de la polimerasa a partir de una biblioteca de ADNc de *A. gallica*. La secuenciación de varios fragmentos de PCR subclonados reveló una fracción significativa de los clones de ADNc que todavía contienen intrones. Para las cuatro citocromo P450 monooxigenasas, sin embargo, se obtuvieron clones de ADNc correctos y se subclonaron para determinar la expresión heteróloga en *Saccharomyces cerevisiae*. Para facilitar el análisis de la correcta expresión heteróloga y ubicación en el retículo endoplasmático de levadura, las cuatro citocromo P450 monooxigenasas se marcaron con histidina en el extremo C-terminal.

Expresión heteróloga de citocromo P450 monooxigenasas y examen funcional.

40 Para la expresión heteróloga, los cuatro constructos de expresión construidos que contienen las citocromo P450 monooxigenasas aisladas de *A. gallica* se transformaron en CEN.PK2-1C de levadura. Tras la inducción de la expresión heteróloga con L-galactosa, se recogieron las células de levadura recombinantes y se examinó la ubicación correcta del P450 expresado de manera heteróloga por medio de la preparación de proteína microsómica de levadura. El análisis de inmunotransferencia que usa un anticuerpo primario específico de hexa-histidina mostró la ubicación correcta de la proteína heteróloga expresada (con un peso molecular esperado de 57 a 60 kDa) con respecto a la fracción microsómica (véase la figura 3).

45 Para pruebas funcionales de las monooxigenasas dependientes del citocromo P450 clonadas, la NADPH:citocromo P450 reductasa de *Taxus chinensis* se coexpresó en la levadura recombinante. Para un examen funcional inicial del citocromo P450 clonado de *A. gallica* se tomó un enfoque de alimentación *in vivo*. Para el experimento de alimentación *in vivo*, se le añadieron 15.000 cpm de 6-protoiludeno marcado con ^3H a un cultivo de levadura de 5 ml inducido y los cultivos se incubaron durante la noche. Entonces se analizaron los extractos orgánicos de los cultivos de levadura individuales por medio de radiocromatografía en capa fina (radio-CCF). El análisis de radio-CCF para CEN.PK2-1C que expresa la CYP-Arm3 junto con la citocromo P450 reductasa de *T. chinensis* mostró una conversión casi completa del 6-protoiludeno marcado con tritio de alimentación en un supuesto producto hidroxilado más polar. Además, con CYP-Arm2, se observó un producto más polar, sin embargo la conversión se produjo en este caso en un grado mucho menor. De manera interesante, la reextracción del 6-protoiludeno proporcionado a partir del cultivo de células de levadura usando n-pentano demostró ser difícil, sin embargo parecía mejorar significativamente para derivados hidroxilados.

55 Para análisis adicional usando CG/EM, se alimentaron las cepas de levadura recombinante con 6-protoiludeno "frío" derivadas de la reacción de síntesis basada en biocatálisis usando difosfato de farnesilo y protoiludeno sintasa recombinante. Los cultivos de levadura recombinante inducidos y alimentados *in vivo* se extrajeron entonces de manera similar con n-pentano y se analizó el extracto orgánico por medio de CG/EM. El CG-cromatograma de CYP-

Arm3 mostró en comparación con el control negativo (que sólo expresa la NADPH:citocromo P450 reductasa de *T. chinensis*) un pico adicional (figura 4). El espectro de masas obtenido mostró masas correspondientes a un producto de hidroxiprotoiludeno.

Biosíntesis total de hidroxiprotoiludeno mediante S. cerevisiae recombinante.

- 5 A continuación, se elucidó la estructura del posible producto de hidroxiprotoiludeno mediante análisis por RMN. Puesto que para la elucidación estructural adecuada del compuesto se requieren varios miligramos de producto puro, fue necesario obtener suficiente material. Para este fin, y para la purificación del compuesto a homogeneidad y análisis, se modificó por ingeniería genética una cepa de *Saccharomyces cerevisiae* que podía realizar biosíntesis total del supuesto hidroxiprotoiludeno.
- 10 Para la fermentación total, se empleó la cepa de *S. cerevisiae* CEN.PK2-1C transformada con vectores de expresión episomales que contienen la NADPH:citocromo P450 reductasa de *T. chinensis*, las citocromo P450 monooxigenasas CYP-Arm3, una versión truncada de la HMG-CoA reductasa (tHMGR) de levadura y la protoiludeno sintasa de *A. gallica*. Se coexpresó la NADPH:citocromo P450 reductasa de *T. chinensis* de manera heteróloga con el fin de aumentar la reducción y, por tanto, la actividad catalítica de la citocromo P450 monooxigenasa. Se incluyó la versión truncada, desregulada de la HMG-CoA reductasa de levadura para aumentar el flujo a través de la ruta de ácido mevalónico para proporcionar difosfato de farnesilo adicional para reforzar la biosíntesis de protoiludeno (véase la figura 6).

La tabla 1 a continuación muestra un resumen de los plásmidos transformados y los genes expresados de manera heteróloga en la cepa de *S. cerevisiae* recombinante para la producción de hidroxiprotoiludeno:

20 Tabla 1:

Plásmido	ORF	Promotor	Auxotrofia
pCM183:: <i>P450-Red-tc</i>	Citocromo P450 reductasa	P _{cyc1}	Triptófano (<i>TRP1</i>)
pRS315:: <i>tHMGR-sc</i>	3-hidroxi-3-metilglutaril-CoA-reductasa	P _{PGK}	Leucina (<i>LEU2</i>)
pRS423:: <i>Pro1HIS6</i>	protoiludeno sintasa	P _{Gal1}	Histidina (<i>HIS3</i>)
PYES-DEST52:: <i>CYP-Arm3HIS6</i>	citocromo P450 monooxigenasa CYP-ARM3	P _{Gal1}	Uracilo (<i>URA3</i>)

Se llevó a cabo la fermentación para la síntesis total del supuesto producto de hidroxiprotoiludeno usando la cepa metabólica modificada por ingeniería genética en una escala de 2,8 l. Para la purificación del producto CYP-Arm3, se empleó el sobrenadante del cultivo, separado de la biomasa.

25 *Purificación de hidroxiprotoiludeno y elucidación de la estructura*

- Se extrajo el sobrenadante de cultivo obtenido mediante adición de polvo de sílice C18 RP e incubación a 4°C durante Se recuperó el material de sílice C18 RP mediante filtración y se secó mediante liofilización. Se realizó la extracción usando un aparato Soxhlet con n-pentano como disolvente orgánico. Se concentró el extracto resultante a vacío. Entonces se purificó el extracto en bruto mediante cromatografía sobre gel de sílice 60, gel de sílice con funcionalidad cloropropilo y mediante HPLC semipreparativa de fase inversa. Después de tres etapas de purificación, pudieron obtenerse aproximadamente 40 mg de producto puro para análisis de RMN.

Para la determinación estructural se realizaron tanto análisis de RMN de ¹H como de ¹³C. La tabla 2 a continuación muestra los desplazamientos de RMN de ¹H y de ¹³C completos y sus asignaciones como resultado del análisis de espectros de RMN bidimensionales:

35

Tabla 2:

Número	Señales de ¹ H	Señales de ¹³ C
1	1,35 (dd) 1,4 (dd)	41,5 (T)
2	2,4 (dt)	46,7 (D)
3	-	45,6 (S)
4	1,84 (2H, m)	36,6 (T)
5	2,5 (m) 2,7 (m)	25,0
6	-	141,5 (S)
7	-	126,8 (S)
8	3,98 (d, ancho)	74,5 (D)
9	2,2 (m)	51,1 (D)
10	1,16 (t)	46,9 (T)
11	-	40,05 (S)
12	1,06 (3H, s)	20,4 (Q)
13	1,64 (3H, s)	11,1 (Q)
14	1,11 (3H, s)	29,6 (Q)
15	0,99 (3H, s)	22,1 (Q)
-OH	1,47 (ancho)	-

5 El análisis mostró que el compuesto aislado era 8-hidroxi-6-protoiludeno, que es por lo que CYP-Arm3 ha demostrado actuar como una 8-alfa-hidroxilasa. Además, los resultados mostraron, por primera vez, la caracterización funcional, junto con la elucidación de la estructura molecular del producto, de una catalización mediada por citocromo P450 monooxigenasa a un alcohol de sesquiterpeno a partir de un basidiomiceto.

En una siguiente etapa, se elucidó si la expresión adicional de CYP-Arm2 podría conducir a una estructura de protoiludeno dihidroxilada.

10 Para este fin, el conjunto de plásmidos que iba a emplearse tuvo que adaptarse para la expresión de un quinto gen foráneo. La siguiente tabla 3 muestra los plásmidos transformados y los genes expresados de manera heteróloga en la cepa de *S. cerevisiae* recombinante para la producción de protoiludeno dihidroxilado:

Tabla 3:

Plásmido	ORF	Promotor	Auxotrofia
pCM183:: <i>P450-Red-tc</i>	Citocromo P450 reductasa	P _{cyc1}	Triptófano (<i>TRP1</i>)
pRS315:: <i>thmgr-sc</i>	3-hidroxi-3-metilglutaril-CoA-reductasa	P _{PCK}	Leucina (<i>LEU2</i>)
pRS423:: <i>Pro1HIS6::CYP-Arm3HIS6</i>	protoiludeno sintasa y citocromo P450 monooxigenasa CYP-Arm3	2 x P _{Gal1}	Histidina (<i>HIS3</i>)
PYES-DEST52:: <i>CYP-Arm2HIS6</i>	Citocromo P450 monooxigenasa CYP-ARM2	P _{Gal1}	Uracilo (<i>URA3</i>)

5 Para la fermentación de un protoiludeno dihidroxilado, se llevó a cabo una fermentación a una escala de 3 l, de manera análoga a la fermentación descrita anteriormente para la producción del 8-alfa-6-protoiludeno monohidroxilado. Para la purificación del producto, el sobrenadante de cultivo, separado de la biomasa, se sometió a una extracción con cloroformo. Se analizó el extracto concentrado por medio de CG/EM y CL/EM/EM, de manera análoga al proceso de aislamiento del 8-alfa-6-protoiludeno.

Purificación de protoiludeno dihidroxilado y elucidación de la estructura

Se extrajo el sobrenadante de cultivo obtenido en cloroformo-d1 y se secó al aire, se disolvió en disolvente deuterado, y se analizó por medio de CG/EM.

10 Para la determinación estructural adicional, se realizaron análisis tanto de RMN de ^1H como de ^{13}C directos. La tabla 4 a continuación muestra los desplazamientos de RMN de ^1H y de ^{13}C completos y sus asignaciones como resultado del análisis de espectros de RMN bidimensionales:

Tabla 4:

Número	Señales de ^1H	Señales de ^{13}C
1	1,21 (dd) 1,32 (dd)	41,2 (T)
2	2,32 (dt)	46,1 (D)
3	-	45,6 (S)
4	1,78 (2H, m)	36,1 (T)
5	2,58 (m) 2,7 (m)	24,8
6	-	141,9 (S)
7	-	126,9 (S)
8	4,07 (d, amplio)	74,8 (D)
9	2,25 (m)	51,1 (D)
10	1,06 (t) 1,7(dd)	46,3 (T)
11	-	39,8 (S)
12	0,98 (3H, s)	20,2 (Q)
13	4,15 (3H, s)	59,0 (Q)
14	1,02 (3H, s)	29,5 (Q)
15	0,89 (3H, s)	27,0 (Q)
-OH	(ancho)	-

15 El análisis de los espectros mostró que el compuesto aislado era 8-alfa-13-hidroxi-6-protoiludeno. Como consecuencia, la citocromo P450 monooxigenasa CYP-Arm2 se ha caracterizado que actúa como una 13-hidroxilasa.

Con la biosíntesis introducida de manera heteróloga descrita, se ha producido por primera vez un alcohol de sesquiterpeno dihidroxilado en *S. cerevisiae* y se ha caracterizado estructuralmente.

Es posible la identificación de la agrupación de genes de *A. gallica* descrita en el presente documento, la identificación de otras rutas de biosíntesis de sesquiterpenos en basidiomicetos.

Lista de secuencias

<110> Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V.

<120> Agrupación de genes para síntesis de meleóidos y su aplicación

<130> Documento 2067P101WO

<160> 9

<170> PatentIn versión 3.5

<210> 1

<211> 538

<212> PRT

<213> *Armillaria gallica*

<400> 1

Met Ala Leu Phe Ser Ala Tyr Ala Leu Ala Phe Ser Leu Leu Met Val
1 5 10 15

Pro Leu Ile Leu Tyr Ile Leu Arg Ile Gly Arg Arg Glu Ser Gly Leu
20 25 30

Pro Pro Gly Pro Ser Thr Leu Pro Leu Val Gly Asn Leu His Gln Leu
35 40 45

Pro His Gly Ser Leu His Leu Gln Phe Thr Ala Trp Ala Lys Glu Phe
50 55 60

Gly Gly Ile Phe Ser Leu Lys Phe Gly Pro Gly Thr Val Ile Val Ala
65 70 75 80

Thr Ser Pro Arg Ala Val Arg Glu Leu Ile Asp Gln Lys Ser Ala Ser
85 90 95

Thr Ser Asp Arg Pro Pro Ser His Phe Ser Asn Val Ile Thr Gly Gly
100 105 110

Asn Asn Ile Gly Phe Ala Arg Tyr Ser Asp Tyr Trp Arg Arg Gly Arg
115 120 125

Arg Val Met His Ser Met Leu Thr Lys Lys Ala Cys Val Asn His Leu
130 135 140

Thr Ile Gln Arg Ala Glu Ala Ser Gln Leu Met Tyr Asp Tyr Leu Val
145 150 155 160

Glu Pro Lys Val Asn Leu Ser His Gln Ile Leu Val Ala Asn Val Gly
165 170 175

ES 2 747 844 T3

Lys Ser Asp Val Gln Cys Gln Glu Phe Val Ala His Gly Gln Arg Tyr
 180 185 190

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

His Thr Ser Pro Leu Val Thr Ala Phe Phe Gln Met Gln His Glu Trp
 210 215 220

Thr His Leu Leu Thr Pro Gly Ala His Pro Pro Val Asp Met Val Pro
 225 230 235 240

Phe Leu Lys Tyr Ile Pro Gly Ser Trp Lys Gln Ile Cys Ala Lys Met
 245 250 255

Lys Val Ser Gln Glu Glu Leu Tyr Gly Gly Met Ile Asp Ala Cys Ala
 260 265 270

Lys Arg Val Glu Arg Gly Ile Arg Asn Gly Cys Phe Leu Glu Gly Glu
 275 280 285

Leu Glu Asn Lys Asp Val Asp Arg Gly Leu Leu Arg Gly Met Cys Ser
 290 295 300

Ala Leu Met Glu Gly Gly Ser Asp Ser Thr Ser Ile Tyr Leu Gln Ser
 305 310 315 320

Phe Ile Leu Met Leu Val Ala His Pro Asp Val Gln Ala Lys Ala Arg
 325 330 335

Ala Glu Ile Asp Ser Val Val Gly Leu Asp Arg Leu Pro Asp Ile Glu
 340 345 350

Asp Met Asp Asn Leu Pro Tyr Val Ser Ala Val Ile Lys Glu Val Leu
 355 360 365

Arg Leu Arg Pro Ile Thr Ala Leu Gly Ala Pro His Tyr Ser Thr Gln
 370 375 380

Pro Glu Val Val Asp Gly Tyr Val Ile Pro Lys Asn Ser Met Ile Phe
 385 390 395 400

Met Asn Gln Trp Gly Met Leu His Asp Pro Asp Ser Tyr Asp Gln Pro
 405 410 415

Glu Ser Phe Met Pro Glu Arg Phe Leu Thr Ser Pro Phe Gly Thr Arg
 420 425 430

ES 2 747 844 T3

Pro Gly Ala Asp Asp Thr Gly Arg Ser Lys Asp Ile Tyr Phe Gly Gly
435 440 445

Gly Arg Arg Ile Cys Val Gly Met His Phe Gly Gln Asn Ser Leu Ala
450 455 460

Ile Thr Ser Met Tyr Leu Ile Trp Ala Phe Asn Leu Thr Asn Ala Ile
465 470 475 480

Asp Pro Gln Thr Lys Asn Pro Ile Pro Val Asp Ile Asn Glu Tyr Asp
485 490 495

Ile Ser Leu Asn Thr Ile Pro Lys Pro Phe Lys Cys Asp Phe Gln Val
500 505 510

Arg Ser Glu Glu His Lys Arg Met Val Glu Asn Ala Phe Ala Glu Ser
515 520 525

Arg Gln Val Leu Ala Pro Phe Glu Gln Asp
530 535

<210> 2

<211> 513

<212> PRT

<213> *Armillaria gallica*

<400> 2

Met Thr His Ala Ser Ser Ala Trp Phe Leu Ala Ala Val Ala Ile Val
1 5 10 15

Thr Phe Ile Val Val Arg Arg Ile Arg Ser Ser Trp Arg Lys Leu Pro
20 25 30

Pro Gly Pro Arg Gly Leu Pro Ile Val Gly Asn Leu Leu Gln Leu Arg
35 40 45

Ser Lys Gln Trp Leu Thr Phe Thr Glu Leu Gly Lys Lys Tyr Gly Asp
50 55 60

Leu Met Tyr Phe Asn Val Ala Gly Gln Pro Leu Ile Val Ile Asn Ser
65 70 75 80

Leu Arg Val Ala Thr Asp Leu Leu Asp Arg Ala Lys Phe Ser Asp Arg
85 90 95

Pro Arg Asn Ile Val Ala Ser Asp Ile Met Thr Arg Gly Met Phe Val
100 105 110

ES 2 747 844 T3

Ala Phe Ala Pro Tyr Gly Asn Ala Trp Arg His Met Arg Lys Ala Ala
115 120 125

His Glu Gly Leu Asn Lys Asn Ile Val Asn Gln Tyr His Pro Ile Gln
130 135 140

Ile Lys Glu Ala Val Leu Leu Ala Asn Asp Leu Leu Ala Glu Pro Asn
145 150 155 160

Arg Trp Val Ser His Val Arg Arg Thr Ala Ala Ser Thr Ile Met Ser
165 170 175

Ile Val Tyr Asp Lys Pro Ala Thr Ser Glu Gln Asp Pro Ser Ile Lys
180 185 190

Arg Ile Asn Asp Phe Ala Ala Arg Leu Thr Arg Ala Ala Met Pro Gly
195 200 205

Ala His Phe Val Glu Ser Phe Pro Trp Met Leu Arg Ile Pro Ser Lys
210 215 220

Tyr Ala Lys Trp Lys Arg Glu Ala Glu Gly Trp Tyr Ala Lys Asp Ser
225 230 235 240

Ser Met Phe Glu Ser Leu Phe His Ser Val Lys Asp Arg Val Ala Glu
245 250 255

Gly Asn Asn Arg Pro Ser Phe Val Ala Thr Leu Ile Gln Gly Ala Gly
260 265 270

Arg His Gly Leu Thr Asp His Glu Ser Ser Trp Leu Ala Gly Ala Met
275 280 285

Tyr Thr Ala Gly Val Glu Ser Ser Ser Ala Ala Ile Ser Trp Trp Met
290 295 300

Leu Ala Met Ile Leu Tyr Pro Asp Ala Gln Lys Arg Ala Gln Ala Glu
305 310 315 320

Leu Asp Lys Val Val Gly Arg Asp Arg Leu Pro Ala Phe Ser Asp Tyr
325 330 335

Glu His Leu Pro Tyr Val Arg Ala Met Val Lys Glu Thr Leu Arg Trp
340 345 350

Arg Ala Val Asp Pro Val Gly Leu Pro His Arg Ser Thr Glu Asp Asp

ES 2 747 844 T3

Gly Val Met Ser Trp Trp Thr Leu Ala Met Ile Val Tyr Pro Glu Thr
305 310 315 320

Gln Lys Arg Ala Gln Ala Glu Leu Asp Ala Val Val Gly Arg Asp Arg
325 330 335

Leu Pro Ser Phe Ala Asp Tyr Glu His Leu Pro Tyr Ile Arg Ala Met
340 345 350

Val Lys Glu Ala Leu Arg Trp Arg Met Val Asp Pro Val Gly Leu Pro
355 360 365

His Thr Ser Thr Glu Asp Asp Val Tyr Asp Gly Tyr Phe Ile Pro Ala
370 375 380

Gly Thr Ile Ile Ile Ala Asn Val Trp His Leu Asn Arg Asp Pro Glu
385 390 395 400

Ile Tyr Gly Pro Asp Ala Glu His Phe Asn Pro Ala Arg His Leu Asp
405 410 415

Lys Asp Gly Lys Leu Ala Pro Gly Pro Ala Asp Thr Lys Glu Glu Ser
420 425 430

His Val Thr Tyr Gly Phe Gly Arg Arg Ile Cys Val Gly Arg His Val
435 440 445

Ala Asn Asn Ser Leu Phe Ile Asp Ile Ala Met Met Leu Trp Ala Met
450 455 460

Asn Ile Glu Arg Ala Thr Asp Glu Asn Gly Val Pro Leu Pro Leu Asp
465 470 475 480

Val Asp Gly Cys Val Glu Asp Gly Leu Val Thr Arg Pro Val Pro Phe
485 490 495

Lys Ala Lys Ile Thr Pro Arg Phe Arg Glu Ala Gln Ala Ile Val Glu
500 505 510

Gln Glu Arg Glu Leu Leu Gly Tyr His
515 520

<210> 4

<211> 509

<212> PRT

<213> *Armillaria gallica*

<400> 4

ES 2 747 844 T3

Met Asp Ser Ala Ser Leu Ala Val Val Val Trp Ala Ile Leu Leu Val
1 5 10 15

Leu Trp Leu Arg Arg Ile Phe Gly Gln Arg Ser Ser Leu Pro Leu Pro
20 25 30

Pro Ala Pro Pro Gly Tyr Pro Val Ile Gly Asn Leu Leu Asp Leu Ala
35 40 45

Asn Asn Asp Val His Ile Arg Ala Arg His Trp Ser Arg Asn Phe Asp
50 55 60

Asp Asp Val Ile Ser Leu Lys Val Leu Gly Lys Thr Met Ile Ile Leu
65 70 75 80

Asn Ser Pro Thr Ala Val Ser Asp Leu Phe Asp Lys Arg Ala Ser Asn
85 90 95

Tyr Ser Asp Arg Pro Asp Met Pro Met Ile Val Asp Leu Met Gly Trp
100 105 110

Asp Trp Thr Phe Ala Leu Met Arg Tyr Gly Pro Arg Trp Lys Glu His
115 120 125

Arg Arg Val Phe Asn Asn His Phe Asn Ile Gly Thr Ser Gly Ala Ser
130 135 140

Glu Asp Arg His Ile Gln Leu Arg Ile Cys Arg Glu Leu Leu Ser Leu
145 150 155 160

Met Phe Gln Ser Pro Ser Lys Tyr Leu Glu Asn Leu Arg His Tyr Thr
165 170 175

Gly His Ile Ile Leu Lys Arg Thr Tyr Gly His Thr Val Val Asp Glu
180 185 190

Lys Asp Pro Tyr Ile Arg Leu Val Glu Ala Ala Ser Gln Ser Thr Ser
195 200 205

Glu Ala Ala Val Pro Gly Ala Phe Leu Val Asp Leu Phe Pro Ser Met
210 215 220

Lys Tyr Ile Pro Glu Trp Val Pro Gly Ala Gln Phe Lys Arg Lys Ala
225 230 235 240

Arg Glu Trp Arg Lys Leu Ser Glu Ala Met Ile Asn Ala Pro Tyr Asp
245 250 255

ES 2 747 844 T3

Met Ala Lys Gly Lys Phe Asp Glu Gly Asn Ala Glu Pro Cys Phe Val
 260 265 270

Ser Ala Cys Leu Glu Gln Asn Lys Thr Ala Ser Gly Gln Gly Leu Ser
 275 280 285

Glu Glu Leu Ile Lys Asp Thr Ala Ala Val Ala Tyr Ala Ala Gly Ala
 290 295 300

Asp Thr Ser Val Ser Thr Leu Thr Thr Phe Ile Leu Ala Met Thr Leu
 305 310 315 320

Tyr Pro Asp Val Gln Lys Ala Ala Gln Ala Glu Leu Asp Ala Leu Leu
 325 330 335

Gly Gly Glu Arg Leu Pro Asp Phe Gly Asp Lys Thr Arg Leu Pro Tyr
 340 345 350

Val Thr Ala Ile Leu Lys Glu Val Leu Arg Trp Ile Pro Val Leu Pro
 355 360 365

Met Ala Val Pro His Arg Ala Val Asn Ala Asp Thr Tyr Lys Gly Tyr
 370 375 380

Tyr Ile Pro Ala Gly Ala Phe Val Tyr Gly Asn Ala Trp Ala Ile Leu
 385 390 395 400

His Asn Pro Asp Ile Phe Ala Asp Pro Glu Thr Phe Arg Pro Asp Arg
 405 410 415

Phe Ile Glu Asn Pro Thr Leu Leu Asn Pro Ile Asp Asn Gly Val Phe
 420 425 430

Gly Phe Gly Arg Arg Ala Cys Ala Gly Arg Val Met Ala Leu Asp Thr
 435 440 445

Met Trp Ile Ala Met Ala Ser Ile Leu Ala Val Phe Asp Ile Ser Lys
 450 455 460

Ala Val Asp Glu Arg Gly Asn Glu Ile Ser Pro Pro Val Lys Leu Ser
 465 470 475 480

Pro Gly Thr Ile Ser His Pro Ala Pro Phe Pro Cys Ile Ile Lys Pro
 485 490 495

Arg Ser Lys Ala Ala Leu Glu Leu Ile Thr Gln Asp Asp
 500 505

<210> 5

<211> 1617

<212> ADN

<213> *Armillaria gallica*

<400> 5

ES 2 747 844 T3

atggccctct tctcagcata tgctctcgcg ttctcacttt tgatgggtccc cttgatcctc 60
 tacataactta ggatcgggag aaggagagat gggcttcccc cagggcogtc tacgctaccc 120
 ttggtaggta acctccacca gctcccacac ggtagtctgc atctacaatt cactgcctgg 180
 gcgaaggaat tcggagggat cttctcgctg aaatttggtc ccggaacggt cattgtcgcg 240
 acgagccctc gcgccgttcg tgagctaatt gatcagaaga gtgcaagcac ctcagatcgt 300
 ccaccgtcac acttttcgaa tgtcatcacc ggcgggaaca acattgggtt tgctcgctac 360
 tcggactact ggcgacgagg acgtcgtgtc atgcactcga tgctcacgaa gaaggcgtgc 420
 gtaaatactc tgactatcca acgcgccgaa gcttcgcagt tgatgtacga ctacctgtc 480
 gagcctaagg tgaacctctc ccatcaaatc ctagttgcaa acgtggggaa atctgacgtt 540
 caatgtcagg aattcgtcgc ccatggtcag cgttacgcca attcgggtgat cacttcaatc 600
 cttgctggta cccggtcgcc gcaccatacc agcccgtcgc tcaccgcttt cttccaaatg 660
 cagcatgaat ggactcacct cctcacgcct ggcgccacc cccctgtcga tatggtgccc 720
 ttcttgaaat acatcccagg aagttggaag caaatttgcg ccaaaatgaa ggtctcccag 780
 gaggagctat acggtgggat gatcgtatgcc tgcgcaaagc gcgtogaacg cggtatagc 840
 aatggatgct ttttggaagg tgagctagag aacaaggatg tggaccgtgg attgctacga 900
 gggatgtggt ccgcgctcat ggagggcgga tcggattcga cgtctatata tttgcagagc 960
 ttcaattctta tgcttgtagc gcaccagat gtccaagcaa aggctcgcgc agagatagac 1020
 tcggctcgtc gactggatcg attgcctgac attgaagata tggacaatct tccctatgtg 1080
 tcagccgtaa ttaaagagg actgcgtctt cgacctatca cagcgtggg agcaccacat 1140
 tatagtaccc agcctgaagt tgttgacggc tacgtgattc caaagaactc gatgatcttc 1200
 atgaatcaat ggggcatgct acacgaccca gactcctatg accaaccaga atcattcatg 1260
 cccgagcgtc tcctgacatc gccatttggc acaaggcctg gagcggatga tacgggcccg 1320
 agtaaagata tttacttcgg tggcgggagg cgcactctgtg ttggcatgca tttcggacag 1380
 aactcgttag ccattacgag catgtatctt atttgggcgt tcaacttgac caatgctatc 1440
 gaccacaga ccaagaacct aatccctgtg gatatcaatg aatatgacat cagcctcaat 1500
 acgataccga aacctttcaa gtgcgacttc caagtgagga gtgaagaaca caaacgcatg 1560
 gtcgagaatg cgtttgaga atcgcgccag gttttggcgc catttgagca ggattag 1617

<210> 6

<211> 1542

<212> ADN

<213> *Armillaria gallica*

<400> 6

ES 2 747 844 T3

atgacacacg catcgtctgc ttggtttctg gctgctgtgg ctatcgtcac cttcatcgtc 60
 gttcgtcgca tccgtagctc ttggcgcaag ctcccacctg gacctagagg ccttcctatc 120
 gttggcaatc ttctccagct acgtagcaaa caatggctca catttaccga gctgggaaag 180
 aaatatggcg accttatgta cttcaatgtc gccggacagc ctctcatagt catcaactcc 240
 ctgaggggtg ccacagatct cttagatcgt gcaaagtttt cggaccgccc acgcaacatc 300
 gtcgcctcag acataatgac tagaggcatg ttcgtcgcgt tcgctcctta tgggaatgcc 360
 tggcgccata tgcgtaaagc agctcatgaa ggtctcaaca aaaacatcgt gaaccaatac 420
 catccaatcc aatcaagga agctgtcctt cttgcaaacg atttgctggc ggaaccaaac 480
 cgatgggtat cgcacgttcg ccgtaccgcc gcttctacta tcatgtccat cgtatacgat 540
 aagcctgcaa cgtccgaaca agaccatca atcaagcgca tcaacgactt tgctgcacgg 600
 ttaaccctg ctgctatgcc tgggtgctcac tttgtggagt cttttccctg gatgctgcgt 660
 atccccagca aatacgcaaa gtggaaaagg gaggcagaag gctgggatgc caagattcg 720
 tccatgtttg aaagcctctt ccacagcgtc aaggatcgag tagccgaagg taataatcgt 780
 cccagcttcg tggctactct gattcaaggc gccggaagac atgggctgac tgaccatgaa 840
 agctcttggc tggcaggggc catgtacacg gctgggtgctg aatcgagctc agccgctata 900
 tcttgggtgga tgcttgctat gatcctttac cctgatgcgc agaaacgagc ccaagcagaa 960
 ctcgataagg ttgtcggctg agatcgactc cccgctttct cagactatga gcatttaccg 1020
 tatgttcgtg caatggttaa ggagactctc agatggcgtg cagtggatcc agttgggctc 1080
 cctcatcggg cgaccgagga cgatatctat aacggatact tcattcccgc cggtagtatc 1140
 ctcatcgcca atgtttggca tattaacaga gaccagaga attatggact tgacgccgaa 1200
 cacttcgatc cggcccggca ccttgacgag acaggccaac tcgctccggt cgcgggact 1260
 aaggaagaga accacgtttc cttcggtttt ggtcggcgta tctgcgttgg acgacatatt 1320
 gctaatagaca ctctgtttgc tgtcatagcc acgttgcttt gggctataca catcgaaccc 1380
 gccacggacg agaagggcgc ttcccttccg ttagatggtg acggatgtat tgaagatgga 1440
 gtggtcgttc gcccgttacc cttcaaagtc aaggtaactc caaggttttc ggaagcacag 1500
 gccattgtgg agcaagagcg ggagcttctt ggccactact aa 1542

<210> 7

<211> 1566

<212> ADN

<213> *Armillaria gallica*

<400> 7

ES 2 747 844 T3

atgactcggg tattctcaga ggaggtgtcg ccgatatgga ttttgactgc aatcgtagtc 60
 gtcgcctaca ccaccgttcg ctaccttcgc agcccttggc gtaacctccc cccgggacca 120
 agaggccttc ctctcatcgg aaaccttttg gaacttcgtg gcaagcaatg gctcacgttt 180
 actgagctag ggaaaaaata tggcgatctc atgtacttca atgtcgccgg gcagccgctg 240
 gtggtcctca actcccagaa ggtggcagcc gacctgcttg accgtcgcg c tggaaagtac 300
 tccgaccgcc cacgcaacat tgtcgcgtcc gacatcatga ctggaggaaa cttagtcgta 360
 ttcacccgct atggggatgt ctggcgccga atgcgcaagg cggctcacga aggtctcaac 420
 aagggtgtcg tatacaaata tcatcccatt cagactgcgg aagccgttct tcttacggct 480
 ggcgtgctcg cggagcccga aaaatggaat tcacatcttc gtcgcaccgc cgcacatgcc 540
 atcatgtcca tggatatacga tacgcctcca acttcagaac aagatccatc agtcaagaac 600
 attaatgaat ttgtggcacg attaactcgc gctgctatgc caggcgccca cttcgtggaa 660
 ttcttcctt ggatgcggtg catccctagc aagtacgcaa agtggaagcg tgaggcagaa 720
 gaatcatacg ctaaagattc tgcgatgttt gaagggctct tcaacggtgt caaggatcgc 780
 gtagccaagg gtgacgagcg tcccagtttg gcgagtactc tgattcagga cgcgggaaga 840
 cacgatttaa ctgacagaga aaattcgtgg ctggccggga ccatgtatgc tgcgggtgct 900
 gagaccactt caggtgttat gtcctggtgg actcttgcaa tgattgtcta ccccgaaca 960
 cagaaacgag cacaagcaga actcgatgct gtcgttggtc gcgatcgttt gccgtctttc 1020
 gccgactatg aacatttgcc atacattcgt gcaatggtca aggaggccct caggtggcgc 1080
 atggttgacc cagtcgggct tcctcataca tcgacagaag acgatgtata tgacggatac 1140
 ttcatacctg ccggtacat catcatcgcc aacgtctggc acctcaacag agaccagag 1200
 atctatggc ccgatgccga acatttcaac cctgcccggc acctcgacaa ggacggcaag 1260
 ctggcccctg gacccgcaga cacgaaggag gagagtcaag tcaacctcgg atttgggcga 1320
 cggattttgtg tgggacggca cgttgcaaac aactctttgt tcattgacat agcaatgatg 1380
 ctttgggcta tgaacatcga acgtgctacc gacgagaatg gtgttcccct cccgttgat 1440
 gttgacggat gcgttgaaga tggactagtc actcgtccag tccctttcaa ggccaagata 1500
 actccaaggt tccgggaagc tcaggcaata gtagaacaag agcgagaact ccttgatata 1560
 cactga 1566
 <210> 8
 <211> 1530
 <212> ADN
 <213> *Armillaria gallica*
 <400> 8

ES 2 747 844 T3

atggattctg cgtctcttgc cgttgcgtt tgggctatcc ttctcgttt gtggctcaga 60
 aggatthttcg gtcaacgctc ttccctccct ctccccctg cacctcctgg ttatcccgtt 120
 ataggcaacc tgcttgactt ggctaacaat gacgtgcata tcagagccag gcaactgtcc 180
 cgcaatthttg atgacgatgt catctctctc aaagtctctgg gaaagacat gataatcttg 240
 aactcgccca ccgcagctc tgatctcttt gacaagcgag cctctaacta ctcagatcgg 300
 cctgacatgc caatgattgt tgacttgatg ggatgggact ggacatttgc cctcatgcgg 360
 tacgggcccgc gttggaagga acacagacgg gtcttcaaca accacttcaa tattggcact 420
 tctggtgcgt cagaagaccg gcatattcag ctacgcatct gtcgcgagct cttgtccttg 480
 atgttccaat ctctctcga ataccttgag aaccttcggc actacacagg tcatattatt 540
 ctcaaacgca cgtatggaca cacggtagtc gatgagaagg acccttatat ccgcctggtt 600
 gaggcagcgt cccaaagtac atctgaagcc gcagtcaccag gaggcttctt tgtcgatcta 660
 ttcccacgca tgaagtatat tcctgagtgg gtcccggggc cgcaattcaa gagaaaggct 720
 agggagtggc gcaaactctc agaagctatg atcaacgcac cgtatgacat ggccaagggg 780
 aaatthgatg aaggaaatgc cgaaccttgc tttgtctctg cttgtctcga gcagaacaag 840
 accgcttccg gtcaagggtt gagcgaagaa ctcatthaaag acaccgctgc cgtcgcataat 900
 gctgccggtg ctgatacgag cgtctccact ctaacgacat ttattctcgc catgacactt 960
 tatcccgatg ttcagaaagc agcacaagcc gaactcgatg cactacttgg cggcgaacgt 1020
 ttacccgatt ttggcgataa aactcggctg ccatacgtca ctgcaatatt aaaggaggtc 1080
 cttcgggtgga ttctctgttt accgatggcc gtacctcacc gagccgthaa tggggacact 1140
 taaaaaggat actacatccc ggctgggtgcc ttctgttacg gaaacgcatg ggctatcctt 1200
 cataaccccc acatctthtgc agaccccgag acattcagac cagacagatt catcgagaac 1260
 ccaacgthac tcaaccgat agacaatggg gtatthgggt ttggaagacg cgcgtgtgca 1320
 gggagggtta tggcgcttga caccatgtgg atagccatgg catccattct agccgtcttc 1380
 gacatthtca aagccgtcga tgaacgcggg aatgagatct cacctcccgt aaagctcagt 1440
 ccgggaacga taagtcaccc ggcaccatth ccgtgtatta tcaaacctcg atctaaggct 1500
 gcccttgaac tcattacgca ggatgactga 1530

<210> 9

<211> 22204

<212> ADN

<213> *Armillaria gallica*

<220>

<221> misc_feature

<222> (168)..(168)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (215)..(215)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (1520)..(1520)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (1551)..(1551)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (1554)..(1554)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (8100)..(8100)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (8105)..(8106)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (8980)..(8980)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (8982)..(8982)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (9701)..(9701)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (9703)..(9703)

<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (10603)..(10603)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (11508)..(11508)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (13489)..(13489)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (14372)..(14372)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (14396)..(14396)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (14431)..(14431)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (14474)..(14474)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (14484)..(14484)
<223> n es a, c, g, o t
<220>
<221> misc_feature
<222> (18683)..(18683)
<223> n es a, c, g, o t

ES 2 747 844 T3

<220>

<221> misc_feature

<222> (19415)..(19416)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (19480)..(19480)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (19619)..(19619)

<223> n es a, c, g, o t

<220>

<221> misc_feature

<222> (19723)..(19723)

<223> n es a, c, g, o t

<400> 9

```
atggccctct tctcagcata tgctctcgcg ttctcacttt tgatgggtccc cttgatcctc      60
tacatactta ggatcgggag aagggagagt gggcttcccc cagggccgtc tacgctaccc      120
ttggtaggta acctccacca gctcccacac ggtagtctgc atctacantg agtgggtggca      180
agatgtccca attctcccct gaacatgtcc tgatngattc actgcctggg cgaaggaatt      240
cggagggatc ttctcggtaa gaactatgaa actcgtcgtgta aattgtactc atattctaaa      300
```

ES 2 747 844 T3

tctacgccag ctgaaatttg gtcccggaac ggtcattgtc gcgacgagcc ctccgcccgt 360
 tcgtgagcta attgatcaga agagtgcaag cacctcagat cgtccaccgt cacacttttc 420
 gaatgtcatc accggcggga acaacattgg gtttgctcgc tactgtgcmc tcttctctat 480
 ttcgtcattt tctattgtcg gctaaatctc ggttcacagc ggactactgg cgacgaggac 540
 gtcgtgtcat gcactcgatg ctcacgaaga aggcgtgcmg aaatcatctg actatccaac 600
 gcgccgaagc ttcgcagttg atgtacgact accttgtcga gcctaaggty aacctctccc 660
 atcaaactct agttgcaaac gtggggaaat ctgacgttca atgtcaggaa ttcgtcgcgc 720
 atggtcagcmg ttacgccaat tcggtgatca cttcaatcct tgcctggtacc cggctcgcgc 780
 accataccag cccgctcgtc accgctttct tccaaatgca gcatgaatgg actcacctcc 840
 tcacgcctgg cgcaccacc cctgtcgata tgggtccctt cttgaaatac atcccaggaa 900
 gttggaagca aatttgcmcc aaaatgaagg tctcccagga ggagctatac ggtgggatga 960
 tcgatgcctg cgcmaaagcmg gtcgaaccmg gtatacgcaa tggatgcttt ttggaaggty 1020
 agctagagaa caaggatgty gaccgtggat tgcctacgta tgtcttaaca gtacgcgtyc 1080
 cccattttac aaatttatat tttacagagg gatgtgttcc gcgctcatgg agggcggatc 1140
 ggattcagcmg tctatatatt tgcagagctt cattcttatg cttgtagcmg acccagatgt 1200
 ccaagcaaag gctcgcgcmg agatagactc ggtcgtcgmga ctggatcmg tgcctgacat 1260
 tgaagatatg gacaatcttc cctatgtgty agccgtaatt aaagaggtaa gctcttcgct 1320
 tcgcctttgt tctatgccag tgtgttcatt acgtatatca taggtactgcm gtcttcgacc 1380
 tatcacagcmg ctgggagcmg cacattatag taccagcct gaagttgtyc gtcattctat 1440
 ttttaccttc cgcacaattca caaagtgatg tttgctcact tcaggttgac ggctacgtga 1500
 ttccaaagaa ctcmgatgn ttcatgaatc aatgtacgta cggtaacttc nctntaccgt 1560
 tgcgaactaa agctcccatc ttattgcagm gggcatgcta cacgaccag actcctatga 1620
 ccaaccagaa tcattcatgcm cgcagcmgtt cctgacatcm ccatthggca caaggcctgm 1680
 agcggatgat acgggcccgm gtaaagatat ttacttcggt ggcgggaggg taagcagcm 1740
 gcttgthtat gtacatcggm gtcgtctaac tttttctaag cgcactctgty ttggcatgca 1800
 tttcggacag aactcgmctag taggtthaat ttaacgthc gatatgccat tacgcmct 1860
 cactgggacm gaaataggcc attacgagca tgtatcttat ttgggcmthc aacthgacca 1920
 atgctatcmga cccacagacc aagaacccaa tccctgtgga tatcaatgaa tatgacatcm 1980
 tgcgtgthgt ctcttatgcm agthttaccg gtaccthgac aaaagththc agagcctcaa 2040
 tacgataccm gaaactthca agthcgmactt ccaagthgagm agthgaagaa acaaacgmct 2100
 ggtcmgagaa gcmththcmg aatcmgcmca ggtththggcm ccatthgagcm aggattagag 2160

ES 2 747 844 T3

gaacatggct tgaaggaaaa tttggttagt gtctacgctg cgacgctatt cccattaaac 2220
 ttagtgagtt taccagctag aataacagta tacagtatac cacatatcac cttagaatth 2280
 tgtattacaa ttgctccatc tgcacgggta tgtacgagag aagggttatc agaataatthc 2340
 ttcatattgc agtggaccac ctgctgact caactttgaa gacacattat cagctcggta 2400
 tactaagcgg gggccgaagc aaatcctatg cctaccttat ttttgccatc ggccggactc 2460
 tggaaaataa cgttagaatt tcgaactcaa caggccgatt tgtagtggta gccgaagatg 2520
 tctgcgttcc tgttcccatt gaaccagga aggatcgatc gaaatgtcct tgaaacactc 2580
 ggtcggtgat gcgcgggcac ccgaacctga agagcacaac ttctcgctac ccgtttggcc 2640
 atcggggagc ccacgcaagt ttgaagttgc agattaccg taatagaccg tagcccgaca 2700
 agcgcgcccc gaagtgatca agtggaagt actgattgcg aactagcag ggtcatttgt 2760
 ggagctgaaa ccgaggtgga tttctgaaaa gaaaaagaaa accctagatt accattgcta 2820
 acggaagttt cagtcttgca tctaaaaaca agggaacctg ttccattggt ggtggggcca 2880
 tatttttggga aagcgcaaca ctgggcgttt ggcattgcbc gtggatagtc aaacttcaac 2940
 acagtaagca tgcggctgtg taagttacgt cgtgacgcat atgcbacctt gctattcgat 3000
 tgtatcaatt ctattactat aaaccgttgc ttgcagtgat agcgggaagca aatccgagcc 3060
 ttccgggtgat ccgtatcagg tattgccaat ctcgatgatt tggacctatt gtaagcgcac 3120
 ccggcaataa gtaccctcgc acagatgaaa attcaataaa gttacctca accgggtact 3180
 ccagcctcca ggtcttggct catgtccggt gcctgtggag aggctatgbc ccgatttcta 3240
 catagcacat acatacgcac cagccaatta tgattcaata tggcttaggc agttcgttcc 3300
 cacgtcgata tatcaggcaa cgcttcaaca cgataataat aataaaagcc ccgcbagcgc 3360
 tctcgcctga ggcttgttgc tgtgtggcat gtgccatac tctattatgt cactttaaac 3420
 tttcaacctg attatatcag ggcttgtggc catggggatg ctcbgtgctgc acggcaaaaa 3480
 tcagcgggag cctggcttga tcctaatttc tgcctcttat gaacatttga agtttgaacc 3540
 ctcbgcccc aacgcaaacg gaaaatttgc catcaagtac aaagttgtht gcatctgaaa 3600
 tgtactaata catgggcctt caagcaatth tggcactthg agaagcctga aacggtaaga 3660
 gcacggtatg cttattthtt ttttacttcc ttcbgtgctth ctcbgagccgc tgcgcaatcc 3720
 tgccaggata cacaagtaag tgtgtccaac acaaacgaca tcbctccgaa aatagtagca 3780
 ataatgthac gggtaacagag tctccgcctt atagcbtcta gagcaagtga gtcgtgacac 3840
 agtattatat gcattctgct ctgtctaac tcccagacta acacctthga acaatgatca 3900
 aagagaccga ctggagcgtg aatacgcact tcaagatact tcaactatth ccaacagatc 3960
 atcaaagtta tgcgatgggt ccattgattt tacaagctth aaaaaaggaa ggtggcagaa 4020
 gaagtgcctgc tgtgcctggt gtcgagccat ctcbgtactct ctcbatggacc ggtaaagcaa 4080

ES 2 747 844 T3

cgtcatgtat atctgaccat ccctaaacct acccggtgcc aaagcaaaca attatgactg 4140
 tgataaccac tgacaattgc tgcaaagcca tttcatagcc atggtactaa tacttgcttg 4200
 tgcttaacaa tactgagcat ggtgcagtac tttcaatgcc tggttttctt gttcaaagtg 4260
 actggaatct ccccgagacc ctagcaaagt cgacgcgata gaaagaaaaa gaagactgac 4320
 cagaggaggg ttccccggtc gaacggccga ctgataaaat atacatacaa aatttaggca 4380
 aactaaccgc taaaccaagc aagctcaaca catcaccagc acatgctccc tacggtcgca 4440
 tccacttgaa aacttgaaaa cttgaattct ccacagacat aacacgtaag cccaggtctt 4500
 tgggctttga ggaaagatgt tataatagag cacagtgatg gtcatatata aagacaccga 4560
 gttgatatca gctatcgcca ggctgatcat gtcggcaggt ccttacaaaa agagtcaaga 4620
 ccgccagcag ccagtcatcc cctcatttga ccctagttgt cgcccgtaac acttggctta 4680
 taggaggtcg cagcccgcca cttcataacc ctagattctc tgatcttcaa tgccttgctt 4740
 tgagacgtag ttgcgacaac cctcatcgac ggcattgctat cgggtgaattg gtctagcaag 4800
 aaagtggctt ttatcatacg tatttctccc gagttgctca gcgtactgga gaactgtaa 4860
 ttgaagattt ccaggctgtc gtcacagcgc cccagtatca tccagcgtgg gtcgcttctt 4920
 cagccccacc tcttgacgta tactttcatt acatatgaag aactcctgcg tggaaagcaa 4980
 gggctctgaa ttggctgata ctttacttcc ttacatgaaa gtatatcaca cactgagaag 5040
 ataaaatttt actgaatcat ttccctttag tagtggccaa gaagctcccg ctcttgctcc 5100
 acaatggcct gtgcttccga gaaccttga gttaccttga ctttgaaggg taacgggcca 5160
 ctgcaaggca tacgatcaat aaatgtcgta cgtaatggcg aacgagaact acgcacacga 5220
 ccactccatc ttcaatacat ccgtcaacat ctaacggaag ggaagcgccc ttctcgtccg 5280
 tggcgggttc gatgtgtata gcccaaagca acgtggctat gacagcaaac agagtgtcat 5340
 tagcaatatg tcgtccaacg cagatacgcc gaccaaaacc gaaggaaacg tggcttcgca 5400
 tagattagcc tatgtcgtat tgaaaaggca catgaacgaa cttctcttcc ttagtgcggc 5460
 cgaccggagc gagttggcct gtctcgtcaa ggtgccgggc cggatcgaag tgttcggcgt 5520
 caagtccata attctctggg tctctgttaa tatgccaaac attggcgtat aggatactac 5580
 cggcctgtat atgtcagatg agtatgacag cagcgaactg aacgtcgttg acgcttacgg 5640
 gaatgaagta tccgttatag atatcgtcct cggtcgaccg atgagggagc cctggaaaat 5700
 cactgtcagt gtcgtgccac acgcatctcg agagtcatac gcaccaactg gatcctgcac 5760
 ggaagacaac aattcgcata aaacggagga gtcgagagag aaaacttacc actgcacgcc 5820
 atctgagagt ctcccttaacc attgcacgaa catacggtaa atgctcatag tctgagaaa 5880
 cggggagtcg atctcgaccg acaaccttat cgagttctgc ttgggctcgt ttctcgcgat 5940

ES 2 747 844 T3

cagggtaaag gatcatagca agcatccacc aagatatagc ggctgagctc gattcggcac 6000
 cagccgtgct accgaataat cagatgggag gaaactaaac gcagtctaaa agacgtacta 6060
 catggcccct gccagccaag agctttcatg gtcagtcagc ccatgtcttc cggcgccttg 6120
 aatcagagta gccacgaagc tgggacgatt attaccttcg gccttatata gaaaagatta 6180
 gcactagctg tcgatcgcac catcggtaac gtactactcg atccttgacg ctgtggaaga 6240
 ggctttcaaa catggacgaa tctttggcat accagccttc tgctccctt ttccactttg 6300
 cgtatctacc aatcagcca tcagccgctt tcccgtgacg ttttgtataa ccttacttgc 6360
 tggggatacg cagcatccag ggaaaagact ccacctttgg gttgcaggca tcaatcagac 6420
 gctgggcttg acagatgaga gtctaacca caaagtgagc accaggcata gcagcacggg 6480
 ttaaccgtgc agcaaagtcg ttgatgcgct tgattgatgg gtcttgctcg gacgttgcag 6540
 gcttatcgtc tacgatggac atgatagtag aagcggcggc acggcgaacg tgcgataccc 6600
 atcggttggg ttccgccagc aaatcgtttg caagaaggac agcttccttg atttggattg 6660
 gatggtattg gttcacgatg tttttggtga gaccttcagc agctgcttta cgcatatggc 6720
 gccagctaag ctgcgatatc aggtacgcat actggatcag tgagatcaaa agcttacgca 6780
 ttcccataag gagcgaacgc gacgaacatg cctctagtca ttatgtctga ggcgacgatg 6840
 ttgocgtgggc ggtccgaaaa ctttgcacga tctaagagat ctgtggcaac cctcagggag 6900
 ttgatgacta tgagaggctg tccggcgaca ttgaagtaca taaggctgcc tagagcatag 6960
 cgtgaaacca gaaatcatgc ccaaacgaag agcgcgcata ccatatttct ttcccagctc 7020
 ggtaaatgtg agccattgtt tgctacgtag ctggagaaga ttgccaacga taggaaggcc 7080
 tctaggtcca ggtgggagct tgcgccaaga gctacggatg cgacgaacga cgatgaaggt 7140
 gacgatagcc acagcagcca gaaaccaagc agacgatgag tgtgtcatca tggctcggaa 7200
 aagaagagat agcgttggtg gttcttcttg atgtgaaaga ctaactcatc acgcgcacg 7260
 aatccctagg gtgagcgtca cattttcttc ccgccatcag gtttcacggg aagggttctg 7320
 ccgctaacaa aggatatggt taagctcagg agtgtatgtc tggctccaga ctttgattag 7380
 ttgatatcat agtgctccta gaacagtctg ctggctcacc ttcgaaccca tttgatccg 7440
 agcgcgtgaaa atagttgaca gttgacagat gaaaaataag tatcataaaa agtaccaaga 7500
 cccccactta caaagttaga gggtgaaaat cctccagaaa caatatctcg ctgagctggt 7560
 ggcgacgtcg tggttgtgag ctcgatatca aattgactga tgtcatgcat ggacgtaatt 7620
 tcaatatctg cagccacggt gtgcgattgt acccgcaacg gcaaccgtcc tactggccta 7680
 tggctagcac cgatgagtag aatgtagtat actagagctt tttgggagcg atggcatttg 7740
 gtaaaratgg aaattatttg gaactttgta cctactctcc aaggaagcac ctcgtttcac 7800
 atcttccaag aaactgtcct tgcttgggat actttaaagt aagggtttat aattcattct 7860

ES 2 747 844 T3

aaatccgatg acgcctcacc aacgggtccac aggactttca attggctcat atcatcgttt 7920
 ctcgacgaaa aacgtttccg aggaatattc gaatttcagt ctacagtaga cgatgtcatt 7980
 tcctatagtg tcgtagtcgt agattttgta ttacttagag gacaacagag tggaagggtc 8040
 acgtactttt tccacttccg tgacgtattc cctgtaatac gcgtccttct cattatcagn 8100
 aaggngagg cacttagctt taacgacaat atccacagat tctacaccag caatatgagc 8160
 ttcacctaac tccaacttcg tcttggcccc gagatggcaa aagtgagtgg ttacagctgt 8220
 cgatgacggc cagacctatt cctcaattag tcctgtgaca gtattatcga tcgatgggaa 8280
 gtgcaactct ggtcagaagt gcaactttcg ccaatgtag cctcctacag gccgcattta 8340
 cgacttgttc ctctgcaa tttaaagtag agtacctata gatgatacct acaagtatgt 8400
 taagtcagat cggttttccg ttgagataat tataatcagg aatttctaca tcgtgcccaa 8460
 tcgaaaaaaaa tgaatagcac atcgcgttct tgcttcgtat cgatacagtg aaatgccggg 8520
 tggttctgag agatttatac aagcttggtg tatattttta gagtgttttg acaaaaacga 8580
 atgcgagcag acctgaagta aagtttgac atgcagtcga agctcagcat ggcaagctca 8640
 gccgctccgc gcttgagccg atcgcggctt ggtcgcgatt gcggtgtacc ggcagcatcg 8700
 gcacttcggg tcatatttcg tcgcagttgc acttaccgcg tattttgacc caccacgacg 8760
 gcatcactat tatgggttcg aatctcagac ggcgactcag catatttcgg caaggtctgg 8820
 ctcagtgccg cagtcatcca agccgcattg ctgatgctgc aatcgagccg cagtgcgaca 8880
 tatttgaaag atggatgcga gaccaagtat gtacgttgca aatagtagtg cgaaataagt 8940
 aggtgaagtt ttaagcactg gtccacctgt gttgacttcn gnaagttgcc tctaagccga 9000
 gtctcaccgg tcggcaagcg agtcggttaa aaatcctaata taggtacagt aagatttgat 9060
 cgacgcaaaa ttttcgatca tctcttcccc ccttctgcaa ttcgtttctc cctcatttct 9120
 gagtcctttg catcggaaat agtatcgta ggatgctgcg taagacaaga tggccgagtg 9180
 gctgtggcta aggcaccagg ctaaggaaac aactgttatac aaaataactg cccccgaaa 9240
 ccaaactgat attcgagtgc cagtgaagag ttatctagtg tcgtgacttc ggctctgcaa 9300
 ggaagcgaca ggaactttcc gtaatccagg cgagtacgat taccgttttt ttgagggttag 9360
 agataccgtg ttgtaacata tgcatttggt gttcaggact tcgttttttg cattcaactg 9420
 tatggatcga tatccagcga attacataac cacgtcgtcc ggtatcgata ttctgtgcta 9480
 tgtcatcata tcccgtacga ccttttgact ctccgggtgg tatgatatcg tgttttttcc 9540
 gtcgcttaat ccaacgattc cttacaaggg gagccgtag aacctatgca ttacaaatgc 9600
 gaagcagtaa ttcggccatt tctggcctca ggaagcccgt ttcctaggca tatctatgcc 9660
 ccaaaccgga ttagaggtga gctggctgct aacgatctgt ncncatgacat ctagtggcac 9720

ES 2 747 844 T3

tcaagagccg gttgatttcc attcatagct tagtattcaa caaggggccg ggatgccctt 9780
atgtccagca ttgaatgggt tactgcatca ctactcacac gtaacaagga cagatgggtc 9840
tgctgttcaa tcctgcttgc catgatagcc actccccaga gaatttgttt gtcacgaaca 9900
gcaaacacgg gcaagctaata gtgattcttt cgattgcaca tcaactcgtt cttcgggtcaa 9960
atgtagcagt gcatatcacc tctttctccg ggaatgcagg tgttacgttc ggtctatatac 10020
tttcaactggc atcaacttgg cttggaatgt gcattaatct agaaactccg ggggtgccatt 10080
caccctacgg tatgctctgc cagtgactat gtccatattc atcctctcaa aagaccagc 10140
gaagaacctt cttacctgog cctcgggtcta cttcatgaac cactcacaaa gatccaagaa 10200
ctcgataggt accgtcggaa gtgggcttac tagcatcctt cgtatcttca tcccgtacat 10260
cccgcaaacg cactgtctct gtccagccaa tctgacagct gaatttcatac agtgtgtcat 10320
ttcagacaat gtcactccat gcggccctat tcttgctcca cttctcccgt tgatgagggt 10380
gactgcgaac tcaccgcaca gtttgatctt gctctatcat ccccatgagc ttgggaacgc 10440
acgtaatctt agatgagaag gatacccacc agcttgcatg ggggctgcgt atcttacttg 10500
atcgcgaaaag aaggggtccaa gtcatttggga agctagtagc acaaggcatc ctcacatgat 10560
atcctcggtg ccgacataga aagtgggcca gtgaaggttt ggncgaagat tgttgattgg 10620
ttggatatcg aaccttggc actcctcggg tategtaacc tcgtatgtgt tgttaatacac 10680
tgccgagcaa atagcttctt cgaagccatt acgttctttg ctcagcgttt ggtccatcga 10740
gtatttagag cggggtccat caagctgttc tgcgcgctg gtatgactgc tatggtgggc 10800
tgttgacatc tcgatggatt ctggaatctc attgaaaaa aagactatgc agcgcgtggt 10860
gaatggttcg gcgtaggtat ttgctaggta ttacagcaa caaaaatatt acgcttgatg 10920
acgcgtttgt ggggtggttta tggctattta ggccgcatc tcggaggcca tttcgtgcgc 10980
ttgaagtccc caaaacaagg tttatgacga tccgagcttc cttttcatcg tccacgaaac 11040
cgcccgtta cggatgcaa accgctgtaa agaaattcat tacgactagt ctactatcgt 11100
aacaggtccc ccttaccacc gtcagctaaa agggagagag tagcatatat ttattctcta 11160
agaatagatc ttgatacata agctggtcac tggtcagacc agactcggag gttcaacaaa 11220
aacggatgac acacgtccac gtggctgccg aatcttcttc cgactcgggt atccatgcat 11280
tcatctccgt cgcgtagtca gcgtcaaaaa tggcgcatac aaagagagtg cagatatgaa 11340
gaccagaaat ctgtgggogg cgcagctcag agctggatcg tgggcacaaa tttggacggt 11400
ttctgcttca cttgacctca tctttctctc taagatttat tcattgattt atttcgacgt 11460
tgcaaacgca cacatgagag ttgtcaaaact gcatagcctt catggaangg cgttctcaat 11520
aaacaatact cgggtgcagag ttatctacag aagcaaaagc taggatatgt gcctaggtaa 11580
gccatthtca atttgatctt ggtacatcag cgcacagct caagctcact tacttgccgc 11640

ES 2 747 844 T3

agattcacat caatgttccg ctcaggtagc agagtataac catataagcg attatacggg 11700
 atggggactc ctgtacgtag gcaatcgaga ccagtgacat tattccgatt aaatgattct 11760
 acagcggagt acactctaca tagtgaagtg agaaagaaac ctgatagtcc tcagtgatat 11820
 ccaaggagtt ctcgctcttg ttctactatt gcctgagctt cccggaacct tggagttatc 11880
 ttggccttga aagggactgg acgactgcaa cgaagacgat taacaacgac ttgaacacgg 11940
 aacagtgcag gacttacgtg actagtccat cttcaacgca tccgtcaaca tccaacggga 12000
 ggggaacacc attctcgtcg gtagcacgtt cgatgttcat agcccaaagc atcattgcta 12060
 tgtcaatgaa caaagagttg tttgcaacgt gccgtcccac acaaatccgt cgcccaaadc 12120
 cgtaggtgac gtgactgcac aaattcagct tattatgtgt gggaactgga tatcaacgaa 12180
 cctctcctcc ttcgtgtctg cgggtccagg ggcagcttg ccgtccttgt cgaggtgccg 12240
 ggcagggttg aaatgttcg catcgggccc atagatctct gggctctctgt tgaggtgcca 12300
 gacgttggcg atgatgatgg taccggcctg tggtagatga gcacgtactg catcgagcgt 12360
 aatccacact tacaggtatg aagtatccgt catatacatc gtcttctgtc gatgtatgag 12420
 gaagccctag agaccgacta tcagaatcag gcataatctgc ctcgagagtc atacgtaccg 12480
 actgggtcct gtaacggcag cgtcaattgg tctgcgttgt tcgacaccac ggagagagcg 12540
 gacttacaac catgcgccac ctgagggcct ccttgaccat tgcacgaatg tatggcaaat 12600
 gttcatagtc ggcgaaagac ggcaaacgat cgcgaccaac gacagcatcg agttctgctt 12660
 gtgctcgttt ctgtgtttcg gggtagacia tcattgcaag agtccaccag gacataaac 12720
 ctgaagtggc ctcagcaccg gcagcactgt tgaagctta gaggaggaaa tttagaagac 12780
 gaagaagacg tactacatgg tcccggccag ccacgaattt tctctgtcag ttaaactcgtg 12840
 tcttcccgcg tctgaatca gagtactcgc caaactggga cgctcgtcac ccttggcctt 12900
 ttcgaaacgc atgggtatca actatcaagg attcaatggc tgtgcacgta ctacgcgatc 12960
 cttgacaccg ttgaagagcc cttcaaacat cgcagaatct ttagcgtatg attcttctgc 13020
 ctcacgcttc cactttgcgt acctaagtag gaaattgtca atgcacgctt acgcaggacc 13080
 tcatcaaact cacttgctag ggatgtaccg catccaaggg aagaattcca cctttggagc 13140
 cacagacatg attaacggga aggaaaaccg gtagagagaa acttacgaag tgggcgcctg 13200
 gcatagcagc gcgagttaat cgtgccacaa attcattaat gttcttgact gatggatctt 13260
 gttctgaagt tggaggcgta tcgtatacca tggacatgat ggcagatgcg gcggtgcgac 13320
 gaagatgtga attccattht tcgggctccg cgagcacgcc agccgtaaga agaacggctt 13380
 ccgcagctcg aatgggatga tatttgata cgacaccctt gttgagacct tcgtgagccg 13440
 ccttgccgat tcggcgccag ctaagthtca aatgtcagat ggaggaacng aatacgaagt 13500

ES 2 747 844 T3

ggaaggacct tacacatccc catagcgggt gaatacgact aagtttcctc cagtcatgat 13560
 gtcggacgcg acaatgttgc gtgggcggtc ggagtacttt ccagcgcgac ggtcaagcag 13620
 gtcggctgcc accttctggg agttgaggac caccagcggc tgcccggcga cattgaagta 13680
 catgagatcg cctagagcca agtcagctaa tctacatctt acgcagctcg aaagaacaga 13740
 ccatattttt tccctagctc agtaaactg agccattgct tgccacgaag ttccaaaagg 13800
 tttccgatga gaggaaggcc tcttggtccc ggggggaggt tacgccaagg gctgccaagg 13860
 tagcgaacggt tggtgtaggc gacgactacg attgcagtca aaatccatat cggcgacacc 13920
 tcctctgaga atatccgagt catgacggct tatttccgac aagaaataat cttgaaggat 13980
 ttgagagaaa gggggggggg gtgataaact cgtcctgtgg atgccgtgtc tctcttgat 14040
 ggtttatfff ctcacacctt agatccactc tattaggtat tacaccgctc atggatctgg 14100
 aagctcggcg tcagtgttta ggagcggagt aatgattcgc ccggtttggt tcaatcactg 14160
 atcagaacaa catctgagtt gatcggggtc tgatctgatt cgttgcaata ccctggttcg 14220
 atatccccgg gccgctacag cttcaatatt atgtatttcg gatatacaata caatggtgac 14280
 agcttgagta cacatctgct ccgagcaccg agttcgagaa aagataccaa cttccccgta 14340
 acggtagata cgcgcgtagt aacggcatag gnagacgata ggtacttggt attgcnttac 14400
 ggtcctacat ttgccatctc atcggctcct nagccacacg gccactatcg gagtcgcaaa 14460
 agcatgtagc accnttatff ggtntgcgcg attttcggca tacgatggat ttcgctgagt 14520
 caaggtcaca gtcggtggag gcattcggag accctggttc gcttggttta caaaatggga 14580
 ggacacaaggt tattgtgtgg agccacattc ataatgtaca atgttaatac tgtcttgccc 14640
 ttatcgacia tcaatcttct cattcaactt cggttaacttt tgactcttct cgggcggcca 14700
 tggcggcaat cccacgattc acggattcta gacgtaactt actcaatgcc tctagctgat 14760
 aaatacgtag ctgcagttgc atagctccgg tgtaggcacg ctgaatttta agttgaacgt 14820
 gctttgaggt gtttgaaagc cactggaaac attatgttcg atctgatagc agaagttcgt 14880
 aagccagctt gagactcggg gacagttggc gcactgagcg ctctgatgta acaaagccgc 14940
 agactccacc acccatgagt aagatcgtga tttcagcaat gctagcagta cagactctac 15000
 gcggatattt gtcactagta acaaaatgac tggtaacaa aggcgtcaca ggacagagcc 15060
 tttcaciaaac agcaagagta aggacatatc aagacgacia actaatagta ggatgaatcc 15120
 acacaactct cagtcgaaca caatcccgcc atcgacattg atctgattgt cccatttaag 15180
 tcacaatagc aaccataatc acaaaagagc ttactgtttg tccctggtaaa agaacaatgt 15240
 caaaatatag ctgatgacga gaaggaaatg cataccggtg atgagacttg cctccttgga 15300
 tgcaaggaag gaaaccacia ctgcgatatc ctcgggcagc gcgatgagc ccaaggccgc 15360
 tgttttatta agcttcgaaa gtgtaagata tgacgtgatg aaagtgaat tgacgcacct 15420

ES 2 747 844 T3

cttgctcaag aactcccttg ggtgtgccta taatggtttc gaagtgatca tctggaaata 15480
 aggagaatgt cggataaaag tagaatagaa caatgctctc gttcatacac ataggtgtct 15540
 tcaactggccc tggggcaaaa gcattgacgg agattccatg aggaccgaac tccaaagctg 15600
 ataaaaaaag tattcatgag cgggaacctt tcatatgtat tgcacggagc tccccgcaa 15660
 cttgggtcag tcctcgtata ccgaatttgc tggctgcata tgcactgaac agcggttgac 15720
 ctgaagatct atcaagatag atacagatta tccgcaacga ctgctcacct ttcttcccg 15780
 ttcccgaaca cgcgcctaga atatcattga cggggaagct ctgaagacgt cgacatacag 15840
 cttaccgatg atcctcccac cttttccttg agtgaccatt tgtttgccgg cgtgcttgta 15900
 gcataagaaa gtaccacgca ggtttacaga ggtaatgcgg tcaaagtcct cggcagtagc 15960
 taggtgacga tgagaaatct ggaagaaag gagctatgaa atcgtactat cgaaaaaaga 16020
 cttcataatg gatatcccag cgttgcagga ccatcttga gaaatgagat ctcgtttgcg 16080
 acaagcaaca gccgaacgta ccacgtcgag tccgcctaata tcttcaacaa cttttgaaat 16140
 catctcttca acctcggttt cctgcgagac gtccgcaaaa atgtagcagc ccttgcgacc 16200
 aacggcctct atagcggcga gcgcagaaac gaggttcgct tcgtttgctt gaagatcatt 16260
 gagcgcgacg tcaaatacgt ccgatgccag tcgaagtgcg atggcacggc caatgccctg 16320
 tccagcacc cgtgacaagc cgactctgcc ctgagaaact gaagtggacg ccattgtgtc 16380
 gaggagaagt ataagggagt aatggcccgt ggccaccttt atatatgatg gtcattagat 16440
 aaggtcagca atcagtcgtg tcgagttagc ggtggttcgg ttcaagcggg aataataagg 16500
 ctttgccaa acatgtgtaa tgtgctgttt caaattcgaa gtccatctac ttgtcatcta 16560
 caccgtctac actgtgaaat attcctcgaa actaatttac tctaccctaa tatataatat 16620
 atacagcgca aacgtactcg ctagagatga aatccgtcaa caatttgagg accaagttcc 16680
 tccgacacca tcttgggcat cagcgttatc cacctcctct ggatgatttc tggacccttc 16740
 tttccaaagt atctctcgct ctogaagctc cactggtcat tcgctctcac ccaattacct 16800
 agcccatcac agtatcgagc aacctgcag tcaatgggac ctccccatct gggcactttc 16860
 ttgaagtttt ccatgaacct ggactcgagc tgcttgtggt attcgtcgac ccaaatcatg 16920
 gcacctgcca cgtctgtcct gagctcgttc attgcgatag tgacgatggt gtggctagag 16980
 accgcgtcag aatgataatg acgggcgaaa tgctaaagac tcaccgcgtc tcatcccag 17040
 cttgttcaag gttatacga acgacatcct tgaaagtata agaggtcagt tccctttttc 17100
 tgtggtggat aaagacggat tgetcacgtt tcccaagcag agcatatcaa tacagctctt 17160
 tctcaagttt ctttgataac cggatgattg atgactttcg tcgggtaact gcatgtcaaa 17220
 ctcgagaagt gcgaacgatg gtttcgcgcc gatggtggtg cggcgaactt cgaggtacga 17280

ES 2 747 844 T3

ctcgatgcct cggacgtggg agttggtgcg gtcggcggcc tgttgtagca cagactccag 17340
gtactcgtcg aagggtgcga tgaagcgctt ctgogcttga gtgctcgtg tcttctttgc 17400
gaggtcccag aatctatagc gtacgatcag tccatggatc atgggtatct aaatgtgtgg 17460
tactgactgt cgagatactt ctctccaat ccaactctccg gcgggcccag gcttctctgt 17520
attccgaatc gcatccatga ctatgtcctt ctgtgcgcgc actggataca acagtcagaa 17580
taactgttat ctcaaggaga atttaatggt gagcttacct tcctcctccg tcgagacgtc 17640
cgagtactcg tcgataacga agaaactgtg tgatcgcatac aacaaaatat ccccgagcgc 17700
tggcttgaag acctacaggt tcatgagatc acaccgccta cggagacgtg ctagagacga 17760
aaaaggatta ggctccatgg acgggcagtt ctctgcatc actcgccttc gtcccagg 17820
gggtacgcga atgaggccag gagatctgag tgatggtgta cagtcagtct agctctaata 17880
agcgcctgag tcttgactta ctgaagtcgc agcggtcgaa agcttcctga gccttcgttt 17940
ggaaagccta acgctcttgg tcagaatatt cgacataaac gtattctact tccaacttac 18000
tcggaagctt ttgcccagg ccgctgatgc cttcttgact tcggcatagt gggggtgag 18060
atggcggggc cattgccagt tggcaagagt atcaggaagg aagatgcgtt gagacattga 18120
gaatagtgga aggtgtggg gatttctgca aggtcagtct ggaacgctct taaggatagg 18180
aataattacc ggtcatttta aaaggcaatc ggagatgaac ggcttgatg agcctgaatg 18240
aaaccggtg caaagccat agatggcact gggtcacgag attcactctc attccggtg 18300
ggaacttctc tatgtgccgt agctaccgtg cctcagtgta acataagccg ggcgcgacct 18360
tggcgtaact ccggtatcta actccatacg atgatccttc aagcgcact ctgcgcgaac 18420
attcctgcgc ggagttggaa atgcacctga aaacggaaat gccgcctcac ctctatgcca 18480
tgacatacga atttgaactg agattggtcg ttcattgaat atgttagct atagtgaaca 18540
taagaacagg gcgaatattg tttccctcac cctgggtcga tgatccatga tcggcgttct 18600
ggtggtttta cactgtttca taacgctcct gtatcaacgg taagtcggcc tcatcgcagc 18660
ggctcgcaat catcctcact gnttgactt taggcgagtc tcgacgttgc gaagcggag 18720
tgcaaaagta aaagtgttaa aaaaggccgg tgtattatgc aagaacggag tcgaatcaaa 18780
agtaccaaca gacttcttca atattcactc cctttcgggtg agagcggctc acctaccgtg 18840
acgacgatcc gtggcatgtc agcgtctatg gaccggaact tagcgcagct ttattgatga 18900
aactagcgtc tacttatcgg aaattgttta aataaaggat tgtatgagcc gataagtagc 18960
attcaggcta tgcacatgcc tccattggtc ctccgcatct tgtgagagtt cagggtacac 19020
tctcaaagtc ttgcgtagta ttccggcaga tataacgata accgatctga tccatcttcc 19080
acgcatcttg atgctaaagt aatataccac acttaccgcc gcctcagcgt ctgattcatg 19140
agagtttctc gctgactaat aaccttgaa gtattcgagc tcctaccgac ctccggcac 19200

ES 2 747 844 T3

cgccactaga agcataactgc gaaccatagc aaagtctcaa cgcgccaggt acattcacac 19260
 tgattagga atttagtgtt ccgtagctcc agttggaaga aacagtgccca ccggaaccga 19320
 acaattgacc acgcgtacta cttagtctaa gtgggaatct cggatcgatc attgatccac 19380
 aaaaatcgaa aataactcaag tatgtagttg acatnnaaga aatgatcgtg caataattca 19440
 agaagctcgt caaaagcgtc aaaatcggcg ctcagagccn attcacgaga aatdddgcca 19500
 aatcagctgg actgagggcg agatgttcgt ctattccaac ccgaaaccct gaggtggacg 19560
 aaaggcggac ggcgatgcat gcggaaatga tggtagcggat aatggcaac gccggcccng 19620
 gttagcgtat tttgcttttt acctccgcc ttgtcgaaca gctgagcagg gccttgttct 19680
 atgaaatttc acctttgaaa agagatcgtt acaggatgtt tgntgtggat cagttcatgt 19740
 ttcgaatctg caccggccg gtgtcggctc cgatccaaac ccgcgacatg gaacctttcg 19800
 gagcaatggg aactgattac tgtagtttcg acccaatggg tatgcctctc tgtactttcc 19860
 tatgtaggca aaccgagcat cgggaaatag tacatctcca gggtgatttc acctcctcgc 19920
 ttctcctcgc tcatggattc tgcgtctctt gccgttgtcg tttgggctat ccttctcgtt 19980
 ttgtggctca gaaggatttt cggcacaacgc tcttccctcc ctctcccccc tgcacctcct 20040
 ggttatcccg ttataggcaa cctgcttgac ttggctaaca atgacgtgca tatcagagcc 20100
 aggcactggg cccgcaattt tggtagtctc atcgtgcaag ttaataacag tgaataaaca 20160
 gcttgatctc tcagatgacg atgtcatctc tctcaaagtt ctgggaaaga ccatgataat 20220
 cttgaactcg cccaccgcag tctctgatct ctttgacaag cgagcctcta actactcaga 20280
 tcggcctgac atgccaatga ttgttgactt gtgcgtaatt tctgacttac tcgtaacaac 20340
 ctttgacacg ttattgatca ggatgggtgc gtctcacctg gatgcggtac aagagcaact 20400
 tttatgagat gttttccgta taggatggga ctggacattt gccctcatgc ggtacgggcc 20460
 gcgttggaag gaacacagac gggctctcaa caaccacttc aatattggca cttctggtgc 20520
 gtcagaagac cggcatattc agctacgcat ctgtcgcgag ctcttgtcct tgatgttcca 20580
 atctccttcg aaataccttg agaaccttcg gcagtacgtc ggtgtggttt ctggaagaga 20640
 ccgctcctta tcgacttcat tttctaagct acacaggtca tattattctc aaacgcacgt 20700
 atggacacac ggtagtcgat gagaaggacc cttatatccg cctgggtgag gcagcgtccc 20760
 aaagtacatc tgaaggtaat tcacttctc aaattdgtct taatgatggc agttctcata 20820
 agttttgtag ccgcagctcc aggagccttc cttgtcgatc tattcccatc gagtacgttg 20880
 ataaaattca gtcattcagc ctgccaatat ctgctctgta gtgaagtata ttctgagtg 20940
 ggtcccgggc gcgcaattca agagaaaggc tagggagtgg taagggtctt ctggtgatat 21000
 tcttagctct ctccatttat gaactggacg ctataggcgc aaactctcag aagctatgat 21060

ES 2 747 844 T3

caacgcaccg tatgacatgg cgaaggggaa atttgtatgt tcaactttga tcggatcacg 21120
ctgcgaacta aattctgttg catcgtgaag gatgaaggaa atgccgaacc ttgctttgtc 21180
tctgcttgtc tcgagcagaa caagaccgct tccggccaag ggttgagcga agaactcatt 21240
aaagacaccg ctgccgtcgc atatgctgcc ggtcagtttg tctctatttc gtttcgcac 21300
cctattctca tgaacttgat aattaggtgc tgatacggta caattcctct gggccttcat 21360
tcgacttaca acacttaca attataacag agcgtctcca ctctaacgac atttattctc 21420
gccatgacac tttatcccga tgttcagaaa gcagcacaag ccgaactcga tgcattactt 21480
ggcggcgaac gtttaccoga ttttgccgat aaaactcggc tgccatacgt cactgcaata 21540
ttaaaggagg tccttcggtg gattcctgtt ttaccgatgg gtgagtacga gccgggtttc 21600
ttaatttaca cgacccaacc acgttttagc cgtacctcac cgagccgtta atgcggacac 21660
ttacaaagga tactacatcc cggctggtgc cttcgtgtac ggaaacgcat ggtgagtcta 21720
ttgtcgacat atgaatgctg cctaaacctg tgctagggct atccttcata accccgacat 21780
ctttgcagac cccgagacat tcagaccaga cagattcatc gagaacccaa cgttactcaa 21840
cccgatagac aatgggggat ttgggtttgg aagacggtga gcaggaacat cattgtctcg 21900
ttccagaatt gcctaatttc atcaacagcg cgtgtgcagg gagggtaatg gcgcttgaca 21960
ccatgtggat agccatggca tccattctag ccgtcttcga catttcgaaa gccgtcgatg 22020
aacgcgggaa tgagatctca cctcccgtaa agctcagtcc gggaacgata aggtaagcgc 22080
tgttagataa agactacatt tctggctgac cgttggtgct atcaaaagtc acccggcacc 22140
atcccggtgt attatcaaac ctcgatctaa ggctgccctt gaactcatta cgcaggatga 22200
ctga 22204

REIVINDICACIONES

1. Microorganismo huésped que se transforma para comprender y expresar un polinucleótido que codifica un polipéptido con actividad citocromo P450 monooxigenasa y que comprende o consiste en una secuencia nucleica seleccionada del grupo que consiste en (i), (ii) y (iii):
- 5 (i): a) SEQ ID NO 5 a 8 de la lista de secuencias adjunta; b) una secuencia de ácido nucleico complementaria a SEQ ID NO 5 a 8; c) secuencias de ácido nucleico que se hibridan en condiciones rigurosas con las secuencias de ácido nucleico definidas en a) y b) o sus hebras complementarias;
- 10 (ii) un vector que contiene una secuencia de ácido nucleico de (i), que codifica un polipéptido con actividad citocromo P450 monooxigenasa, estando la secuencia de ácido nucleico unida operativamente a secuencias de control reconocidas por un microorganismo huésped transformado con el vector, preferiblemente el vector es un vector de expresión,
- (iii) SEQ ID NO 9 de la lista de secuencias adjunta;
- transformándose adicionalmente el microorganismo huésped para expresar una 2-hidroxi-3-metilglutaril-CoA (HMG-CoA) reductasa, y
- 15 seleccionándose del grupo que consiste en hongos, incluyendo levadura y bacterias.
2. Microorganismo huésped de la reivindicación 1, en el que el polinucleótido que codifica el polipéptido con actividad citocromo P450 monooxigenasa se adapta al uso de codón de la respectiva célula.
3. Microorganismo huésped de cualquiera de las reivindicaciones 1 ó 2, en el que el microorganismo de célula huésped se selecciona de *Saccharomyces spp.* en particular *Saccharomyces cerevisiae*; *Escherichia coli*; *Aspergillus nidulans*; homobasidiomicetos, en particular del género *Armillaria*, en particular *Armillaria gallica*, *Armillaria mellea*.
- 20 4. Microorganismo huésped de cualquiera de las reivindicaciones 1 a 3, en el que el microorganismo huésped se transforma adicionalmente para comprender una protoiludeno sintasa, en particular una protoiludeno sintasa de uno de los siguientes: *Armillaria spp.*, en particular de *Armillaria gallica*, y/o para expresar al menos uno de los siguientes: (i) una NADPH-citocromo P450 reductasa, en particular una NADPH:citocromo reductasa de *Taxus chinensis*.
- 25 5. Método para producir un protoiludeno hidroxilado y/o un éster arílico de tipo protoiludeno sesquiterpenoide, comprendiendo el método las etapas de
- 30 a) proporcionar un microorganismo huésped que se selecciona del grupo que consiste en hongos, incluyendo levadura y bacterias, y que se ha transformado para expresar de manera heteróloga SEQ ID NO 9, o para expresar de manera heteróloga al menos todo lo siguiente: i) una protoiludeno sintasa, y ii) al menos una citocromo P450 monooxigenasa, en la que la citocromo P450 monooxigenasa tiene una secuencia de aminoácidos que se selecciona de
- 1) una de SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 3 y SEQ ID NO 4,
- 35 2) una secuencia de aminoácidos de una variante alélica de una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO 1 a 4, en la que dicha variante alélica está codificada por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8, o
- 40 3) una secuencia de aminoácidos de un ortólogo de una secuencia de aminoácidos mostrada en cualquiera de SEQ ID NO 1 a 4, en la que dicho ortólogo está codificado por una molécula de ácido nucleico que se hibrida en condiciones rigurosas con la hebra opuesta de una molécula de ácido nucleico mostrada en SEQ ID NO 5 a 8,
- en el que al menos uno de (i) o (ii) es ajeno a dicho microorganismo huésped, y
- 45 en el que el microorganismo huésped se transforma adicionalmente para expresar una 2-hidroxi-3-metilglutaril-CoA (HMGCoA) reductasa;
- b) hacer crecer, en condiciones de nutrientes adecuadas que permiten la producción de los ésteres arílicos de tipo protoiludeno sesquiterpenoide, el microorganismo huésped proporcionado en la etapa a); produciendo así el protoiludeno hidroxilado y, según sea el caso, el éster arílico de tipo protoiludeno sesquiterpenoide.
- 50 6. Método de la reivindicación 5, en el que el microorganismo huésped se ha transformado con SEQ ID NO 5 a 8 de la lista de secuencias adjunta para expresar una citocromo P450 monooxigenasa.

7. Método de la reivindicación 5 ó 6, que comprende además la etapa c):
- c) aislar dicho protoiludeno hidroxilado y/o dicho éster arílico de tipo protoiludeno sesquiterpenoide del microorganismo huésped o del medio de crecimiento.
- 5 8. Método de cualquiera de las reivindicaciones 5 a 7, en el que la protoiludeno sintasa y/o la citocromo P450 monooxigenasa es una protoiludeno sintasa y/o una citocromo P450 monooxigenasa, respectivamente, de uno de los siguientes: *Armillaria spp.*, en particular de *Armillaria gallica*.
9. Método de cualquiera de las reivindicaciones 5 a 8, en el que en la etapa a) se proporciona un microorganismo que se ha transformado adicionalmente para expresar una NADPH-citocromo P450 reductasa, preferiblemente una NADPH:citocromo reductasa de *Taxus chinensis*.
- 10 10. Método de cualquiera de las reivindicaciones 5 a 9, en el que el éster arílico de tipo protoiludeno sesquiterpenoide se selecciona de meleólidos y armililorselinatos.
11. Método de cualquiera de las reivindicaciones 5 a 10, en el que el protoiludeno hidroxilado se selecciona de protoiludeno monohidroxilado, en particular 8-hidroxi-6-protoiludeno, protoiludeno dihidroxilado, en particular 8,13-hidroxi-6-protoiludeno.

15

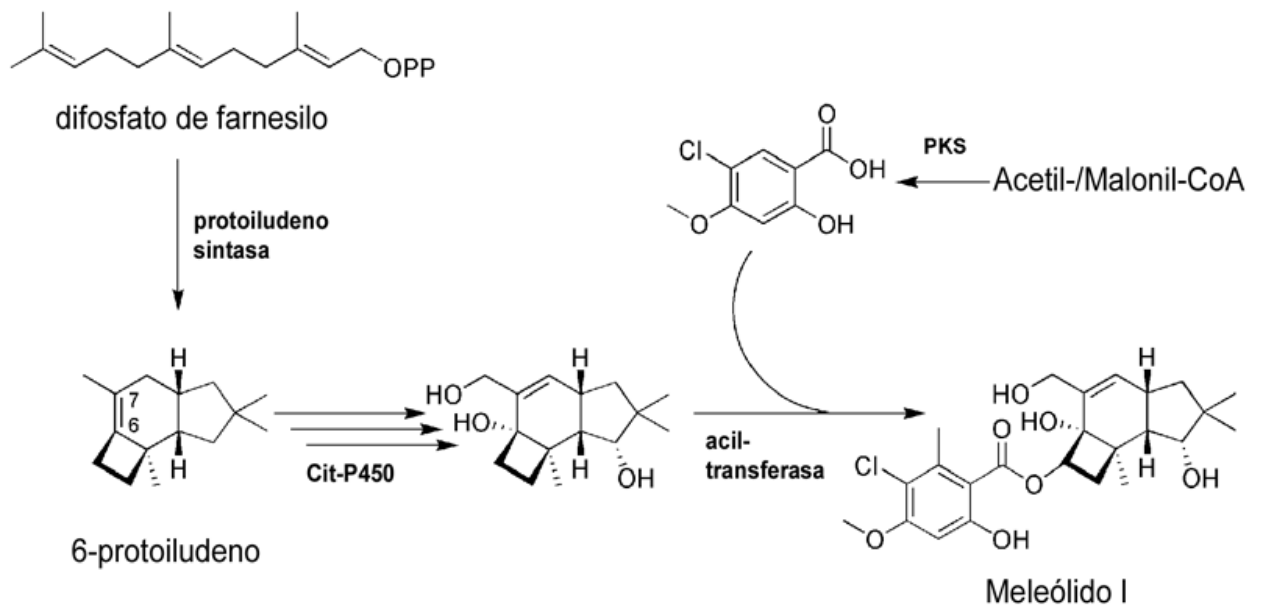


Fig. 1

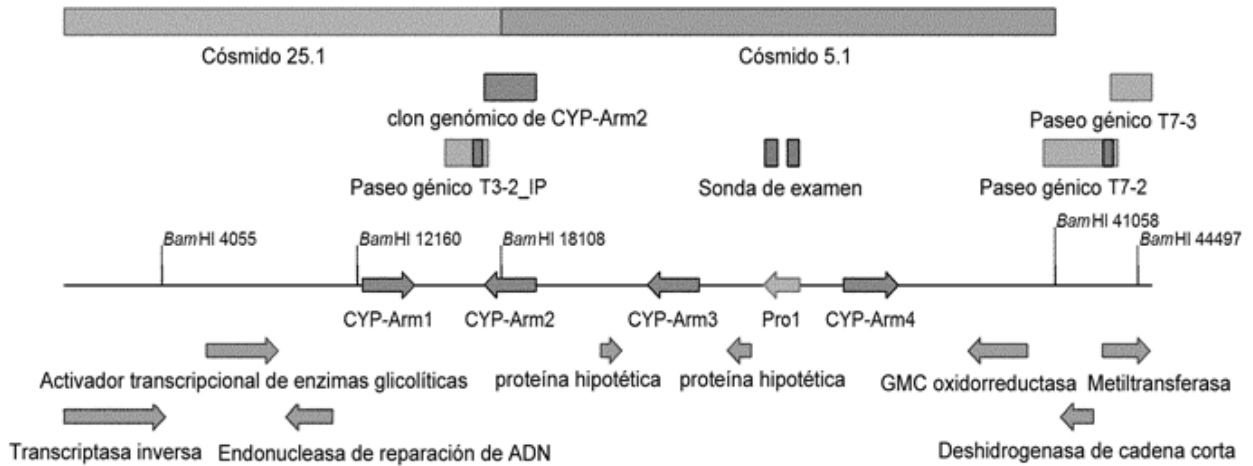


Fig. 2A

SEQ ID NO.1: CYP-Arm1

MALFSAYALAFSLLMVPLILYILRIGRRESGLPPGPSTLPLVGNLHQLPHGSLHLQFTAWAKEFGGIF
 SLKFGPGTVIVATSPRAVRELIDQKSASTSDRPPSHFSNVITGGNNIGFARYSDYWRRGRRVMHS
 MLTKKACVNHLTIQRAEASQLMYDYLVPEKVNLSHQILVANVGKSDVQCQEFVAHGQRYANSVITS
 ILAGTRSPHHTSPLVTAFFQMQUEWTHLLTPGAHPPVDMVPFLKYIPGSWKQICAKMKVSQEELY
 GGMIDACAKRVERGIRNGCFLEGELENKDVDRGLLRGMCSALMEGGSDSTSIYLQSFILMLVAHP
 DVQAKARAEIDSVVGLDRLPDIEDMDNLPYVSAVIKEVLRRLRPITALGAPHYSTQPEVVDGYVIPKN
 SMIFMNQWGMLHDPDSYDQPESFMPERFLTSPFGTRPGADDTGRSKDIYFGGRRICVGMHFG
 QNSLAITSMYLIWAFNLTAIDPQTKNPIVDINEYDISLNTIPKPKCDFQVRSEEHKRMVENAFAE
 SRQVLAPFEQD

Fig. 2B

SEQ ID NO.2: CYP-Arm2

MTHASSAWFLAAVAIVTFIVVRRIRSSWRKLPPEGPRGLPIVGNLLQLRSKQWLTFTELGKKYGDLM
YFNVAGQPLIVINSLRVATDLLDRAKFSRPRNIVASDIMTRGMFVAFAPYGNAWRHRMKAHEGL
NKNIVNQYHPIQIKEAVLLANDLLAEPNRWVSHVRRTAASTIMSIVYDKPATSEQDPSIKRINDFAAR
LTRAAMPGAHFVESFPWMLRIPSKYAKWKREAEGWYAKDSSMFESLFHSVKDRVAEGNNRPSFV
ATLIQGAGRHLTDHESSWLAGAMYTAGVESSAAISWWMLAMILYPDAQKRAQAELDKVVGRD
RLPAFSDYEHLPYVRAMVKETLRWRAVDPVGLPHRSTEDDIYNGYFIPAGSILIANVWHINRDPEN
YGLDAEHFDPARHLDETGQLAPVAGTKEENHVSFGFGRRICVGRHIANDTLFAVIATLLWAIHIEPA
TDEKGASPLDVDGCIEDGVVVRPLPFVKVTPRFSEAQAIVEQERELLGHY

Fig. 2C

SEQ ID NO.3: CYP-Arm3

MTRIFSEEVSPIWILTAIVVAYTTVRYLRSPWRNLPPGPRGLPLIGNLLELRGKQWLTFTELGKKY
GDLMYFNVAGQPLVVLNSQKVAADLLDRRAGKYSRPRNIVASDIMTGGNLVVFTRYGDVWRRM
RKAHEGLNKGVVYKYHPIQTAEAVLLTAGVLAEPEKWNSHLRRTAASAIMSMVYDTPPTSEQDP
SVKNINEFVARLTRAAMPGAHFVEFFPWMRYIPSKYAKWKREAEESYAKDSAMFEGLFNGVKDR
VAKGDERPSLASTLIQDAGRHDLDRENSWLAGTMYAAGAETTSGVMSWWTLAMIVYPETQKRA
QAELDAVVGRDRLPSFADYEHLPYIRAMVKEALRWRMVDVGLPHTSTEDDVYDGYFIPAGTIIAN
VWHLNRDPEIYGPDAEHFNPARHLDKDGKLAGPADTKEESHVTYGFGRRICVGRHVANNSLFIDI
AMMLWAMNIERATDENGVPLPLDVDGCVEDGLVTRPVPFKAKITPRFREAQAIVEQERELLGYH

Fig. 2D

SEQ ID NO.4: CYP-Arm4

MDSASLAVVWAILLVLWLRRIFGQRSSLPLPPAPPGYVIGNLLDLANNDVHIRARHWSRNFDDD
VISLKVLGKTMILNSPTAVSDFDKRASNYSDRPDMPMIVDLMGWDWTFALMRYGPRWKEHRRV
FNNHFNIGTSGASEDRHIQLRICRELLSLMFQSPSKYLENLRHYTGHIILKRTYGHTVVDEKDPYIRL
VEAASQSTSEAAVPGAFLVDLFPMSKYIPEWVPGAQFKRKAREWRKLSEAMINAPYDMAKGFDF
EGNAEPCFVSACLEQNKTASGQGLSEELIKDTAAVAYAAGADTSVSTLTFFILAMTLYPDVQKAAQ
AELDALLGGERLPDFGDKTRLPYVTAILKEVLRWIPVLPMAVPHRAVNADTYKGYYPAGAFVYGN
AWAILHNPDI FADPETFRPDRFIENPTLLNPIDNGVFGFGRRACAGRVMALDTMWIAMASILAVFDI
SKAVDERGNEISPPVKLSPGTISHPAPFPCIIPRSKAALELITQDD

Fig. 2E

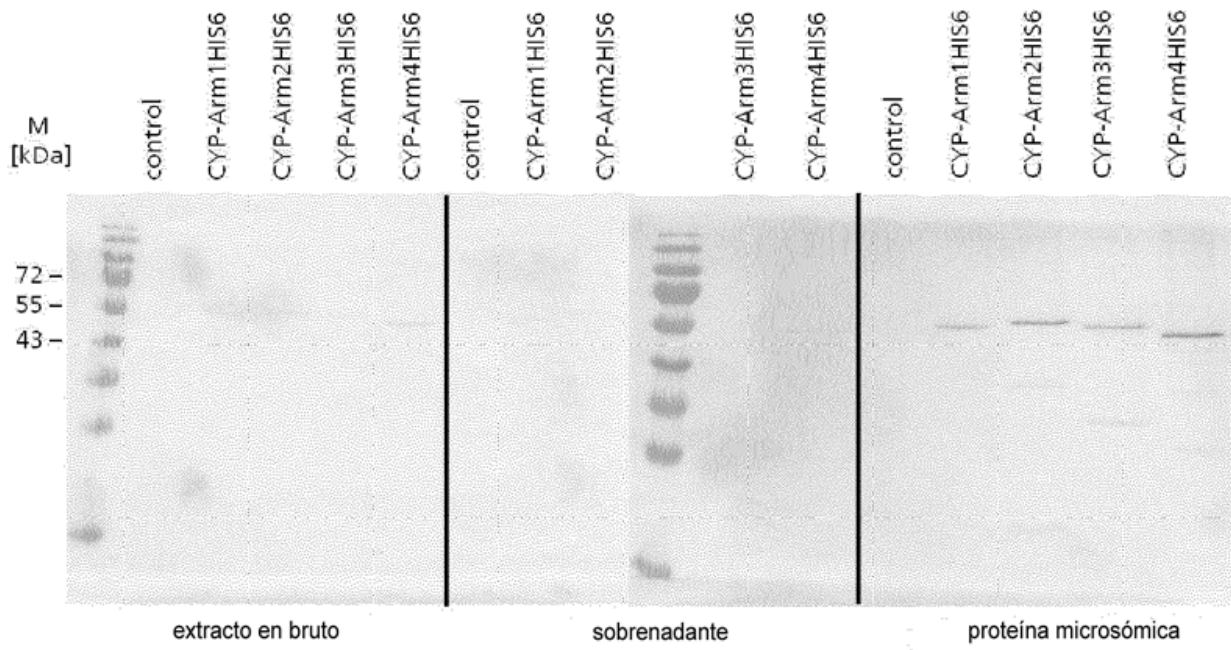


Fig. 3

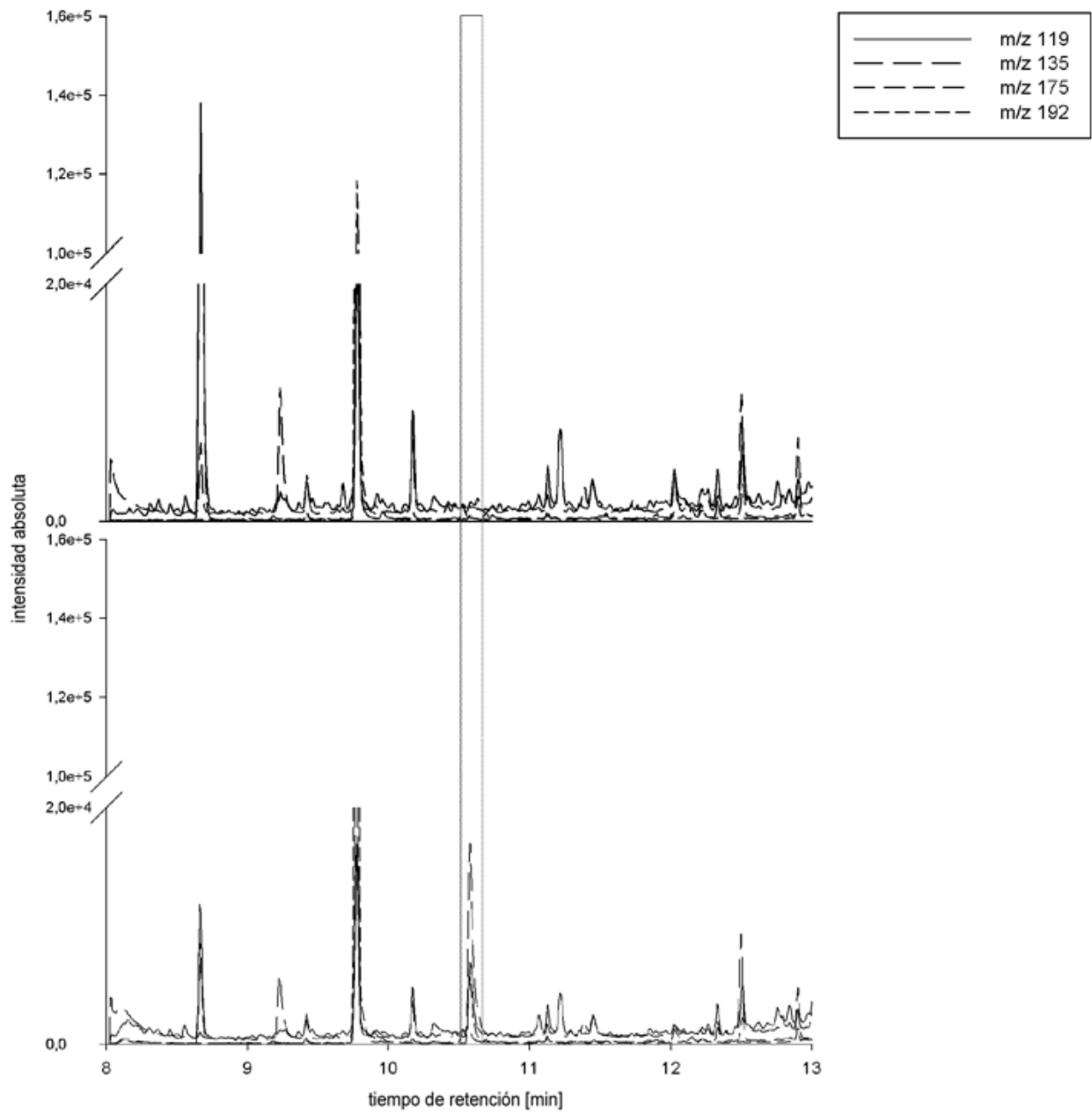


Fig. 4

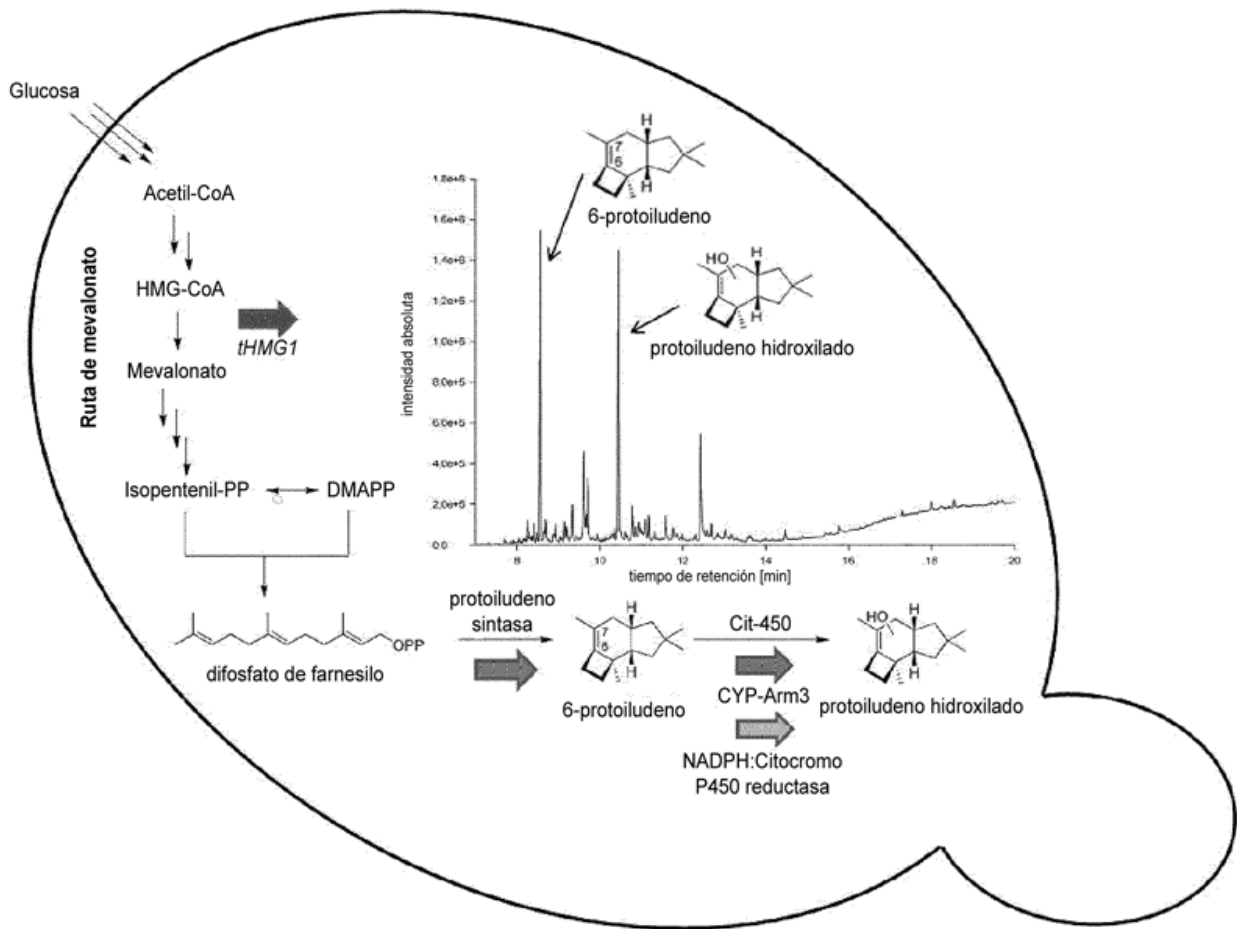


Fig. 5

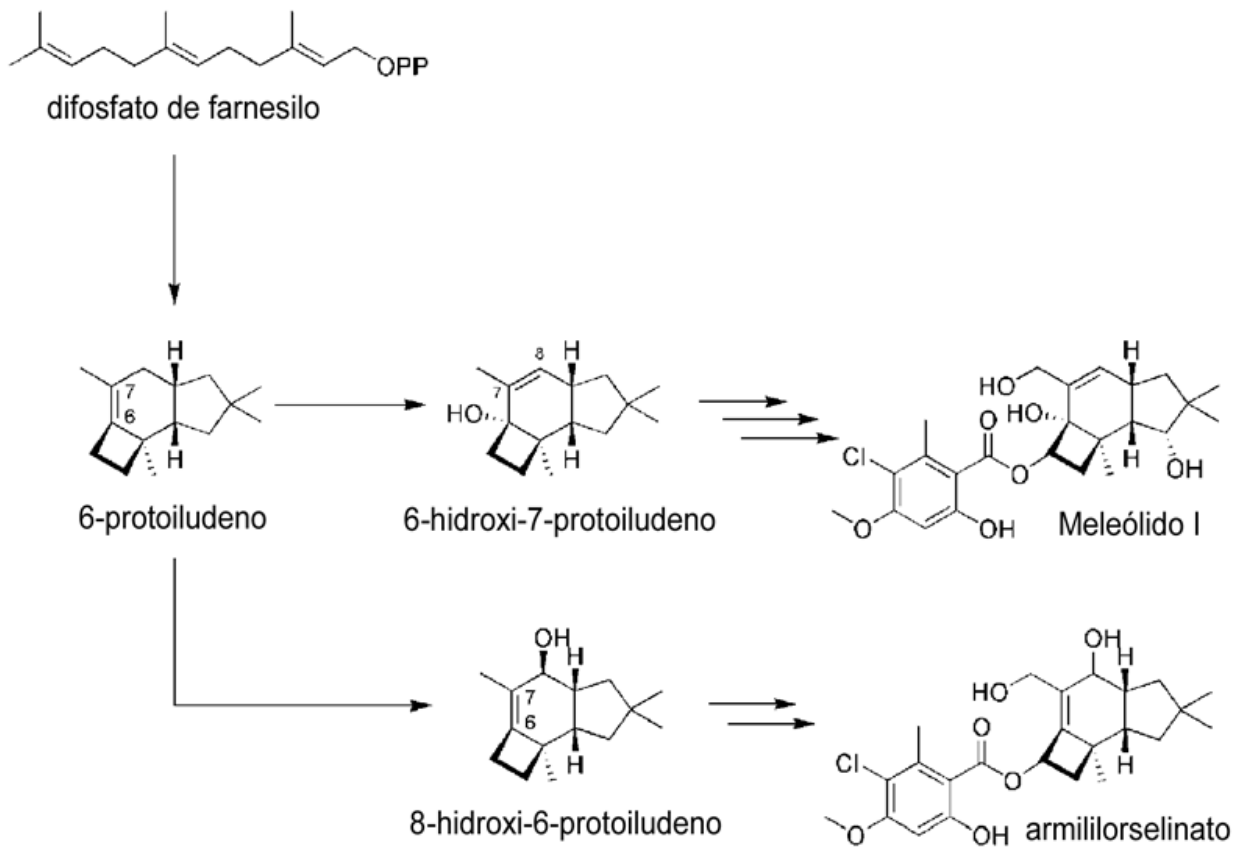


Fig. 6