

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 386 359**

51 Int. Cl.:  
**C12N 1/15** (2006.01)  
**C12N 1/16** (2006.01)  
**C12N 1/18** (2006.01)  
**C12N 5/10** (2006.01)  
**C12N 15/10** (2006.01)  
**C12N 15/31** (2006.01)  
**C12N 15/12** (2006.01)  
**C12N 15/63** (2006.01)  
**C12N 15/64** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05775484 .8**  
96 Fecha de presentación: **21.07.2005**  
97 Número de publicación de la solicitud: **1778831**  
97 Fecha de publicación de la solicitud: **02.05.2007**

54 Título: **Células huésped modificadas genéticamente y uso de las mismas para producir compuestos isoprenoides**

30 Prioridad:  
**27.07.2004 US 592009 P**

45 Fecha de publicación de la mención BOPI:  
**17.08.2012**

45 Fecha de la publicación del folleto de la patente:  
**17.08.2012**

73 Titular/es:  
**THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
1111 FRANKLIN STREET, 12TH FLOOR  
OAKLAND, CA 94607-5200, US**

72 Inventor/es:  
**KEASLING, Jay D.;  
PARADISE, Eric M. y  
KIRBY, James**

74 Agente/Representante:  
**Ponti Sales, Adelaida**

ES 2 386 359 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

CÉLULAS HUÉSPED MODIFICADAS GENÉTICAMENTE Y USO DE LAS MISMAS PARA PRODUCIR COMPUESTOS ISOPRENOIDES

5

CAMPO DE LA INVENCION

**[0001]** La presente invención se encuentra en el campo de la producción de compuestos isoprenoides y, en particular, células huésped que son modificadas genéticamente para producir compuestos isoprenoides.

10

ANTECEDENTES DE LA INVENCION

**[0002]** Los isoprenoides constituyen un grupo extremadamente amplio y diverso de productos naturales que tienen un origen biosintético común, es decir, un único precursor metabólico, isopentenil difosfato (IPP). Los compuestos isoprenoides también se refieren como "terpenos" o "terpenoides". Se han descrito alrededor de 40.000 isoprenoides. Por definición, los isoprenoides están formados de las denominadas unidades de isopreno (C5). El número de átomos de C presentes en los isoprenoides es normalmente divisible por cinco (C5, C10, C15, C20, C25, C30 y C40), aunque se han descrito isoprenoides irregulares y politerpenos. Los miembros importantes de los isoprenoides incluyen carotenoides, sesquiterpenoides, diterpenoides, y hemiterpenos. Los carotenoides incluyen, por ejemplo, licopeno,  $\beta$ -caroteno, y similares, muchos de los cuales actúan como antioxidantes. Los sesquiterpenoides incluyen, por ejemplo, artemisinina, un compuesto que tiene actividad contra la malaria. Los diterpenoides incluyen, por ejemplo, taxol, un agente quimioterapéutico contra el cáncer.

15

20

**[0003]** Los isoprenoides comprenden la familia más numerosa y estructuralmente diversa de productos naturales. En esta familia, los terpenoides aislados de plantas y otras fuentes naturales se utilizan como compuestos de aromas y fragancias comerciales, así como fármacos contra la malaria y el cáncer. Una mayoría de los compuestos terpenoides que se utilizan actualmente son productos naturales o sus derivados. Los organismos de origen (por ejemplo, árboles, invertebrados marinos) de muchos de estos productos naturales no son susceptibles de un cultivo a gran escala necesario para producir cantidades comercialmente viables ni de manipulación genética para una mayor producción o derivatización de estos compuestos. Por lo tanto, los productos naturales deben producirse semisintéticamente a partir de análogos o sintéticamente utilizando síntesis química convencional. Además, muchos productos naturales presentan estructuras complejas y, como resultado, son actualmente caros o imposibles de sintetizar. Dichos productos naturales deben extraerse de sus fuentes nativas, tales como árboles, esponjas, corales y microbios marinos; o producirse sintéticamente o semisintéticamente a partir de precursores abundantes. La extracción de un producto natural a partir de una fuente nativa está limitada por la disponibilidad de la fuente nativa; y la producción sintética o semisintética de productos naturales puede sufrir un rendimiento bajo y/o un coste elevado. Dichos problemas de producción y disponibilidad limitada de la fuente natural pueden limitar el desarrollo comercial y clínico de dichos productos.

25

30

35

**[0004]** La biosíntesis de productos naturales de isoprenoides en células huésped modificadas podría aprovechar el potencial comercial y terapéutico sin explotar de estos recursos naturales y producir productos de química fina y farmacéuticos menos caros y más ampliamente disponibles. El obstáculo principal para la biosíntesis de terpenoides a nivel elevado es la producción de precursores de terpenos. En *Saccharomyces cerevisiae*, el mecanismo del mevalonato proporciona la producción de isopentenil difosfato (IPP), que se puede isomerizar y polimerizar en isoprenoides y terpenos de valor comercial. También se producen otros precursores valiosos, incluyendo farnesil difosfato (FPP) y geraniogerani difosfato (GPP). Sin embargo, mucho del flujo de reacción está dirigido hacia las posteriores etapas no deseadas del mecanismo de esterol, dando lugar a la producción de ergosterol.

40

45

**[0005]** Existe la necesidad en la técnica de células huésped mejoradas productoras de isoprenoides y precursoras de isoprenoides que proporcionen una producción a nivel elevado de compuestos isoprenoides, así como de los precursores de poliprenil difosfato de dichos compuestos. La presente invención está dirigida a esta necesidad y proporciona ventajas relacionadas.

50

Literatura

55

**[0006]** Publicación de patente de Estados Unidos No. 2004/005678; Publicación de patente de Estados Unidos No. 2003/0148479; Martin et al. (2003) Nat. Biotech. 21(7):796-802; Polakowski et al. (1998) Appl. Microbiol. Biotechnol. 49: 67-71; Wilding et al. (2000) J Bacteriol 182(15): 4319-27; Publicación de patente de Estados Unidos No. 2004/0194162; Donald et al. (1997) Appl. Env. Microbiol. 63:3341-3344; Jackson et al (2003) Organ. Lett. 5:1629-1632; Publicación de patente de Estados Unidos No. 2004/0072323; Publicación de patente de Estados Unidos No. 2004/0029239; Publicación de patente de Estados Unidos No. 2004/0110259; Publicación de patente de Estados Unidos No. 2004/0063182; Patente de Estados Unidos No. 5,460,949; Publicación de patente de Estados Unidos No. 2004/0077039; Patente de Estados Unidos No. 6,531,303; Patente de Estados Unidos No. 6,689,593; Hamano et al. (2001) Biosci. Biotechnol. Biochem. 65:1627-1635; T. Kuzuyama. (2004) Biosci. Biotechnol. Biochem. 68(4): 931-934; T. Kazuhiko. (2004) Biotechnology Letters. 26: 1487-1491; Brock et

60

65

al (2004) Eur J. Biochem. 271: 3227-3241; Choi, et al (1999) Appl. Environ. Microbio. 65 4363-4368; Parke et al., (2004) Appl. Environ. Microbio. 70: 2974-2983; Subrahmanyam et al (1998) J. Bact. 180: 4596-4602; Murli et al. (2003) J. Ind. Microbiol. Biotechnol. 30: 500-509; Szkopinska et al, (2000) Biochemical And Biophysical Research Communications, vol. 267, no. 1, 473-477; Kontoyiannis et al (1999) Abstract Of The Interscience Conference On Antimicrobial Agents And Chemotherapy, vol. 39, Septiembre 1999, páginas 586-586, Wentzinger et al (2002) Plant Physiology. vol. 130, no. 1,334 - 346, EP-0982404A1.

#### DESCRIPCIÓN RESUMIDA DE LA INVENCION

10 **[0007]** La presente invención proporciona una célula huésped de *Saccharomyces cerevisiae* que produce un precursor de isoprenoide o compuesto isoprenoide a través del mecanismo de mevalonato, en el que dicha célula es modificada genéticamente para comprender: (a) un ácido nucleico heterólogo integrado en el cromosoma de la célula huésped que codifica una 3-hidroxi-3-metilglutaril coenzima-A reductasa (HMGR) truncada que carece del dominio que abarca la membrana y que mantiene su dominio catalítico C-terminal; (b) una farnesil difosfato sintasa (FPPS) heteróloga integrada en el cromosoma de la célula huésped para incrementar el nivel de actividad de dicha FPPS; y (c) un ácido nucleico heterólogo integrado en el cromosoma de la célula huésped para disminuir el nivel de actividad de escualeno sintasa; donde la célula huésped comprende además una sesquiterpeno sintasa heteróloga y es capaz de producir un isoprenoide que deriva de la acción de dicha sesquiterpeno sintasa, y donde las modificaciones genéticas proporcionan la producción del isoprenoide derivado de la acción de dicha sesquiterpeno sintasa a un nivel que por lo menos aproximadamente el 50% superior al nivel del isoprenoide en una célula de control que no comprende las modificaciones genéticas. Se proporcionan métodos para la producción de un compuesto isoprenoide o un compuesto precursor de isoprenoide en dicha célula huésped. Los métodos implican generalmente cultivar dicha célula huésped modificada genéticamente bajo condiciones que promueven la producción de niveles elevados de un compuesto isoprenoide o compuesto precursor de isoprenoide y aislar el compuesto precursor de isoprenoide o compuesto isoprenoide del medio de cultivo. Se hace referencia aquí a continuación a una "célula huésped" como una célula huésped de *S. cerevisiae* según la reivindicación 1.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

30 **[0008]** La figura 1 es una representación esquemática del mecanismo de mevalonato en *Saccharomyces cerevisiae*. Se muestran las estructuras de los intermedios y nombres de los genes que codifican varias enzimas en el mecanismo.

35 **[0009]** La figura 2 es una representación esquemática de una parte del mecanismo de biosíntesis de esterol en un organismo que expresa amorfadieno sintasa (ADS). Se muestran las estructuras de los intermedios y nombres de los genes que codifican varias enzimas en el mecanismo.

40 **[0010]** Las figuras 3A y 3B representan la producción de amorfadieno por *S. cerevisiae* durante 96 horas de cultivo que expresa amorfadieno sintasa (ADS) (♦); ADS y 3-hidroxi-3-metilglutaril coenzima-A reductasa (tHMGR) truncada (•); ADS y upc2-1 (■); y ADS y ecm22-1 (▲). Los datos se muestran como la producción total (3A) y normalizados para la densidad celular (3B). Los datos son las medias ± desviación estándar (n=3).

45 **[0011]** La figura 4 representa la producción de amorfadieno en la cepa EPY212 de *S. cerevisiae* desarrollada en concentraciones de metionina de 0, 0,1, 0,3, 0,5 y 1 después de 64 y 87 horas de cultivo. Los datos son los promedios de las medias de las dos muestras.

50 **[0012]** La figura 5 representa la producción de amorfadieno por *S. cerevisiae* por varias cepas de levadura durante 144 horas de cultivo expresante. Los datos son las medias ± desviaciones estándar (n=3).

**[0013]** La figura 6 representa una secuencia de nucleótidos que codifica una HMGR truncada.

**[0014]** Las figuras 7A y 7B representan una secuencia de aminoácidos de una HMGR truncada.

#### DEFINICIONES

60 **[0015]** Los términos "isoprenoide," "compuesto isoprenoide," "terpeno," "compuesto terpeno," "terpenoide," y "compuesto terpenoide" se utilizan indistintamente en la presente invención. Los compuestos isoprenoides están formados de varias cantidades de unidades los denominados isopreno (C5). El número de átomos de C presentes en los isoprenoides es normalmente divisible uniformemente por cinco (por ejemplo, C5, C10, C15, C20, C25, C30 y C40). Se han descrito isoprenoides y politerpenos irregulares y también se incluyen en la definición de "isoprenoide". Los compuestos isoprenoides incluyen pero sin limitación, monoterpénos, sesquiterpenos, triterpenos, politerpenos, y diterpenos.

65 **[0016]** Tal como se utiliza aquí, el término "prenil difosfato" se utiliza indistintamente con "prenil pirofosfato" e incluye monoprenil difosfatos que tiene un único grupo prenilo (por ejemplo, IPP y DMAPP), así como poliprenil

difosfatos que incluyen 2 o más grupos prenilo. Los monoprenil difosfatos incluyen isopentenil pirofosfato (IPP) y su isómero dimetilalil pirofosfato (DMAPP).

5 **[0017]** Tal como se utiliza aquí, el término "terpeno sintasa" se refiere a cualquier enzima que modifica enzimáticamente IPP, DMAPP, o un poliprenil pirofosfato, se manera que se produce un compuesto terpenoide. El término "terpeno sintasa" incluye enzimas que catalizan la conversión de un prenil difosfato en un isoprenoide.

10 **[0018]** La palabra "pirofosfato" se utiliza indistintamente aquí con "difosfato." De este modo, por ejemplo, los términos "prenil difosfato" y "prenil pirofosfato" son intercambiables; los términos "isopentenil pirofosfato" y "isopentenil difosfato" son intercambiables; los términos farnesil difosfato" y farnesil pirofosfato" son intercambiables; etc.

15 **[0019]** El término "mecanismo del mevalonato" o "mecanismo del MEV" se utiliza aquí para referirse al mecanismo biosintético que convierte el acetil-CoA en IPP. La ruta del mevalonato comprende enzimas que catalizan las siguientes etapas: (a) condensar dos moléculas de acetil-CoA a acetoacetil-CoA; (b) condensar acetoacetil-CoA con acetil-CoA para formar HMG-CoA; (c) convertir HMG-CoA en mevalonato; (d) fosforilar mevalonato en mevalonato 5-fosfato; (e) convertir mevalonato 5-fosfato en mevalonato 5-pirofosfato; y (f) convertir mevalonato 5-pirofosfato en isopentenil pirofosfato. La ruta del mevalonato se ilustra esquemáticamente en la figura 1.

20

25 **[0020]** Tal como se utiliza aquí, el término "prenil transferasa" se utiliza indistintamente con los términos "isoprenil difosfato sintasa" y "poliprenil sintasa" (por ejemplo, "GPP sintasa," "FPP sintasa," "OPP sintasa," etc.) para referirse a una enzima que cataliza la condensación 1'-4 consecutiva de isopentenil difosfato con sustratos cebadores alílicos, dando lugar a la formación de prenil difosfatos de varias longitudes de cadena.

30 **[0021]** Los términos "polinucleótido" y "ácido nucleico," utilizados indistintamente aquí, se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean ribonucleótidos o desoxinucleótidos. De este modo, este término incluye, pero sin limitación, ADN o ARN de monocadena, doble cadena o múltiples cadenas, ADN genómico, ADNc, híbridos ADN-ARN, o un polímero que comprende bases de purina y pirimidina u otras bases de nucleótidos naturales, modificados química o bioquímicamente, no naturales o derivadas.

35 **[0022]** Tal como se utiliza aquí, los términos "operón" y "unidad de transcripción individual" se utilizan indistintamente para referirse a dos o más regiones codificantes contiguas (secuencias de nucleótidos que codifican un producto génico, tal como un ARN o una proteína) que son regulados de forma coordinada por uno o más elementos de control (por ejemplo, un promotor). Tal como se utiliza aquí, el término "producto génico" se refiere a ARN codificado por ADN (o viceversa) o proteína que es codificada por un ARN o ADN, donde un gen comprenderá habitualmente una o más secuencias de nucleótidos que codifican una proteína, y pueden incluir también intrones y otras secuencias de nucleótidos no codificantes.

40

45 **[0023]** Los términos "péptido," "polipéptido," y "proteína" se utilizan indistintamente aquí y se refieren a una forma polimérica de aminoácidos de cualquier longitud, que pueden incluir aminoácidos codificados y no codificados, aminoácidos modificados o derivatizados química o bioquímicamente, y polipéptidos que tienen esqueletos peptídicos modificados.

50 **[0024]** El término "natural" tal como se utiliza aquí aplicado a un ácido nucleico, una célula, o un organismo, se refiere a un ácido nucleico, célula, u organismo que se halla en la naturaleza. Por ejemplo, una secuencia de polipéptido o polinucleótido que está presente en un organismo (incluyendo virus) que se pueden aislar de una fuente en la naturaleza y que no se ha modificado intencionadamente por el ser humano es natural.

55 **[0025]** El término "ácido nucleico heterólogo," tal como se utiliza aquí, se refiere a un ácido nucleico en el que por lo menos una de las frases siguientes es cierta: (a) el ácido nucleico es foráneo ("exógeno") a (es decir, no natural hallado en) un microorganismo huésped o célula huésped determinados; (b) el ácido nucleico comprende una secuencia de nucleótidos que se halla de forma natural en (por ejemplo, es "endógena a") un microorganismo huésped o célula huésped determinados (por ejemplo, el ácido nucleico comprende una secuencia de nucleótidos endógena al microorganismo huésped o célula huésped); sin embargo, en el contexto de un ácido nucleico heterólogo, la misma secuencia de nucleótidos que la hallada endógenamente se produce en una cantidad no natural (por ejemplo, mayor que la esperada o mayor que la hallada de forma natural) en la célula, o un ácido nucleico que comprende una secuencia de nucleótidos que difiere en la secuencia de la secuencia de nucleótidos endógena, pero codifica la misma proteína (que tiene la misma o sustancialmente la misma secuencia de aminoácidos) que la hallada endógenamente se produce en una cantidad no natural (por ejemplo, mayor que la esperada o mayor que la hallada de forma natural) en la célula; (c) el ácido nucleico comprende dos o más secuencias de nucleótidos que no se hallan en la misma relación entre sí en la naturaleza, por ejemplo, el ácido nucleico es recombinante. Un ejemplo de un ácido nucleico heterólogo es una secuencia de nucleótidos que codifica HMGR unida operativamente a un elemento de control de la transcripción (por ejemplo, un promotor) a la que una secuencia codificante de HMGR endógena (natural) no está normalmente unida

60

65

operativamente. Otro ejemplo de un ácido nucleico heterólogos un plásmido con un elevado número de copias que comprende una secuencia de nucleótidos que codifica HMGR. Otro ejemplo de un ácido nucleico heterólogo es un ácido nucleico que codifica HMGR, donde una célula huésped que no produce normalmente HMGR se modifica genéticamente con el ácido nucleico que codifica HMGR; dado que los ácidos nucleicos que codifican HMGR no se hallan de forma natural en la célula huésped, el ácido nucleico es heterólogo a la célula huésped modificada genéticamente.

**[0026]** "Recombinante," tal como se utiliza aquí, significa que un ácido nucleico particular (ADN o ARN) es el producto de varias combinaciones de etapas de clonación, restricción, y/o unión que dan lugar a un constructo que tiene una secuencia estructural codificante y no codificante diferenciable de los ácidos nucleicos endógenos hallados en sistemas naturales. En general, las secuencias de ADN que codifican la secuencia codificante estructural se pueden ensamblar a partir de fragmentos de ADNc y enlazadores de oligonucleótidos cortos, o a partir de una serie de oligonucleótidos sintéticos, para proporcionar un ácido nucleico sintético que es capaz de expresarse a partir de una unidad transcripcional recombinante contenida en una célula o en un sistema de transcripción y traducción libre de células. Dichas secuencias se pueden proporcionar en forma de un marco de lectura abierto no interrumpido por secuencias no traducidas internas o intrones, que están presentes normalmente en genes eucariotas. El ADN genómico que comprende las secuencias relevantes también se pueden utilizar en la formación de un gen o unidad transcripcional recombinante. Las secuencias de ADN no traducidas pueden estar presentes en 5' ó 3' con respecto al marco de lectura abierto, donde dichas secuencias no interfieren con la manipulación o la expresión de las regiones codificantes, y pueden de hecho actuar para modular la producción de un producto deseado mediante varios mecanismos (véanse las "secuencias reguladoras de ADN" a continuación).

**[0027]** De este modo, por ejemplo, el término polinucleótido o ácido nucleico "recombinante" se refiere a aquel que no es natural, por ejemplo, se produce mediante la combinación artificial de dos segmentos de secuencia separados a través de la intervención humana. Esta combinación artificial se realiza a menudo mediante medios de síntesis química, o mediante la manipulación artificial de segmentos aislados de ácidos nucleicos, por ejemplo, mediante técnicas de ingeniería genética. Esto se realiza normalmente para sustituir un codón por un codón redundante que codifica el mismo aminoácido o un aminoácido conservativo, mientras que normalmente introduce o elimina un sitio de reconocimiento de secuencia. Alternativamente, se realiza para unir juntos segmentos de ácidos nucleicos de funciones deseadas para generar una combinación deseada de funciones. Esta combinación artificial se realiza a menudo mediante medios de síntesis química, o mediante la manipulación artificial de segmentos aislados de ácidos nucleicos, por ejemplo, mediante técnicas de ingeniería genética

**[0028]** Por "constructo" se entiende un ácido nucleico recombinante, generalmente ADN recombinante, que ha sido generado con el objetivo de expresar una secuencia o secuencias de nucleótidos específicas, o va a utilizarse en la construcción de otras secuencias de nucleótidos recombinantes.

**[0029]** Tal como se utiliza aquí, el término "ácido nucleico exógeno" se refiere a un ácido nucleico que no está normalmente o se halla de forma natural en y/o es producido por una bacteria, organismo o célula determinada en la naturaleza. Tal como se utiliza aquí, el término "ácido nucleico endógeno" se refiere a un ácido nucleico que se halla normalmente en y/o es producido por una bacteria, organismo o célula determinada en la naturaleza. Un "ácido nucleico endógeno" también se refiere como un "ácido nucleico nativo" o un ácido nucleico que es "nativo" a una bacteria, organismo o célula determinada. Por ejemplo, un ADNc generado a partir del ARNm aislado de una planta y que codifica una terpeno sintasa representa un ácido nucleico exógeno para *S. cerevisiae*. En *S. cerevisiae*, las secuencias de nucleótidos que codifican HMGS, MK, y PMK en el cromosoma serían ácidos nucleicos "endógenos".

**[0030]** Los términos "secuencias reguladoras de ADN," "elementos de control," y "elementos reguladores," utilizados indistintamente aquí, se refieren a secuencias de control transcripcional y traduccional, tales como promotores, potenciadores, señales de poliadenilación, terminadores, señales de degradación de proteínas, y similares, que proporcionan y/o regulan la expresión de una secuencia codificante y/o producción de un polipéptido codificado en una célula huésped.

**[0031]** El término "transformación" se utiliza indistintamente aquí con "modificación genética" y se refiere a un cambio genético permanente o transitorio inducido en una célula después de la introducción nuevo ácido nucleico (es decir, ADN exógeno a la célula). El cambio ("modificación") genética se puede llevar a cabo mediante la incorporación del nuevo ADN en el genoma de la célula huésped o mediante el mantenimiento transitorio o estable del nuevo ADN como elemento episomal. Cuando la célula es una célula eucariota, en general se consigue un cambio genético permanente mediante la introducción del ADN en el genoma de la célula. En células procariontas, los cambios permanentes se pueden introducir en el cromosoma o a través de elementos extracromosómicos, tales como plásmidos y vectores de expresión, que pueden contener uno o más marcadores seleccionables para ayudar en su mantenimiento en la célula huésped recombinante.

**[0032]** "Unido operativamente" se refiere a una yuxtaposición en la que los componentes descritos están en una relación que les permite actuar en su manera pretendida. Por ejemplo, un promotor está unido operativamente a

una secuencia codificante si el promotor afecta a su transcripción o expresión. Tal como se utiliza aquí, los términos "promotor heterólogo" y "regiones de control heterólogas" se refieren a promotores y otras regiones de control que no están normalmente asociados con un ácido nucleico concreto en la naturaleza. Por ejemplo, una "región de control de la transcripción heteróloga a una región codificante" es una región de control de la transcripción que no está normalmente asociada con la región codificante en la naturaleza.

**[0033]** Una "célula huésped," tal como se utiliza aquí, indica una célula de *S. Cerevisiae* in vivo o in vitro que puede utilizarse o se ha utilizado como un receptor para un ácido nucleico (por ejemplo, un vector de expresión que comprende una secuencia de nucleótidos que codifica uno o más productos génicos tal como se define en la reivindicación 1), e incluye la progenie de la célula original que se ha modificado genéticamente por el ácido nucleico. Se entiende que la progenie de una célula individual puede no ser necesariamente completamente idéntica en morfología o en complemento de ADN genómico o total a la original, debido a una mutación natural, accidental o deliberada. Una "célula huésped recombinante" (también referida como "célula huésped modificada genéticamente") es una célula huésped en que se ha introducido un ácido nucleico heterólogo, por ejemplo, un vector de expresión. Por ejemplo, una célula huésped de *S. Cerevisiae* en cuestión es una célula huésped modificada genéticamente debido a la introducción en una célula huésped adecuada de un ácido nucleico heterólogo, por ejemplo, un ácido nucleico exógeno que es foráneo a la célula huésped, o un ácido nucleico recombinante que no se halla normalmente en la célula huésped.

**[0034]** Tal como se utiliza aquí, el término "aislado" pretende describir un polinucleótido, un polipéptido, o una célula que está en un medio diferente de aquel en que el polinucleótido, el polipéptido o la célula aparece de forma natural. Una célula huésped modificada genéticamente aislada puede estar presente en una población mixta de células huésped modificadas genéticamente.

**[0035]** Se pueden preparar cassettes de expresión que comprenden una región o regiones de inicio de la transcripción o de control de la transcripción (por ejemplo, un promotor), la región codificante para la proteína de interés, y una región de terminación de la transcripción. Las regiones de control de la transcripción incluyen aquellas que proporcionan la sobreexpresión de la proteína de interés en la célula huésped modificada genéticamente; aquellas que proporcionan una expresión inducible, de manera que cuando se añade un agente inductor al medio de cultivo, se induce o aumenta la transcripción de la región codificante de la proteína de interés hasta un nivel superior al de antes de la inducción.

**[0036]** Un ácido nucleico es "hibridable" a otro ácido nucleico, tal como ADNc, ADN genómico, o ARN, cuando una forma de cadena única del ácido nucleico se puede hibridar a otro ácido nucleico bajo las condiciones apropiadas de temperatura y fuerza iónica de la solución. Las condiciones de hibridación y lavado son conocidas y se ejemplifican en Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), particularmente el capítulo 11 y la Tabla 11.1 en la misma; y Sambrook, J. y Russell, W., *Molecular Cloning: A Laboratory Manual*, Third Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (2001). Las condiciones de temperatura y fuerza iónica determinan la "rigurosidad" de la hibridación. Las condiciones de rigurosidad se pueden ajustar para cribar fragmentos moderadamente similares, tales como secuencias homólogas de organismos relacionados de forma distante, hasta fragmentos altamente similares, tales como genes que duplican enzimas funcionales de organismos estrechamente relacionados. Las condiciones de hibridación y los lavados de post-hibridación son útiles para obtener las condiciones de rigurosidad determinadas deseadas de la hibridación. Un grupo de lavados post-hibridación ilustrativos es una serie de lavados que empiezan con 6 x SSC (donde SSC es NaCl 0,15 M y tampón citrato 15 mM), SDS al 0,5% SDS a temperatura ambiente durante 15 minutos, a continuación se repite con 2 x SSC, SDS al 0,5% a 45°C durante 30 minutos, y a continuación se repite dos veces con 0,2 x SSC, SDS al 0,5% a 50°C durante 30 minutos. Otras condiciones de rigurosidad se obtienen utilizando temperaturas más elevadas en las que los lavados son idénticos a los anteriores a excepción de la que la temperatura del final de dos lavados de 30 minutos en 0,2 x SSC, SDS al 0,5%, se incrementa hasta 60°C. Otro grupo de condiciones de rigurosidad elevada utiliza dos lavados finales en 0,1 x SSC, SDS al 0,1% a 65°C. Otro ejemplo de condiciones de hibridación astringentes es la hibridación a 50°C o superior y 0,1xSSC (cloruro sódico 15 mM/citrato sódico 1,5 mM). Otro ejemplo de condiciones de hibridación rigurosas es incubación durante la noche a 42°C en una solución: formamida al 50%, 5 X SSC (NaCl 150 mM, citrato de trisodio 15 mM), fosfato de sodio 50 mM (pH 7,6), 5 x solución de Denhardt, sulfato de dextrano al 10%, y 20 µg/ml de ADN de esperma de salmón roto desnaturalizado, seguido del lavado de los filtros en 0,1 x SSC a aproximadamente 65°C. Las condiciones de hibridación rigurosas y las condiciones de lavado post-hibridación son condiciones de hibridación y condiciones de lavado de post-hibridación que son por lo menos tan rigurosas como las condiciones representativas anteriores.

**[0037]** La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, aunque dependiendo de la rigurosidad de la hibridación, son posibles desapareamientos entre bases. La rigurosidad apropiada para hibridar ácidos nucleicos depende de la longitud de los ácidos nucleicos y el grado de complementación, variables conocidas en la técnica. Cuanto mayor es el grado de similitud u homología entre dos secuencias de nucleótidos, mayor es el valor de la temperatura de fusión (T<sub>f</sub>) para híbridos de ácidos nucleicos que tienen aquellas secuencias. La estabilidad relativa (correspondiente a una T<sub>f</sub> superior) de

5 hibridaciones de ácido nucleico disminuye en el siguiente orden: ARN:ARN, ADN:ARN, ADN:ADN. Para híbridos superiores a 100 nucleótidos de longitud, se han derivado ecuaciones para calcular Tf (véase Sambrook et al., supra, 9.50-9.51). Para hibridaciones con ácidos nucleicos más cortos, es decir, oligonucleótidos, la posición de los desapareamientos se vuelve más importante, y la longitud de los oligonucleótidos determina su especificidad (véase Sambrook et al., supra, 11.7-11.8). Habitualmente, la longitud de un ácido nucleico hibridable es por lo menos aproximadamente de 10 nucleótidos. Las longitudes mínimas ilustrativas para un ácido nucleico hibridable son: por lo menos aproximadamente de 15 nucleótidos; por lo menos aproximadamente de 20 nucleótidos; y por lo menos aproximadamente de 30 nucleótidos. Además, el experto en la materia reconocerá que la temperatura y la concentración de sal de la solución de lavado se pueden ajustar según sea necesario según los factores, tales como la longitud de la sonda.

15 **[0038]** El término "sustitución conservativa de aminoácidos" se refiere a la intercambiabilidad de proteínas de residuos de aminoácidos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas que consiste en glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales ahidroxil-alifáticas que consiste en serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida que consisten en asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas que consisten en fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas que consisten en lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre que consisten en cisteína y metionina. Los grupos de sustitución conservativa de aminoácidos de ejemplo son valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina, y asparagina-glutamina.

25 **[0039]** Los "ácidos nucleicos sintéticos" se pueden ensamblar a partir de bloques de construcción de oligonucleótidos que se sintetizan químicamente utilizando procedimientos conocidos por los expertos en la materia. Estos bloques de construcción se ligan e hibridan para formar segmentos génicos que a continuación se ensamblan enzimáticamente para construir el gen completo. "Sintetizado químicamente", en relación con una secuencia de ADN, significa que los nucleótidos del componente se ensamblaron in vitro. La síntesis química manual de ADN se puede realizar utilizando procedimientos bien establecidos, o se puede llevar a cabo una síntesis química automatizada utilizando uno de un conjunto de máquinas disponibles comercialmente. La secuencia de nucleótidos de los ácidos nucleicos se puede modificar para una expresión óptima en base a la optimización de la secuencia de nucleótidos para reflejar la tendencia del codón de la célula huésped. El experto en la materia valora la probabilidad de una expresión satisfactoria si la utilización de un codón es propensa a aquellos codones favorecidos por el huésped. La determinación de codones preferidos se puede basar en un reconocimiento de genes derivados de la célula huésped en la que está disponible la información de la secuencia.

40 **[0040]** Un polinucleótido o polipéptido tiene un cierto porcentaje de "identidad en la secuencia" con otro polinucleótido o polipéptido, lo que significa que, cuando se alinean, ese porcentaje de bases o aminoácidos es el mismo, y en la misma posición relativa, cuando se comparan las dos secuencias. La similitud de secuencias se puede determinar de diferentes maneras. Para determinar la identidad de secuencia, las secuencias se pueden alinear utilizando los métodos programas informáticos, incluyendo BLAST, disponibles en la página web [ncbi.nlm.nih.gov/BLAST](http://ncbi.nlm.nih.gov/BLAST). Véase, por ejemplo, Altschul et al. (1990), J. Mol. Biol. 215:403-10. Otro algoritmo de alineación es FASTA, disponible en el paquete Genetics Computing Group (GCG), de Madison, Wisconsin, Estados Unidos, una subsidiaria totalmente propiedad de Oxford Molecular Group, Inc. Otras técnicas para la alineación se describen en *Methods in Enzymology*, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., una división de Harcourt Brace & Co., San Diego, California, Estados Unidos. De particular interés son los programas de alineación que permiten espacios en la secuencia. El algoritmo Smith-Waterman es un tipo de algoritmo que permite espacios en las alineaciones de secuencia. Véase *Meth. Mol. Biol.* 70: 173-187 (1997). Además, el programa GAP que utiliza el método de alineación de Needleman y Wunsch se puede utilizar para alinear secuencias. Véase *J. Mol. Biol.* 48: 443-453 (1970).

55 **[0041]** Antes de describir la presente invención con más detalle, debe entenderse que la presente invención no se limita a realizaciones particulares descritas, y, por tanto, naturalmente, puede variar. Debe entenderse también que la terminología utilizada aquí es para el objetivo de describir realizaciones particulares únicamente y no pretende ser limitante, ya que el alcance de la presente invención se limitará sólo por las reivindicaciones adjuntas.

60 **[0042]** Cuando se proporciona un intervalo de valores, debe entenderse que cada valor que interviene, hasta la décima parte de la unidad del límite inferior, a menos que el contexto dictamine claramente lo contrario, entre el límite superior y el límite inferior de ese intervalo y cualquier otro valor indicado o que interviene en este intervalo indicado, se encuentra comprendido en la invención. Los límites superior e inferior de estos intervalos más pequeños pueden estar incluidos independientemente en los intervalos más pequeños y también están comprendidos en la invención, sujetos a cualquier límite específicamente excluido en el intervalo indicado. Cuando el intervalo indicado incluye uno o ambos límites, los intervalos que excluyen alguno o ambos de esos límites incluidos también están incluidos en la invención.

[0043] A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados aquí tienen el mismo significado que el entendido habitualmente por un experto en la materia a la que pertenece la invención. Aunque se puede utilizar en la práctica o evaluación de la presente invención cualquier método y material similar a los descritos aquí, los métodos y materiales preferidos se describen a continuación.

[0044] Debe indicarse que, tal como se utiliza aquí y en las reivindicaciones adjuntas, las formas singulares "un/una" y "el/la" incluyen los referentes plurales a menos que el contexto dictamine claramente lo contrario. De este modo, por ejemplo, la referencia a "una célula huésped modificada genéticamente" incluye una pluralidad de dichas células huésped y la referencia a "el compuesto isoprenoide" incluye la referencia a uno o más compuestos isoprenoides y equivalentes de los mismos conocidos por los expertos en la materia, y así sucesivamente. Debe indicarse también que las reivindicaciones se pueden redactar para excluir cualquier elemento opcional. Por tanto, esta afirmación pretende servir como base antecedente para utilizar dicha terminología exclusiva como "solamente", "sólo" y similares en relación con la redacción de los elementos de la reivindicación o utilización de una limitación "negativa".

[0045] Las publicaciones aquí descritas se proporcionan solamente para su divulgación antes de la fecha de presentación de la presente solicitud. Nada en el presente documento debe interpretarse como una admisión de que la presente invención no presenta el derecho de prefechar dicha publicación debido a la invención anterior. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que pueden ser necesarias de confirmar independientemente.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

[0046] La presente invención proporciona células huésped de *S. cerevisiae* modificadas genéticamente que producen precursores de isoprenoides o compuestos isoprenoides según la reivindicación 1. Se proporcionan métodos para la producción de un compuesto isoprenoide o un precursor isoprenoide en una célula huésped modificada genéticamente en cuestión según la reivindicación 3. Los métodos implican generalmente cultivar una célula huésped modificada genéticamente en cuestión bajo condiciones que promueven la producción de niveles elevados de un compuesto isoprenoide o un compuesto precursor de isoprenoide.

[0047] Los mecanismos del mevalonato y esterol de *S. cerevisiae* se representan esquemáticamente en la figura 1 y la figura 2 (cabe indicar que la amorfadieno sintasa (ADS) en la figura 2 no se expresa normalmente en *S. cerevisiae* no modificada genéticamente). Este mecanismo es típico de una amplia variedad de células eucariotas. FPP se convierte en escualeno mediante la escualeno sintasa (ERG9). El escualeno se convierte en ergosterol en las etapas posteriores. En células no modificadas, mucho del flujo metabólico dirige FPP hacia la síntesis de esterol. En una célula huésped modificada genéticamente en cuestión, el flujo metabólico se redirige hacia una mayor producción de los precursores de isoprenoide IPP y FPP.

#### CÉLULAS HUÉSPED MODIFICADAS GENÉTICAMENTE

[0048] La presente invención proporciona células huésped de *S. cerevisiae* modificadas genéticamente según la reivindicación 1, modificadas según la presente invención, una célula huésped modificada genéticamente en cuestión muestra las siguientes características: niveles de actividad incrementada de una o más enzimas del mecanismo de mevalonato; niveles incrementados de la actividad de prenil transferasa; y niveles disminuidos de la actividad de escualeno sintasa.

[0049] Los niveles de actividad incrementados de una o más enzimas del mecanismo de mevalonato, los niveles incrementados de la actividad de prenil transferasa, y los niveles disminuidos de actividad de escualeno sintasa incrementan la producción de isoprenoides o precursores de isoprenoides por una célula huésped modificada genéticamente en cuestión. De este modo, se describe aquí una célula huésped modificada genéticamente que muestra incrementos en la producción de isoprenoides o precursores de isoprenoides, donde la producción de isoprenoides o precursores de isoprenoides se incrementa en por lo menos aproximadamente 50%, por lo menos aproximadamente 60%, por lo menos aproximadamente 70%, por lo menos aproximadamente 80%, por lo menos aproximadamente 90%, por lo menos aproximadamente 2 veces, por lo menos aproximadamente 2,5 veces, por lo menos aproximadamente 5 veces, por lo menos aproximadamente 10 veces, por lo menos aproximadamente 20 veces, por lo menos aproximadamente 30 veces, por lo menos aproximadamente 40 veces, por lo menos aproximadamente 50 veces, por lo menos aproximadamente 75 veces, por lo menos aproximadamente 100 veces, por lo menos aproximadamente 200 veces, por lo menos aproximadamente 300 veces, por lo menos aproximadamente 400 veces, por lo menos aproximadamente 500 veces, o por lo menos aproximadamente  $10^3$  veces, o más, en la célula huésped modificada genéticamente, en comparación con el nivel de compuesto precursor de isoprenoide o compuesto isoprenoide producido en una célula huésped de control que no está modificada genéticamente tal como se describe aquí. La producción de isoprenoides o precursores de isoprenoides se determina fácilmente utilizando métodos conocidos, por ejemplo, cromatografía de gases-espectrometría de masas, cromatografía líquida-espectrometría de masas, cromatografía de iones-espectrometría de masas, detección amperométrica pulsada, espectrometría UV-VIS y similares.



**[0050]** Se describe aquí una célula huésped modificada genéticamente que proporciona una mayor producción de isoprenoide o precursor de isoprenoide por célula, por ejemplo, la cantidad de compuesto isoprenoide o precursor de isoprenoide producida utilizando el método en cuestión es por lo menos aproximadamente 10%, por lo menos aproximadamente 15%, por lo menos aproximadamente 20%, por lo menos aproximadamente 25%, por lo menos aproximadamente 30%, por lo menos aproximadamente 35%, por lo menos aproximadamente 40%, por lo menos aproximadamente 45%, por lo menos aproximadamente 50%, por lo menos aproximadamente 60%, por lo menos aproximadamente 70%, por lo menos aproximadamente 80%, por lo menos aproximadamente 90%, por lo menos aproximadamente 2 veces, por lo menos aproximadamente 2,5 veces, por lo menos aproximadamente 5 veces, por lo menos aproximadamente 10 veces, por lo menos aproximadamente 20 veces, por lo menos aproximadamente 30 veces, por lo menos aproximadamente 40 veces, por lo menos aproximadamente 50 veces, por lo menos aproximadamente 75 veces, por lo menos aproximadamente 100 veces, por lo menos aproximadamente 200 veces, por lo menos aproximadamente 300 veces, por lo menos aproximadamente 400 veces, o por lo menos aproximadamente 500 veces, o  $10^3$  veces, o más, más elevada que la cantidad de compuesto isoprenoide o precursor de isoprenoide producida por una célula huésped que no se ha modificado genéticamente por los métodos en cuestión, en base a una célula. La cantidad de células se mide midiendo el peso celular en seco o midiendo la densidad óptica del cultivo celular.

**[0051]** Se describe aquí una célula huésped modificada genéticamente que proporciona una mayor producción de isoprenoide o precursor de isoprenoide por volumen unitario del cultivo celular, por ejemplo, la cantidad de compuesto isoprenoide o precursor de isoprenoide producida utilizando una célula huésped modificada genéticamente en cuestión es, por lo menos aproximadamente 50%, por lo menos aproximadamente 60%, por lo menos aproximadamente 70%, por lo menos aproximadamente 80%, por lo menos aproximadamente 90%, por lo menos aproximadamente 2 veces, por lo menos aproximadamente 2,5 veces, por lo menos aproximadamente 5 veces, por lo menos aproximadamente 10 veces, por lo menos aproximadamente 20 veces, por lo menos aproximadamente 30 veces, por lo menos aproximadamente 40 veces, por lo menos aproximadamente 50 veces, por lo menos aproximadamente 75 veces, por lo menos aproximadamente 100 veces, por lo menos aproximadamente 200 veces, por lo menos aproximadamente 300 veces, por lo menos aproximadamente 400 veces, o por lo menos aproximadamente 500 veces, o  $10^3$  veces, o más, más elevada que la cantidad de compuesto isoprenoide o precursor de isoprenoide producida por una célula huésped que no se ha modificado genéticamente por los métodos en cuestión, en base a un volumen unitario del cultivo celular.

**[0052]** En algunas realizaciones, un huésped modificado genéticamente en cuestión produce un compuesto isoprenoide o precursor de isoprenoide en una cantidad que varía desde 1  $\mu\text{g}$  de compuesto isoprenoide/ml hasta 100.000  $\mu\text{g}$  compuesto isoprenoide /ml, por ejemplo, desde aproximadamente 1  $\mu\text{g}/\text{ml}$  hasta aproximadamente 10.000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 5000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 4500  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 4000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 3500  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 3000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 2500  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 2000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 1500  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 1  $\mu\text{g}/\text{ml}$  a 1000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 5  $\mu\text{g}/\text{ml}$  a 5000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 10  $\mu\text{g}/\text{ml}$  a 5000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 20  $\mu\text{g}/\text{ml}$  a 5000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 30  $\mu\text{g}/\text{ml}$  a 1000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 40  $\mu\text{g}/\text{ml}$  a 500  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 50  $\mu\text{g}/\text{ml}$  a 300  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 60  $\mu\text{g}/\text{ml}$  a 100  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, 70  $\mu\text{g}/\text{ml}$  a 80  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, desde aproximadamente 1  $\mu\text{g}/\text{ml}$  hasta aproximadamente 1.000  $\mu\text{g}/\text{ml}$ , desde aproximadamente 1.000  $\mu\text{g}/\text{ml}$  hasta aproximadamente 2.000  $\mu\text{g}/\text{ml}$ , desde aproximadamente 2.000  $\mu\text{g}/\text{ml}$  hasta aproximadamente 3.000  $\mu\text{g}/\text{ml}$ , desde aproximadamente 3.000  $\mu\text{g}/\text{ml}$  hasta aproximadamente 4.000  $\mu\text{g}/\text{ml}$ , desde aproximadamente 4.000  $\mu\text{g}/\text{ml}$  hasta aproximadamente 5.000  $\mu\text{g}/\text{ml}$ , desde aproximadamente 5.000  $\mu\text{g}/\text{ml}$  hasta aproximadamente 7.500  $\mu\text{g}/\text{ml}$ , o desde aproximadamente 7.500  $\mu\text{g}/\text{ml}$  hasta aproximadamente 10.000  $\mu\text{g}/\text{ml}$ , o más de 10.000  $\mu\text{g}/\text{ml}$  de compuesto isoprenoide, por ejemplo, desde aproximadamente 10 mg de compuesto isoprenoide/ml hasta aproximadamente 20 mg de compuesto isoprenoide/ml, desde aproximadamente 20 mg de compuesto isoprenoide/ml hasta aproximadamente 50 mg de compuesto isoprenoide/ml, desde aproximadamente 50 mg de compuesto isoprenoide/ml hasta aproximadamente 100 mg de compuesto isoprenoide/ml, o más.

**[0053]** En una realización de ejemplo, el mecanismo metabólico de *Saccharomyces cerevisiae* se modifica para producir sesquiterpenos a partir de farnesil difosfato. Uno de dichos sesquiterpenos, amorfadieno, es un precursor del fármaco contra la malaria artemisina. El amorfadieno, ciclado a partir de farnesil difosfato, se puede utilizar como un ensayo para niveles de precursores de isoprenoides.

**[0054]** En una realización de ejemplo, se incrementan los niveles de actividad de HMGR, una prenil transferasa, Ecm22p y Upc2p y se disminuyen los niveles de actividad de escualeno sintasa. La 3-hidroxi-3-metilglutaril coenzima-A reductasa (HMGR) y una prenil transferasa, por ejemplo, la farnesil difosfato sintasa (FPPS), catalizan reacciones de cuello de botella en un mecanismo de síntesis del amorfadieno. El incremento de la actividad de HMGR y una prenil transferasa, FPPS, superan estos cuellos de botella. En la regulación de la síntesis de esterol son importantes dos factores de transcripción, Ecm22p y Upc2p. Cada uno de estos dos factores se muta en un único aminoácido próximo a sus extremos C-terminales, cuya mutación incrementa la

actividad de cada factor. La escualeno sintasa cataliza la reacción de farnesil difosfato a escualeno en el mecanismo de síntesis de esteroles no deseado. De este modo, para maximizar los grupos de precursores y evitar el flujo excesivo a esteroides, se ha limitado la transcripción de ERG9.

5 Nivel de actividad incrementado de una o más enzimas del mecanismo del mevalonato

10 **[0055]** El mecanismo del mevalonato comprende enzimas que catalizan las siguientes etapas: (a) condensar dos moléculas de acetil-CoA en acetoacetil-CoA, habitualmente mediante la acción de acetoacetil-CoA tiolasa; (b) condensar acetoacetil-CoA con acetil-CoA para formar HMGCoA, habitualmente mediante la acción de HMG sintasa (HMGS); (c) convertir HMG-CoA en mevalonato, habitualmente mediante la acción de HMGR; (d) fosforilar mevalonato en mevalonato 5-fosfato, habitualmente mediante la acción de mevalonato quinasa (MK); (e) convertir mevalonato 5-fosfato en mevalonato 5-pirofosfato, habitualmente mediante la acción de fosfomevalonato quinasa (PMK); y (f) convertir mevalonato 5-pirofosfato en isopentenil pirofosfato, habitualmente mediante la acción de mevalonato pirofosfato descarboxilasa (MPD).

15 **[0056]** También se describe aquí una célula huésped eucariota modificada genéticamente que comprende una o más modificaciones genéticas que dan lugar a uno o más de los siguientes: nivel incrementado de la actividad de HMGS; nivel incrementado de la actividad de HMGR; nivel incrementado de la actividad de MK; nivel incrementado de la actividad de PMK; y nivel incrementado de la actividad de MPD.

20 **[0057]** También se describe aquí una célula huésped modificada genéticamente que es modificada genéticamente de manera que se incrementa el nivel de actividad de una o más de las enzimas del mecanismo de mevalonato. El nivel de actividad de una o más enzimas del mecanismo de mevalonato en una célula huésped modificada genéticamente en cuestión se puede incrementar de varias maneras, que incluyen, pero sin limitación, 1) incrementar la fuerza del promotor del promotor al que está unido operativamente la región codificante de la enzima del mecanismo del mevalonato; 2) incrementar el número de copias del plásmido que comprende una secuencia de nucleótidos que codifica la enzima del mecanismo del mevalonato; 3) incrementar la estabilidad de un ARNm de la enzima del mecanismo del mevalonato (donde un "ARNm de la enzima del mecanismo del mevalonato" es un ARNm que comprende una secuencia de nucleótidos que codifica la enzima del mecanismo del mevalonato); 4) modificar la secuencia del sitio de unión a ribosoma de un ARN de la enzima del mecanismo de mevalonato, de manera que se incrementa el nivel de traducción del ARNm de la enzima del mecanismo del mevalonato; 5) modificar la secuencia entre el sitio de unión a ribosoma de un ARNm de la enzima del mecanismo del mevalonato y el codón de inicio de la secuencia codificante de la enzima del mecanismo del mevalonato, de manera que se incrementa el nivel de traducción del ARNm de la enzima del mecanismo del mevalonato; 6) modificar la región 5' intercistronica completa del codón de inicio de la región codificante de la enzima del mecanismo del mevalonato, de manera que se incrementa la traducción del ARNm de la enzima del mecanismo del mevalonato; 7) modificar la utilización de codones de la enzima del mecanismo de mevalonato, de manera que se incrementa el nivel de traducción del ARNm de la enzima del mecanismo del mevalonato; 8) expresar ARNt de codones raros utilizados en la enzima del mecanismo del mevalonato, de manera que se incrementa el nivel de traducción del ARNm de la enzima del mecanismo del mevalonato; 9) incrementar la estabilidad enzimática de la enzima del mecanismo del mevalonato; 10) incrementar la actividad específica (unidades de actividad por unidad de proteína) de la enzima del mecanismo del mevalonato; 11) expresar una forma modificada de una enzima del mecanismo del mevalonato, se manera que la enzima modificada muestra una solubilidad incrementada en la célula huésped; o 12) expresar una forma modificada de una enzima del mecanismo del mevalonato, de manera que la enzima modificada carece de un dominio a través del cual tiene lugar la regulación. Las modificaciones anteriores se pueden realizar individualmente o combinadas; por ejemplo, se pueden realizar dos o más de las modificaciones anteriores para proporcionar un nivel incrementado de la actividad de la enzima del mecanismo del mevalonato.

50 **[0058]** La enzima HMG-CoA reductasa (HMGR) cataliza una reacción irreversible que reduce la 3-hidroxi-3-metilglutaril Coenzima A (HMG-CoA) en mevalonato. Esta etapa es la etapa comprometida en el mecanismo de la biosíntesis de esteroles. De este modo, la HMGR es un punto principal de regulación en organismos que utilizan de forma natural el mecanismo del mevalonato para producir isoprenoides.

55 **[0059]** También se describe aquí una célula huésped modificada genéticamente en cuestión que es modificada genéticamente de manera que se incrementa el nivel de actividad de HMGR. El nivel de actividad de HMGR en la célula huésped modificada genéticamente se puede incrementar de varias maneras, que incluyen, pero sin limitación, 1) incrementar la fuerza del promotor del promotor al que está unido operativamente la región codificante de HMGR; 2) incrementar el número de copias del plásmido que comprende una secuencia de nucleótidos que codifica HMGR; 3) incrementar la estabilidad de un ARNm de HMGR (donde un "ARNm de HMGR" es un ARNm que comprende una secuencia de nucleótidos que codifica HMGR); 4) modificar la secuencia del sitio de unión a ribosoma de un ARNm de HMGR, de manera que se incrementa el nivel de traducción del ARNm de HMGR; 5) modificar la secuencia entre el sitio de unión a ribosoma de un ARNm de HMGR y el codón de inicio de la secuencia codificante de HMGR, de manera que se incrementa el nivel de traducción del ARNm de HMGR; 6) modificar la región 5' intercistronica completa del codón de inicio de la región codificante de HMGR, de manera que se incrementa la traducción del ARNm de HMGR; 7) modificar la utilización

de codones de HMGR, de manera que se incrementa el nivel de traducción del ARNm de HMGR, 8) expresar ARNt de codones raros utilizados en HMGR, de manera que se incrementa el nivel de traducción del ARNm de HMGR; 9) incrementar la estabilidad enzimática de HMGR; 10) incrementar la actividad específica (unidades de actividad por unidad de proteína) de HMGR; u 11) truncar la HMGR para eliminar un elemento regulador negativo. Las modificaciones anteriores se pueden realizar individualmente o combinadas; por ejemplo, se pueden realizar dos o más de las modificaciones anteriores para proporcionar un nivel incrementado de la actividad de HMGR.

**[0060]** El nivel de HMGR se incrementa mediante la modificación genética de una célula huésped, de manera que produce una forma truncada de HMGR (tHMGR), cuya forma truncada presenta una actividad enzimática incrementada en relación a HMGR de tipo salvaje. tHMGR carece de un dominio que comprende la membrana y, por tanto, es soluble y carece de la retroinhibición de HMGR. tHMGR mantiene su región C-terminal catalítica y, de este modo, mantiene la actividad de HMGR. En algunas realizaciones, la HMGR truncada presenta la secuencia de aminoácidos representada en las figuras 7A y 7B (SEQ ID NO:2). En algunas realizaciones, la HMGR truncada es codificada por un ácido nucleico que comprende la secuencia de nucleótidos representada en la figura 6 (SEQ ID NO:1).

**[0061]** En algunas células huésped aquí descritas, se incrementa el nivel de actividad de una o más de HMGS, MK, y PMK. En *S. cerevisiae*, los genes que codifican HMGS (*ERG13*), MK (*ERG12*), y PMK (*ERG8*) comprende un elemento regulador de esterol que se une a los factores de transcripción Ecm22p y Upc2p, donde, tras la unión de Ecm22p y Upc2p, se activa la transcripción. En algunas realizaciones, el nivel de actividad de una o más de HMGS, MK, y PMK se incrementa aumentando la actividad de Ecm22p y Upc2p. Vik et al. (2001) Mol. Cell. Biol. 19:6395-405.

**[0062]** Normalmente, *S. cerevisiae* no capta esteroides del medio bajo condiciones aeróbicas. Lewis et al. ((1988) Yeast 4:93-106) aislaron un mutante de levadura, *upc2-1* (captación de control), que dio lugar a una captación aeróbica de esterol. El alelo *upc2-1* comprende una transición de guanina a adenina en el marco de lectura abierto designado como YDR213 W en el cromosoma IV. Crowley et al. (1998) J. Bacteriol. 16: 4177-4183. La secuencia de ácidos nucleicos de UPC2 de tipo salvaje es conocida y se puede obtener a través del GenBank No. de acceso Z68194. Este alelo de tipo salvaje está indicado como las coordenadas 889746-892487 en el cromosoma de *S. cerevisiae*. Tal como fue descubierto previamente por Lewis et al., bajo condiciones nativas, el nivel de captación de esterol fue de 10 a 20 veces mayor que con el tipo salvaje isogénico. El mutante dio lugar a una producción incrementada de ergosterol.

**[0063]** Se ha observado que un único cambio de aminoácido próximo al extremo c-terminal de los factores de transcripción Upc2p y Ecm22p incrementa su actividad. En muchas realizaciones, una célula huésped modificada genéticamente en cuestión es modificada genéticamente de manera que Upc2p comprende una sustitución de glicina a ácido aspártico en el aminoácido 888; y Ecm22p comprende una sustitución de glicina en ácido aspártico en el aminoácido 790.

Nivel incrementado de la actividad de preniltransferasa

**[0064]** También se describe aquí una célula huésped eucariota modificada genéticamente que es modificada genéticamente, de manera que se incrementa el nivel de actividad de geranil difosfato sintasa (GPPS) y/o farnesil difosfato sintasa (FPPS).

**[0065]** La enzima farnesil difosfato sintasa (FPPS) cataliza una reacción que convierte el geranil difosfato (GPP) en farnesil difosfato (FPP). También se ha observado que esta etapa es limitante de la velocidad en el mecanismo del mevalonato. De este modo, la FPPS es un punto de regulación en organismos que utilizan de manera natural el mecanismo del mevalonato para producir isoprenoides. Por tanto, y para facilitar la descripción posterior, la modulación de los niveles de actividad de una prenil transferasa se discute en términos de modulación del nivel de actividad de una FPPS.

**[0066]** El nivel de actividad de FPPS en una célula huésped modificada genéticamente se puede incrementar de varias maneras, que incluyen, pero sin limitación, 1) incrementar la fuerza del promotor al que está unido operativamente la región codificante de FPPS; 2) incrementar el número de copias del plásmido que comprende una secuencia de nucleótidos que codifica FPPS; 3) incrementar la estabilidad de un ARNm de FPPS (donde un "ARNm de FPPS" es un ARNm que comprende una secuencia de nucleótidos que codifica FPPS); 4) modificar la secuencia del sitio de unión a ribosoma de un ARNm de FPPS, de manera que se incrementa el nivel de traducción del ARNm de FPPS; 5) modificar la secuencia entre el sitio de unión a ribosoma de un ARNm de FPPS y el codón de inicio de la secuencia codificante de FPPS, de manera que se incrementa el nivel de traducción del ARNm de FPPS; 6) modificar la región 5' intercistronica completa del codón de inicio de la región codificante de FPPS, de manera que se incrementa la traducción del ARNm de FPPS; 7) modificar la utilización de codones de FPPS, de manera que se incrementa el nivel de traducción del ARNm de FPPS; 8) expresar ARNt de codones raros utilizados en FPPS, de manera que se incrementa el nivel de traducción del ARNm de FPPS; 9) incrementar la estabilidad enzimática de FPPS; o 10) incrementar la actividad específica (unidades de

actividad por unidad de proteína) de FPPS. Las modificaciones anteriores se pueden realizar individualmente o combinadas; por ejemplo, se pueden realizar dos o más de las modificaciones anteriores para proporcionar un nivel incrementado de la actividad de FPPS.

##### 5 Nivel disminuido de actividad de escualeno sintasa

[0067] La enzima escualeno sintasa cataliza una reacción que convierte farnesil difosfato en escualeno. Esta etapa es la primera etapa en el mecanismo que lleva de farnesil difosfato a ergosterol. De este modo, al limitar la acción de esta enzima, el FPP se desvía hacia mecanismos de producción de terpenoides utilizando, por ejemplo, terpeno sintasas o GGPP sintasa y posteriores terpeno sintasas.

[0068] Una célula huésped modificada genéticamente en cuestión es modificada genéticamente, de manera que se disminuye el nivel de actividad de escualeno sintasa. El nivel de actividad de escualeno sintasa en la célula huésped modificada genéticamente se puede disminuir de varias maneras, que incluyen, pero sin limitación, 1) disminuir la fuerza del promotor del promotor al que está unido operativamente la región codificante de escualeno sintasa; 2) disminuir la estabilidad de un ARNm de escualeno sintasa (donde un "ARNm de escualeno sintasa" es un ARNm que comprende una secuencia de nucleótidos que codifica escualeno sintasa); 3) modificar la secuencia del sitio de unión a ribosoma de un ARNm de escualeno sintasa, de manera que se disminuye el nivel de traducción del ARNm de escualeno sintasa; 4) modificar la secuencia entre el sitio de unión a ribosoma de un ARNm de escualeno sintasa y el codón de inicio de la secuencia codificante de escualeno sintasa, de manera que se disminuye el nivel de traducción del ARNm de escualeno sintasa; 5) modificar la región 5' intercistronica completa del codón de inicio de la región codificante de escualeno sintasa, de manera que se disminuye la traducción del ARNm de escualeno sintasa; 6) modificar la utilización de codones de escualeno sintasa, de manera que se disminuye el nivel de traducción del ARNm de escualeno sintasa, 7) disminuir la estabilidad enzimática de escualeno sintasa; 8) disminuir la actividad específica (unidades de actividad por unidad de proteína) de escualeno sintasa, o 9) utilizar un promotor reprimible químicamente y reprimir el promotor reprimible químicamente mediante la adición de un producto químico a un medio de crecimiento. Las modificaciones anteriores se pueden realizar individualmente o combinadas; por ejemplo, se pueden realizar dos o más de las modificaciones anteriores para proporcionar un nivel disminuido de la actividad de escualeno sintasa.

[0069] En una realización de ejemplo, la actividad de la escualeno sintasa en *S. cerevisiae* se ha reducido o eliminado. Se han producido mutantes de *ERG9* de levadura que no son capaces de convertir el mevalonato en escualeno. Véase, por ejemplo, Karst et al. (1977) Molec. Gen. Genet. 154:269-277; patente de Estados Unidos No. 5,589,372; y publicación de la patente de Estados Unidos No. 2004/0110257. Las modificaciones genéticas incluyen la disminución de la actividad de la escualeno sintasa mediante el bloqueo o la reducción de la producción de escualeno sintasa, reduciendo la actividad de la escualeno sintasa o mediante la inhibición de la actividad de escualeno sintasa. El bloqueo o la reducción de la producción de escualeno sintasa pueden incluir situar el gen de escualeno sintasa bajo el control de un promotor que requiere la presencia de un compuesto inductor en el medio de crecimiento. Mediante el establecimiento de condiciones tales que el inductor se agota en el medio, se puede desactivar la expresión de escualeno sintasa. Algunos promotores se desactivan mediante la presencia de un compuesto represor. Por ejemplo, los promotores de los genes de *CTR3* o *CTR1* de levadura se pueden reprimir mediante la adición de cobre. El bloqueo o la reducción de la actividad de la escualeno sintasa puede incluir tecnología de escisión similar a la descrita en la patente de Estados Unidos No. 4,743,546. En esta estrategia, el gen *ERG9* se clona entre secuencias genéticas específicas que permiten la escisión controlada específica del gen *ERG9* del genoma. La escisión podría estar provocada por, por ejemplo, un desplazamiento en la temperatura de cultivo del cultivo, como en la patente de Estados Unidos No. 4,743,546, o por alguna otra señal física o nutricional. Dicha modificación genética incluye cualquier tipo de modificación y específicamente incluye modificaciones realizadas mediante tecnología recombinante y mediante mutagénesis clásica. Los inhibidores de escualeno sintasa son conocidos (véase la patente de Estados Unidos No. 4,871,721 y las referencias citadas en la patente de Estados Unidos No. 5,475,029) y se pueden añadir a los cultivos celulares.

[0070] En algunas realizaciones, se modifica el uso de codones de una secuencia codificante de escualeno sintasa, de manera que se disminuye el nivel de traducción del ARNm de *ERG9*. La reducción del nivel de traducción de ARNm de *ERG9* por modificación del uso de codones se consigue mediante la modificación de la secuencia para incluir codones que son raros o que no son comúnmente utilizados por la célula huésped. Están disponibles tablas de uso de codones para muchos organismos que resumen el porcentaje de tiempo que un organismo específico utiliza un codón específico para codificar un aminoácido. Ciertos codones se utilizan más frecuentemente que otros, los codones "raros". El uso de codones "raros" en una secuencia disminuye generalmente su índice de traducción. De este modo, por ejemplo, se modifica la secuencia codificante mediante la introducción de uno o más codones raros, que afectan el índice de traducción, pero no la secuencia de aminoácidos de la enzima traducida. Por ejemplo, hay 6 codones que codifican para arginina: CGT, CGC, CGA, CGG, AGA, and AGG. En *E. coli* se utilizan los codones CGT y CGC mucho más frecuentemente (codificando aproximadamente el 40% de las argininas en *E. coli* cada uno) que el codón AGG (que codifica aproximadamente el 2% de las argininas en *E. coli*). Modificando un codón CGT en la secuencia de un gen en un

codón AGG no cambiaría la secuencia de la enzima, pero probablemente disminuiría el índice de traducción del gen.

#### Generación de una célula huésped modificada genéticamente

5 **[0071]** Una célula huésped modificada genéticamente en cuestión se genera utilizando métodos estándar bien conocidos para los expertos en la materia. En algunos métodos aquí descritos, se introduce un ácido nucleico heterólogo que comprende una secuencia de nucleótidos que codifica una variante de una enzima del mecanismo del mevalonato y/o un ácido nucleico heterólogo que comprende una secuencia de nucleótidos que codifica una variante del factor de transcripción que controla la transcripción de una enzima o enzimas del mecanismo del mevalonato en una célula huésped y sustituye todo o parte de un gen endógeno, por ejemplo, a través de recombinación homóloga. En algunos métodos aquí descritos, se introduce un ácido nucleico heterólogo en una célula huésped parental y el ácido nucleico heterólogo se recombina con un ácido nucleico endógeno que codifica una enzima del mecanismo del mevalonato, una preniltransferasa, un factor de transcripción que controla la transcripción de una o más enzimas del mecanismo del mevalonato, o una escualeno sintasa, modificando genéticamente de este modo la célula huésped parental. En algunos métodos aquí descritos, el ácido nucleico heterólogo comprende un promotor que ha incrementado la fuerza del promotor en comparación con el promotor endógeno que controla la transcripción de la preniltransferasa endógena, y la recombinación da lugar a la sustitución del promotor endógeno con el promotor heterólogo. En otros métodos aquí descritos, el ácido nucleico heterólogo comprende una secuencia de nucleótidos que codifica una HMGR truncada que muestra una actividad enzimática incrementada en comparación con la HMGR endógena y la recombinación da lugar a la sustitución de la secuencia codificante de HMGR endógena por la secuencia codificante de la HMGR heteróloga. En algunos métodos aquí descritos, el ácido nucleico heterólogo comprende un promotor que proporciona una transcripción regulada de una secuencia codificante de la escualeno sintasa unida operativamente y la recombinación da lugar a la sustitución del promotor de escualeno sintasa endógeno por el promotor heterólogo.

#### *Modificaciones genéticas adicionales*

30 **[0072]** Se describe aquí una célula huésped modificada genéticamente en cuestión que comprende una o más modificaciones genéticas además de las descritas anteriormente. Por ejemplo, una célula huésped modificada genéticamente en cuestión se modifica genéticamente de forma adicional con una o más secuencias de nucleótidos que comprenden uno o más ácidos nucleicos que codifican una o más de una preniltransferasa (por ejemplo, una preniltransferasa diferente de FPP y GPP); una terpeno sintasa; y similares.

#### Uso de codones

40 **[0073]** En algunas realizaciones, se modifica la secuencia de nucleótidos que codifica un producto génico (por ejemplo, una preniltransferasa, una terpeno sintasa, etc.) de manera que la secuencia de nucleótidos refleja la preferencia de codón para la célula huésped concreta. Por ejemplo, en algunas realizaciones, se modificará la secuencia de nucleótidos para la preferencia del codón de levadura. Ver, por ejemplo, Bennetzen y Hall (1982) J. Biol. Chem. 257(6): 3026-3031.

45 **[0074]** Tal y como se indica anteriormente, en algunas realizaciones, se modifica el uso de codones de una secuencia codificante de escualeno sintasa, de manera que se disminuye el nivel de traducción del ARNm de ERG9. La reducción del nivel de traducción de ARNm de ERG9 por modificación del uso de codones se consigue mediante la modificación de la secuencia para incluir codones que son raros o que no son comúnmente utilizados por la célula huésped. Están disponibles tablas de uso de codones para muchos organismos que resumen el porcentaje de tiempo que un organismo específico utiliza un codón específico para codificar un aminoácido. Ciertos codones se utilizan más frecuentemente que otros, los codones "raros". El uso de codones "raros" en una secuencia disminuye generalmente su índice de traducción. De este modo, por ejemplo, se modifica la secuencia codificante mediante la introducción de uno o más codones raros, que afectan el índice de traducción, pero no la secuencia de aminoácidos de la enzima traducida. Por ejemplo, hay 6 codones que codifican para arginina: CGT, CGC, CGA, CGG, AGA, and AGG. En *E. coli* se utilizan los codones CGT y CGC mucho más frecuentemente (codificando aproximadamente el 40% de las argininas en *E. coli* cada uno) que el codón AGG (que codifica aproximadamente el 2% de las argininas en *E. coli*). Modificando un codón CGT en la secuencia de un gen en un codón AGG no cambiaría la secuencia de la enzima, pero probablemente disminuiría el índice de traducción del gen.

#### Suministro de acetil-CoA incrementado

60 **[0075]** Dado que la acetil-CoA es un reactivo utilizado por acetoacetil-CoA tiolasa y HMGS en el mecanismo de MEV, en algunas células huésped, aumentos en la zona intracelular de acetil-CoA podría conducir a incrementos en isoprenoide y precursores de isoprenoide. Las modificaciones que incrementarían los niveles de acetil-CoA intracelular incluyen, pero sin limitación, modificaciones que disminuirían la actividad total de la lactato deshidrogenasa en la célula, modificaciones que disminuirían la actividad total de acetato quinasa en la célula,

modificaciones que disminuirían la actividad total de alcohol deshidrogenasa en la célula, modificaciones que interrumpirían el ciclo del ácido tricarbóxico, tales como aquellas que disminuirían la actividad total de 2-cetoglutarato deshidrogenasa, o modificaciones que interrumpirían la fosforilación oxidativa, tales como aquellas que disminuirían la actividad total de la (F1F0)H<sup>+</sup>-ATP sintasa, o combinaciones de las mismas.

5

#### Preniltransferasas

[0076] Las preniltransferasas constituyen un amplio grupo de enzimas que catalizan la condensación consecutiva de IPP que da lugar a la formación de prenil difosfatos de varias longitudes de cadena. Las preniltransferasas adecuadas incluyen enzimas que catalizan la condensación de IPP con sustratos de cebadores alílicos para formar compuestos isoprenoides con aproximadamente 5 unidades de isopreno a aproximadamente 6000 unidades de isopreno o más, por ejemplo, 5 unidades de isopreno a aproximadamente 10 unidades de isopreno, de aproximadamente 10 unidades de isopreno a aproximadamente 15 unidades de isopreno, de aproximadamente 15 unidades de isopreno a aproximadamente 20 unidades de isopreno, de aproximadamente 20 unidades de isopreno a aproximadamente 25 unidades de isopreno, de aproximadamente 25 unidades de isopreno a aproximadamente 30 unidades de isopreno, de aproximadamente 30 unidades de isopreno a aproximadamente 40 unidades de isopreno, de aproximadamente 40 unidades de isopreno a aproximadamente 50 unidades de isopreno, de aproximadamente 50 unidades de isopreno a aproximadamente 100 unidades de isopreno, de aproximadamente 100 unidades de isopreno a aproximadamente 250 unidades de isopreno, de aproximadamente 250 unidades de isopreno a aproximadamente 500 unidades de isopreno, de aproximadamente 500 unidades de isopreno a aproximadamente 1000 unidades de isopreno, de aproximadamente 1000 unidades de isopreno a aproximadamente 2000 unidades de isopreno, de aproximadamente 2000 unidades de isopreno a aproximadamente 3000 unidades de isopreno, de aproximadamente 3000 unidades de isopreno a aproximadamente 4000 unidades de isopreno, de aproximadamente 4000 unidades de isopreno a aproximadamente 5000 unidades de isopreno, o de aproximadamente 5000 unidades de isopreno a aproximadamente 6000 unidades de isopreno o más.

[0077] Las preniltransferasas adecuadas incluyen, pero sin limitación, una E-isoprenil difosfato sintasa, incluyendo, pero sin limitación, geranilgeranil difosfato (GGPP) sintasa, hexaprenil difosfato (HexPP) sintasa, heptaprenil difosfato (HepPP) sintasa, octaprenil (OPP) difosfato sintasa, solanesil difosfato (SPP) sintasa, decaprenil difosfato (DPP) sintasa, chicle sintasa, y gutapercha sintasa; y una Z-isoprenil difosfato sintasa, incluyendo, pero sin limitación, nonaprenil difosfato (NPP) sintasa, undecaprenil difosfato (UPP) sintasa, deshidrodoliquil difosfato sintasa, eicosaprenil difosfato sintasa, caucho natural sintasa, y otras Z-isoprenil difosfato sintasas.

35

[0078] Las secuencias de nucleótidos de numerosas preniltransferasas de una variedad de especies son conocidas y se pueden utilizar o modificar para utilizar en la generación de una célula huésped eucariota modificada genéticamente. Las secuencias de nucleótidos que codifican prenil transferasas son conocidas en la técnica. Véase, por ejemplo, el ARNm de farnesil pirofosfato sintetasa humano (GenBank Acceso No. J05262; *Homo sapiens*); gen de farnesil difosfato sintetasa (FPP) (GenBank Acceso No. J05091; *Saccharomyces cerevisiae*); gen de isopentenil difosfato:dimetilalil difosfato isomerasa (J05090; *Saccharomyces cerevisiae*); Wang y Ohnuma (2000) *Biochim. Biophys. Acta* 1529:33-48; patente de Estados Unidos No. 6,645,747; ARNm de farnesil pirofosfato sintetasa 2 (FPS2) / FPP sintetasa 2 / farnesil difosfato sintasa 2 (At4g17190) de *Arabidopsis thaliana* (GenBank Acceso No. NM\_202836); ARNm de geranilgeranil difosfato sintasa (ggpps) de *Ginkgo biloba* (GenBank Acceso No. AY371321); ARNm de geranilgeranil pirofosfato sintasa (GGPP1) / GGPP sintetasa/ farnesiltransferasa (At4g36810) de *Arabidopsis thaliana* (GenBank Acceso No. NM\_119845); gen de *Synechococcus elongatus* para farnesil, geranilgeranil, geranilfarnesil, hexaprenil, heptaprenil difosfato sintasa (Self-HepPS) (GenBank Acceso No. AB016095); etc.

[0079] También se describe aquí una célula huésped que está modificada genéticamente con un ácido nucleico que comprende una preniltransferasa. Por ejemplo, una célula huésped se puede modificar genéticamente con un ácido nucleico que comprende secuencias de nucleótidos que codifican una preniltransferasa seleccionada entre una GGPP sintasa, una GFPP sintasa, una HexPP sintasa, una HepPP sintasa, una OPP sintasa, una SPP sintasa, una DPP sintasa, una NPP sintasa, y una UPP sintasa.

55

#### Terpeno Sintasas

[0080] Las terpeno sintasas catalizan la producción de compuestos isoprenoides a través de una de las reacciones más complejas conocidas en química o biología. En general, las terpeno sintasas son enzimas de tamaño moderado que tienen pesos moleculares de aproximadamente 40 a 100 kD. Como enzima, las terpeno sintasas se pueden clasificar por tener velocidades de recambio de bajas a moderadas acopladas con una especificidad de reacción exquisita y conservación de la quiralidad. El recambio comprende la unión de sustrato a la enzima, el establecimiento de la conformación del sustrato, la conversión del sustrato en producto y la liberación del producto. Las reacciones se pueden realizar in vitro en disolventes acuosos, habitualmente requieren iones magnesio como cofactores y los productos resultantes, que a menudo son altamente

65

hidrofóbicos, se pueden recuperar mediante la separación en un disolvente orgánico. Patente de Estados Unidos No. 6.890.752.

5 [0081] También se describe aquí una célula huésped modificada genéticamente que es modificada genéticamente adicionalmente con un ácido nucleico que comprende una secuencia de nucleótidos que codifican una terpeno sintasa. Un ácido nucleico con el que una célula huésped es modificada genéticamente puede comprender una secuencia de nucleótidos que codifica una terpeno sintasa que difiere en la secuencia de aminoácidos en uno o más aminoácidos de una terpeno sintasa natural u otra terpeno sintasa original, por ejemplo, una variante de terpeno sintasa. Una "terpeno sintasa original" es una terpeno sintasa que sirve como punto de referencia para la comparación. Las variantes de terpeno sintasas incluyen terpeno sintasas consenso y terpeno sintasas híbridas. El ácido nucleico sintético puede comprender una secuencia de nucleótidos que codifica una terpeno sintasa consenso. El ácido nucleico puede comprender una secuencia de nucleótidos que codifica una terpeno sintasa híbrida.

15 [0082] Se puede utilizar un ácido nucleico que comprende una secuencia de nucleótidos que codifica cualquier terpeno sintasa conocida. Entre las terpeno sintasas adecuadas se incluyen, pero sin limitación, amorfa-4,11-dieno sintasa (ADS), beta-cariofileno sintasa, germacreno A sintasa, 8-epicedrol sintasa, valenceno sintasa, (+)-delta-cadineno sintasa, germacreno C sintasa, (E)-beta-farneseno sintasa, Casbeno sintasa, vetispiradieno sintasa, 5-epi-aristoloueno sintasa, Aristoloueno sintasa, beta-cariofileno, alfa-humuleno, (E,E)-alfa-farneseno sintasa, (-)-beta-pineno sintasa, Gamma-terpineno sintasa, limoneno ciclasa, Linalool sintasa, 1,8-cineol sintasa, (+)-sabineno sintasa, E-alfabisaboleno sintasa, (+)-bornil difosfato sintasa, levopimaradieno sintasa, Abietadieno sintasa, isopimaradieno sintasa, (E)-gamma-bisaboleno sintasa, taxadieno sintasa, copalil pirofosfato sintasa, caureno sintasa, longifoleno sintasa, gamma-humuleno sintasa, Delta-selineno sintasa, beta-felandreno sintasa, limoneno sintasa, mirceno sintasa, terpinoleno sintasa, (-)-camfeno sintasa, (+)-3-careno sintasa, sin-copalil difosfato sintasa, alfa-terpineol sintasa, sin-pimara-7,15-dieno sintasa, ent-sandaaracopimaradieno sintasa, estemero-13-eno sintasa, E-beta-ocimeno, S-linalool sintasa, geraniol sintasa, gamma-terpineno sintasa, linalool sintasa, E-beta-ocimeno sintasa, epi-cedrol sintasa, alfa-zingibereno sintasa, guaiadieno sintasa, cascariladieno sintasa, cis-muroladieno sintasa, afidicolan-16b-ol sintasa, elizabetatrieno sintasa, sandalol sintasa, pachulol sintasa, Zinzanol sintasa, cedrol sintasa, escareol sintasa, copalol sintasa, manool sintasa, y similares.

30 [0083] Las secuencias de nucleótidos que codifican terpeno sintasas son conocidas en la técnica y se puede utilizar cualquier secuencia de nucleótidos que codifique terpeno sintasas para modificar genéticamente una célula huésped. Por ejemplo, las siguientes secuencias de nucleótidos que codifican terpeno sintasas, seguidas de sus números de acceso a Gen Bank y los organismos en que se identificaron, son conocidas y se pueden utilizar: ARNm de (-)-germacreno D sintasa (AY438099; *Populus balsamifera* subsp. *trichocarpa* x *Populus deltoids*); ARNm de E,E-alfa-farneseno sintasa (AY640154; *Cucumis sativus*); ARNm de 1,8-cineolo sintasa (AY691947; *Arabidopsis thaliana*); ARNm de terpeno sintasa 5 (TPS5) (AY518314; *Zea mays*); ARNm de terpeno sintasa 4 (TPS4) (AY518312; *Zea mays*); ARNm de mirceno/ocimeno sintasa (TPS10) (At2g24210) (NM\_127982; *Arabidopsis thaliana*); ARNm de geraniol sintasa (GES) (AY362553; *Ocimum basilicum*); ARNm de pineno sintasa (AY237645; *Picea sitchensis*); ARNm de mirceno sintasa 1e20 (AY195609; *Antirrhinum majus*); ARNm de (E)-β-ocimeno sintasa (0e23) (AY195607; *Antirrhinum majus*); ARNm de E-β-ocimeno sintasa (AY151086; *Antirrhinum majus*); ARNm de terpeno sintasa (AF497492; *Arabidopsis thaliana*); ARNm de (-)-canfeno sintasa (AG6.5) (U87910; *Abies grandis*); gen (por ejemplo, secuencia genómica) de (-)-4S-limoneno sintasa (AF326518; *Abies grandis*); gen de delta-selineno sintasa (AF326513; *Abies grandis*); ARNm de amorfa-4,11-dieno sintasa (AJ251751; *Artemisia annua*); ARNm de E-α-bisaboleno sintasa (AF006195; *Abies grandis*); LARN de gamma-humuleno sintasa (U92267; *Abies grandis*); ARNm de δ-selineno sintasa (U92266; *Abies grandis*); ARNm de pineno sintasa (AG3.18) (U87909; *Abies grandis*); ARNm de mirceno sintasa (AG2.2) (U87908; *Abies grandis*); etc.

50 [0084] Las secuencias de aminoácidos de las siguientes terpeno sintasas se encuentra bajo los números de acceso del GenBank mostrados en paréntesis, junto con el organismo en que se identificaron, después de cada terpeno sintasa: (-)-germacreno D sintasa (AAR99061; *Populus balsamifera* subsp. *trichocarpa* x *Populus deltoids*); D-cadineno sintasa (P93665; *Gossypium hirsutum*); 5-epi-aristoloueno sintasa (Q40577; *Nicotiana tabacum*); E,E-alfa-farneseno sintasa (AAU05951; *Cucumis sativus*); 1,8-cineolo sintasa (AAU01970; *Arabidopsis thaliana*); (R)-limoneno sintasa 1 (Q8L5K3; *Citrus limon*); sin-copalil difosfato sintasa (AAS98158; *Oryza sativa*); una taxadieno sintasa (Q9FT37; *Taxus chinensis*; Q93YA3; *Taxus bacca*; Q41594; *Taxus brevifolia*); una D-cadineno sintasa (Q43714; *Gossypium arboretum*); terpeno sintasa 5 (AAS88575; *Zea mays*); terpeno sintasa 4 (AAS88573; *Zea mays*); terpenoide sintasa (AAS79352; *Vitis vinifera*); geraniol sintasa (AAR11765; *Ocimum basilicum*); mirceno sintasa 1e20 (AA041727; *Antirrhinum majus*); 5-epi-aristoloueno sintasa 37 (AAP05762; *Nicotiana attenuata*); (+)-3-careno sintasa (AA073863; *Picea abies*); (-)-canfeno sintasa (AAB70707; *Abies grandis*); abietadieno sintasa (AAK83563; *Abies grandis*); amorfa-4,11-dieno sintasa (CAB94691; *Artemisia annua*); tricodieno sintasa (AAC49957; *Myrothecium roridum*); gamma-humuleno sintasa (AAC05728; *Abies grandis*); δ-selineno sintasa (AAC05727; *Abies grandis*); etc.

65 Ácidos nucleicos, vectores, promotores

5 [0085] Para generar una célula huésped modificada genéticamente, se introducen en una célula huésped de forma estable o transitoria uno o más ácidos nucleicos que comprenden secuencias de nucleótidos que codifican uno o más productos génicos utilizando técnicas establecidas, incluyendo, pero sin limitación, la electroporación, precipitación de fosfato de calcio, transfección mediada por DEAE-dextrano, transfección mediada por liposoma, choque térmico en presencia de acetato de litio, y similares. Para una transformación estable, un ácido nucleico incluirá generalmente además un marcador seleccionable, por ejemplo, cualquiera de varios marcadores seleccionables conocidos, tales como resistencia a neomicina, resistencia a ampicilina, resistencia a tetraciclina, resistencia a cloramfenicol, resistencia a kanamicina, y similares.

10 [0086] En muchas realizaciones, el ácido nucleico con el que la célula huésped es modificada genéticamente es un vector de expresión que incluye un ácido nucleico que comprende una secuencia de nucleótidos que codifica un producto génico, por ejemplo, una enzima del mecanismo del mevalonato, un factor de transcripción, una preniltransferasa, una terpeno sintasa, etc. Los vectores de expresión adecuados incluyen, pero sin limitación, vectores de baculovirus, vectores de bacteriófagos, plásmidos, fagémidos, cósmidos, fósmodos, cromosomas artificiales bacterianos, vectores virales (por ejemplo, vectores virales basados en virus de vacuna, poliovirus, adenovirus, virus adeno-asociados, SV40, virus de herpes simplex, y similares), cromosomas artificiales basados en P1, plásmidos de levadura, cromosomas artificiales de levadura, y cualquier otro vector específico para huéspedes específicos de interés (tales como levadura). De este modo, por ejemplo, un ácido nucleico que codifica un producto o productos génicos está incluido en cualquiera de un grupo de vectores de expresión para expresar el producto o productos génicos. Dichos vectores incluyen secuencias de ADN cromosómico, no cromosómico y sintético.

25 [0087] Se conocen numerosos vectores de expresión adecuados por los expertos en la materia, y muchos están disponibles comercialmente. Se proporcionan los siguientes vectores a modo de ejemplo; para células huésped eucariotas: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, y pSVLSV40 (Pharmacia). Sin embargo, se puede utilizar cualquier otro plásmido o vector siempre que sea compatible con la célula huésped.

30 [0088] La secuencia de nucleótidos en el vector de expresión está unida operativamente a una secuencia o secuencias de control de la expresión apropiadas (promotores) para dirigir la síntesis del producto génico codificado. Dependiendo del sistema huésped/vector utilizado, se pueden utilizar en el vector de expresión cualquiera de un conjunto de elementos de control adecuados de la transcripción y traducción, incluyendo promotores constitutivos e inducibles, elementos potenciadores de la transcripción, terminadores de la transcripción (véase, por ejemplo, Bitter et al. (1987) *Methods in Enzymology*, 153:516-544).

35 [0089] Ejemplos no limitantes de promotores eucariotas adecuados (promotores que son funcionales en células eucariotas) incluyen CMV inmediatamente temprano, HSV timidina quinasa, SV40 temprano y tardío, LTR de retrovirus y metalotioneína-I de ratón. La selección del vector y promotor adecuados se encuentra en la rutina del experto en la materia. El vector de expresión también puede contener un sitio de unión a ribosoma para el inicio de la traducción y un terminador de la transcripción. El vector de expresión también puede incluir secuencias apropiadas para amplificar la expresión.

45 [0090] Además, los vectores de expresión contendrán en muchas realizaciones uno o más genes marcadores seleccionables para proporcionar una característica fenotípica para la selección de células huésped transformadas, tales como resistencia a dihidrofolato reductasa o neomicina para un cultivo de células eucariotas.

50 [0091] En general, los vectores de expresión recombinante incluirán orígenes de replicación y marcadores seleccionables que permiten la transformación de la célula huésped, por ejemplo, el gen TRP1 de *S. cerevisiae*; y un promotor derivado de un gen altamente expresado para dirigir la transcripción de la secuencia codificante del producto génico. Dichos promotores pueden derivar de operones que codifican enzimas glicolíticos, tales como 3-fosfoglicerato quinasa (PGK), factor  $\alpha$ , fosfatasa ácida, o proteínas de choque térmico, entre otros.

55 [0092] En muchas realizaciones, una célula huésped modificada genéticamente se modifica genéticamente con un ácido nucleico que incluye una secuencia de nucleótidos que codifica un producto génico, donde la secuencia de nucleótidos que codifica el producto génico está unido operativamente a un promotor inducible. Los promotores inducibles son conocidos en la técnica. Los promotores inducibles adecuados incluyen, pero sin limitación, el pL de bacteriófago  $\lambda$ ; Plac; P<sub>trp</sub>; Ptac (promotor híbrido P<sub>trp</sub>-lac); un promotor inducible por isopropil-beta-D-tiogalactopiranosido (IPTG), por ejemplo, un promotor lacZ; un promotor inducible por tetraciclina; un promotor inducible por arabinosa, por ejemplo, P<sub>BAD</sub> (véase, por ejemplo, Guzman et al. (1995) *J. Bacteriol.* 177:4121-4130); un promotor inducible por xilosa, por ejemplo, P<sub>xyl</sub> (véase, por ejemplo, Kim et al. (1996) *Gene* 181: 71-76); un promotor GAL1; un promotor de triptófano; un promotor lac; un promotor inducible por alcohol, por ejemplo, un promotor inducible por metanol, un promotor inducible por etanol; un promotor inducible por rafinosa; un promotor inducible por el calor, por ejemplo, un promotor P<sub>L</sub> lambda inducible por el calor, un promotor controlado por un represor sensible al calor (por ejemplo, vectores de expresión basados en lambda reprimidos por CI857; véase, por ejemplo, Hoffmann et al. (1999) *FEMS Microbiol Lett.* 177(2):327-34); y similares.



**[0093]** En muchas realizaciones, una célula huésped modificada genéticamente es modificada genéticamente con un ácido nucleico que incluye una secuencia de nucleótidos que codifica un producto génico, donde la secuencia de nucleótidos que codifica el producto génico está unida operativamente a un promotor constitutivo. En levadura, se pueden utilizar un conjunto de vectores que contienen promotores constitutivos o inducibles. Para una revisión, véase, Current Protocols in Molecular Biology, Vol. 2, 1988, Ed. Ausubel, et al., Greene Publish. Assoc. & Wiley Interscience, Ch. 13; Grant, et al., 1987, Expression and Secretion Vectors for Yeast, en Methods in Enzymology, Eds. Wu & Grossman, 31987, Acad. Press, N.Y., Vol. 153, pp.516-544; Glover, 1986, DNA Cloning, Vol. II, IRL Press, Wash., D.C., Ch. 3; Bitter, 1987, Heterologous Gene Expression in Yeast, Methods in Enzymology, Eds. Berger & Kimmel, Acad. Press, N.Y., Vol. 152, pp. 673-684; y The Molecular Biology of the Yeast Saccharomyces, 1982, Eds. Strathern et al., Cold Spring Harbor Press, Vols. I y II. Se puede utilizar un promotor de levadura constitutivo, tal como ADH o LEU2 o un promotor inducible, tal como GAL (Cloning in Yeast, Ch. 3, R. Rothstein in: DNA Cloning Vol. 11, A Practical Approach, Ed. DM Glover, 1986, IRL Press, Wash., D.C.). Alternativamente, se pueden utilizar vectores que inducen la integración de secuencias de ADN exógenas en el cromosoma de la levadura.

*Composiciones que comprenden una célula huésped eucariota modificada genéticamente en cuestión*

**[0094]** La presente invención proporciona adicionalmente composiciones que comprenden una célula huésped eucariota modificada genéticamente en cuestión. Una composición en cuestión comprende una célula huésped eucariota modificada genéticamente en cuestión; y en algunas realizaciones comprenderá uno o más componentes adicionales, cuyos componentes se seleccionan basándose en parte en el uso pretendido de las células huésped eucariotas modificadas genéticamente. Los componentes adecuados incluyen, pero sin limitación, sales; tampones, estabilizadores; agentes inhibidores de proteasas; compuestos conservantes de la membrana celular y/o de la pared celular, por ejemplo, glicerol, dimetilsulfóxido, etc.; medios nutricionales apropiados para la célula; y similares.

#### **MÉTODOS PARA LA PRODUCCIÓN DE COMPUESTOS ISOPRENOIDES**

**[0095]** La presente invención proporciona métodos de producción de un compuesto isoprenoide o un precursor de isoprenoide según la reivindicación 3. Los métodos implican, en general, cultivar una célula huésped modificada genéticamente en cuestión de la reivindicación 1 en un medio adecuado.

**[0096]** Los sesquiterpenos incluyen periplanona B, ginkgolide B, amorfadieno, artemisinina, ácido artemisinico, valenceno, nootkatona, epi-cedrol, epi-aristoloqueno, farnesol, gosispol, sanonina, periplanona, y forskolina.

**[0097]** En algunas realizaciones, un método en cuestión comprende además aislar el compuesto isoprenoide de la célula y/o del medio de cultivo.

**[0098]** En general, una célula huésped modificada genéticamente en cuestión se cultiva en un medio adecuado (por ejemplo, caldo Luria-Bertoni, opcionalmente suplementado con uno o más agentes adicionales, tales como un inductor (por ejemplo, donde una o más secuencias de nucleótidos que codifican un producto génico están bajo el control de un promotor inducible), etc.). En algunas realizaciones, una célula huésped modificada genéticamente en cuestión se cultiva en un medio adecuado; y el medio de cultivo se recubre con un disolvente orgánico, por ejemplo, dodecano, que forma una capa orgánica. El compuesto isoprenoide producido por la célula huésped modificada genéticamente se separa en la fase orgánica, de la cual se puede purificar. En algunas realizaciones, donde una o más de las secuencias de nucleótidos que codifican productos génicos están unidas operativamente a un promotor inducible, se añade un inductor al medio de cultivo; y, después de un tiempo adecuado, el compuesto isoprenoide se aísla de la fase orgánica que recubre el medio de cultivo.

**[0099]** En algunas realizaciones, el compuesto isoprenoide se separará de otros productos que pueden estar presentes en la fase orgánica. La separación del compuesto isoprenoide de otros productos que pueden estar presentes en la fase orgánica se consigue fácilmente utilizando, por ejemplo, técnicas cromatográficas estándar.

**[0100]** En algunas realizaciones, el compuesto isoprenoide es puro, por ejemplo, por lo menos aproximadamente 40% puro, por lo menos aproximadamente 50% puro, por lo menos aproximadamente 60% puro, por lo menos aproximadamente 70% puro, por lo menos aproximadamente 80% puro, por lo menos aproximadamente 90% puro, por lo menos aproximadamente 95% puro, por lo menos aproximadamente 98% puro, o más del 98% puro, donde "puro" en el contexto de un compuesto isoprenoide se refiere a un compuesto isoprenoide que está libre de otros compuestos isoprenoides, contaminantes, etc.

#### **EJEMPLOS**

**[0101]** Los siguientes ejemplos se presentan con el fin de proporcionar a los expertos en la materia una información y descripción completas de cómo realizar y utilizar la presente invención, y no están destinados a limitar el alcance de lo que los inventores consideran su invención ni pretenden representar que los experimentos

que siguen son todos o los únicos experimentos realizados. Se han realizado esfuerzos para asegurar la precisión con respecto a los números utilizados (por ejemplo, cantidades, temperatura, etc), pero deben tenerse en cuenta algunos errores y desviaciones experimentales. A menos que se indique lo contrario, las partes son partes en peso, peso molecular es peso molecular promedio en peso, la temperatura es en grados Celsius, y la presión está en o cerca de la atmosférica. Se pueden utilizar abreviaturas estándar, por ejemplo, pb, pares de bases; kb, kilobases; pl, picolitros; s o seg, segundos; min, minutos; h, horas; aa, aminoácidos; kb, kilobases; pb, pares de bases; nt, nucleótidos; i.m., intramuscular(mente); i.p., intraperitoneal(mente); s.c., subcutánea(mente); y similares.

10 Ejemplo 1: Producción de niveles elevados de un compuesto isoprenoide en una célula de levadura modificada genéticamente

## MATERIALES Y MÉTODOS

15 **[0102] Productos químicos.** El dodecano y cariofileno se adquirieron de Sigma-Aldrich (St. Louis, MO). El ácido 5-fluoroórtico (5-FOA) se adquirió de Zymo Research (Orange, CA). Las Mezclas de Suplemento Completas para la formulación de medios Definidos Sintéticos se adquirieron de Qbiogene (Irvine, CA). Todos los otros componentes del medio se adquirieron de Sigma-Aldrich o Becton, Dickinson (Franklin Lakes, NJ).

20 **[0103] Cepas y medios.** Se utilizaron las cepas de *Escherichia coli* DH10B y DH5 $\alpha$  para la transformación bacteriana y la amplificación de plásmidos en la construcción de los plásmidos de expresión utilizados en este estudio. Las cepas se cultivaron a 37°C en medio Luria-Bertani con 100 mg/litro de ampicilina con la excepción de plásmidos de base p $\delta$ -UB que se cultivaron con 50 mg/litro de ampicilina.

25 **[0104]** Se utilizó la cepa de *Saccharomyces cerevisiae* BY4742 (Baker Brachmann et al. (1998) Yeast 14(2):115-132), un derivado de S288C, como la cepa original para todas las cepas de levadura. Esta cepa se desarrolló en medio rico en YPD. Burke et al. Methods in yeast genetics: a Cold Spring Harbor laboratory course manual. 2000, Plainview, NY: Cold Spring Harbor Laboratory Press. Las cepas de levadura modificadas se desarrollaron en medio Definido Sintético (SD) (Burke et al. (2000) supra) con leucina, uracilo, histidina y/o metionina añadidos según fuera apropiado. Para la inducción de genes expresados a partir del promotor GAL1, se desarrollaron cepas de *S. cerevisiae* en galactosa al 2% como la única fuente de carbono.

35 **[0105] Construcción de plásmidos.** Para crear el plásmido pRS425ADS para la expresión de ADS con el promotor GAL1, se amplificó ADS mediante reacción en cadena de la polimerasa (PCR) a partir de pADS (Martin et al. (2003) Nat. Biotechnol. 21(7): p. 796-802) utilizando la pareja de cebadores ADS-SpeI-F/ADS-HindIII-R (Tabla 1). Utilizando estos cebadores, la secuencia de nucleótidos 5'-AAAACA-3' se clonó inmediatamente en dirección 5' del codón de inicio de ADS. Esta secuencia consenso se utilizó para la traducción eficaz (Looman et al. (1993) Nucleic Acids Research. 21(18):4268-71; Yun et al. (1996) Molecular Microbiol. 19(6):1225-39.) de ADS y otros genes inducibles por galactosa utilizados en esta estudio. El producto amplificado se separó con SpeI y HindIII y se clonó en pRS425GAL1 diferido con SpeI y HindIII (Mumberg et al. (1995) Gene 156(1):119-122).

Tabla 1

Cebador	Secuencia (5' a 3')
ADS-SpeI-F	<u>GGACTAGTAAAACA</u> <b>AT</b> GGCCCTGACCGAAGAG (SEQ ID NO:3)
ADS-HindIII-R	CCAAGCTTT <b>CA</b> GATGGACATCGGGTAAAC (SEQ ID NO:4)
HMGR-BamHI-F	CGGGATCCAAAACA <b>AT</b> GGCTGCAGACCAATTGGTG (SEQ ID NO:5)
HMGR-SalI-R	GCGT <b>CGACTT</b> AGGATTTAATGCAGGTGACG (SEQ ID NO:6)
pRS42X-PvuII-SacII-F	CTGCCGCGGGGCCGCAAAATTAAGCCTTC (SEQ ID NO:7)
pRS42X-PvuII-SacII-R	CTCCC <b>GCGGT</b> AGTACGGATTAGAAGCCGC (SEQ ID NO:8)
UPC2-BamHI-F	CGGGATCCAAAACA <b>AT</b> GAGCGAAGTCGGTATACAG (SEQ ID NO:9)
UPC2-SalI-R	GCGT <b>CGACTC</b> ATAACGAAAAATCAGAGAAATTTG (SEQ ID NO:10)
ECM22-BamHI-R	CGGGATCCAAAACA <b>AT</b> GACATCCGATGATGGGAATG (SEQ ID NO:11)
ECM22-SalI-R	GCGT <b>CGACTT</b> ACATAAAAGCTGAAAAGTTTGTAG (SEQ ID NO:12)

Los sitios de restricción están subrayados y la negrita indica un codón de inicio o parada

45 **[0106]** Para la expresión de tHMGR, se construyó el plásmido pRS-HMGR. En primer lugar, se introdujeron sitios de restricción SacII en pRS426GAL1 (Mumberg et al. (1995) Gene 156(1):119-122) en el extremo 5' del promotor GAL1 y el extremo 3' del terminador CYC1. El cassette promotor-sitio de clonación múltiple-terminador de pRS426GAL1 se amplificó por PCR utilizando la pareja de cebadores pRS42X-PvuII-SacII-F/pRS42X-PvuII-SacII-R (Tabla 1). El producto amplificado se clonó directamente en pRS426GAL1 digerido con PvuII para construir el vector pRS426-SacII. El dominio catalítico de HMG1 se amplificó por PCR a partir del plásmido pRH127-3 (Donald et al. (1997) Appl. Environ. Microbiol. 63(9):3341-44) con la pareja de cebadores HMGR-BamHI-

F/HMGR-Sall-R. El producto amplificado se separó con *Bam*HI y *Sall* y se clonó en pRS426-SacII digerido con *Bam*HI y *Xho*I.

5 [0107] El alelo *upc2-1* de UPC2 se amplificó por PCR a partir del plásmido pBD33 utilizando la pareja de cebadores UPC2-BamHI-F/UPC2-Sall-R. El producto amplificado se separó con *Bam*HI y *Sall* se clonó en pRS426-SacII digerido con *Bam*HI y *Xho*I para crear el plásmido pRS-UPC2. Así mismo, el gen *ECM22* que contenía la mutación del tipo *upc2-1* (glicina a aspartato en el residuo 790) se amplificó por PCR a partir del plásmido pBD36 utilizando la pareja de cebadores ECM22-BamHI-F/UPC2-Sall-R. El producto amplificado se separó con *Bam*HI y *Sall* y se clonó en pRS426-SacII digerido con *Bam*HI y *Xho*I para crear el plásmido pRS-ECM22.

15 [0108] Se construyó un plásmido para la integración del cassette de expresión *tHMGR* de pRS-HMGR en el genoma de la levadura utilizando el plásmido pδ-UB (Lee et al. (1997) *Biotechnol Prog.* 13(4):368-373). Se separó pRS-HMGR con *Sac*II y el fragmento del cassette de expresión se extrajo con gel y se clonó en pδ-UB digerido con *Sac*II. Para la integración de *upc2-1*, se creó pδ-UPC2 de una manera idéntica mediante la digestión de pRS-UPC2 con *Sac*II y trasladando el fragmento apropiado a pδ-UB.

20 [0109] Para sustituir el promotor *ERG9* por el promotor *MET3*, se construyó el plásmido pRS-ERG9. El plásmido pRH973 (Gardner et al. (1999) *J. Biol. Chem.* 274(44):31671-31678) contenía un segmento 5' truncado de *ERG9* colocado detrás del promotor *MET3*. El pRH973 se separó con *Apal* y *Cla*I y se clonó en pRS403 digerido con *Apal* y *Cla*I (Sikorski et al. (1989) *Genetics*, 122(1):19-27).

25 [0110] Para la expresión de *ERG20*, se construyó el plásmido pRS-ERG20. Se digirió en primer lugar el plásmido pRS-SacII con *Sall* y *Xho*I que crearon extremos cohesivos compatibles. A continuación, el plásmido se autoligó, eliminando los sitios *Sall* y *Xho*I para crear el plásmido pRS-SacII-DX. Se amplificó *ERG20* por PCR a partir del ADN genómico de BY4742 utilizando la pareja de cebadores *ERG20-Spel-F/ERG20-Sma*I-R. El producto amplificado se separó con *Spe*I y *Sma*I y se clonó en pRS-SacII-DX digerido con *Spe*I y *Sma*I. Para la integración del cassette de expresión de *ERG20*, se separó pRS-ERG20 con *Sac*II y el fragmento del cassette de expresión se extrajo con gel y se clonó en pδ-UB digerido con *Sac*II.

30 [0111] En la tabla 2 se proporciona una descripción de los plásmidos utilizados en este estudio.

Tabla 2

Nombre	Gen expresado	Estado del plásmido	Marcador
pRS425ADS	<i>ADS</i>	Replicón 2-micras	<i>LEU2</i>
pRS-HMGR	<i>tHMGR</i>	Replicón 2-micras	<i>URA3</i>
pRS-UPC2	<i>upc2-1</i>	Replicón 2-micras	<i>URA3</i>
pRS-ECM22	<i>ECM22</i> (mutante <i>upc2-1</i> )	Replicón 2-micras	<i>URA3</i>
pδ-HMGR	<i>tHMGR</i>	Integración	<i>URA3</i>
pδ-UPC2	<i>upc2-1</i>	Integración	<i>URA3</i>
pRS-ERG9	<i>P<sub>MET3</sub>-ERG9</i>	Integración	<i>HIS3</i>
pδ-ERG20	<i>ERG20</i>	Integración	<i>URA3</i>

35 [0112] En la tabla 3 se proporciona una lista de cepas de levadura utilizadas en este estudio y los genotipos relevantes de las cepas.

Tabla 3

BY4742	MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0
EPY201	BY4742 pRS425ADS
EPY203	BY4742 pRS425ADS pRS-HMGR
EPY204	BY4742 pRS425ADS pRS-UPC2
EPY205	BY4742 pRS425ADS pRS-ECM22
EPY206	BY4742 pRS425ADS pRS-ERG20
EPY207	BY4742 pRS425ADS <i>tHMGR</i> (ura+)
EPY209	BY4742 pRS425ADS <i>tHMGR upc2-1</i> (ura+)
EPY212	BY4742 pRS425ADS <i>tHMGR upc2-1 erg9::PMT3-ERG9</i> (ura+)
EPY214	BY4742 pRS425ADS <i>tHMGR upc2-1 erg9::PMT3-ERG9 ERG20</i> (ura+)

45 [0113] **Transformación y construcción de cepas de levadura.** Se utilizó la cepa BY4742 de *S. cerevisiae* (Carrie Baker Brachmann et al. (1998) "Yeast" 14(2):115-132), un derivado de S288C, como la cepa original para todas las cepas de *S. cerevisiae*. La transformación de todas las cepas de *S. cerevisiae* se realizó mediante el método estándar del acetato de litio (Gietz et al. (2002) *Guide to Yeast Genetics and Molecular and Cell Biology*, Pt B., Academic Press Inc: San Diego. 87-96). Se cribaron de tres a diez colonias de cada transformación para la

selección del mayor transformante productor de amorfadieno. La cepa EPY201 se construyó mediante la transformación de la cepa BY4742 con el plásmido pRS425ADS y la selección en placas SD-LEU. Se construyeron las cepas EPY203, EPY204, EPY205, y EPY206 mediante la transformación de la cepa EPY201 con el plásmido pRS-HMGR, pRS-UPC2, pRS-ECM22, y pRS-ERG20, respectivamente. Los transformantes se seleccionaron en placas SDLEU-URA. El plásmido p $\delta$ -HMGR se digirió con *Xho*I antes de la transformación del ADN en la cepa EPY201 para la construcción de EPY207. La cepa EPY207 se cultivó y puso en placas SD-LEU que incluían 1 g/L de selección 5-FOA de la pérdida del marcador *URA3*. El auxótrofo de uracilo resultante se transformó a continuación con ADN plasmídico p $\delta$ -UPC2 digerido con *Xho*I para la construcción de EPY209, que se seleccionó en placas SD-LEU-URA. El plásmido pRS-ERG9 se separó con *Hind*III para la integración de la fusión  $P_{MET3}$ -ERG9 en los locus de *ERG9* de EPY209 para la construcción de EPY212. Esta cepa se seleccionó en placas SD-LEU-URA-HIS-MET. Se cultivó EPY212 y se puso en placas SD-LEU-HIS-MET que contenían 5-FOA para la selección de la pérdida del marcador *URA3*. El auxótrofo de uracilo resultante se transformó a continuación con ADN plasmídico p $\delta$ -ERG20 digerido con *Xho*I para la construcción de EPY214, que se seleccionó en placas SD-LEU-URA-HISMAT.

**[0114] Cultivo de levadura.** Para los experimentos con evolución en el tiempo para la medición de la producción de amorfadieno, se inocularon tubos de cultivo que contenían 5 mL de medio SD (galactosa al 2%) (con omisiones apropiadas de aminoácidos tal como se ha descrito anteriormente) con las cepas de interés. Estas inoculaciones se hicieron crecer a 30°C hasta una densidad óptica a 600 nm ( $DO_{600}$ ) de aproximadamente 1. Se inocularon matraces aireadores de 250 mL que contenían 50 mL de medio SD hasta una  $DO_{600}$  de 0,05 con estos cultivos de siembra. La figura 4 representa cepas desarrolladas en SD-URA-LEU-HIS con metionina al nivel indicada. Los medios para cepas mostradas en la figura 5 contenía SD-URA complementado con metionina hasta una concentración final de 1 mM. Todos los otros experimentos de producción utilizaron SD-URA o SD-URA-LEU cuando era apropiado.

**[0115]** Todos los matraces contenían 5 ml de dodecano. Se tomó una muestra de esta capa de dodecano y se diluyó en acetato de etilo para la determinación de la producción de amorfadieno mediante GC-MS.

**[0116] Análisis GC-MS de amorfadieno.** Se midió la producción de amorfadieno por las diversas cepas mediante GCMS tal como se ha descrito previamente (Martin et al. (2001) *Biotechnology and Bioengineering*, 75(5):497-503) mediante rastreo sólo para dos iones, el ion molecular (204 *m/z*) y el ion 189 *m/z*. Las concentraciones de amorfadieno se convirtieron en equivalentes de cariofileno utilizando una curva estándar de cariofileno y la abundancia relativa de los iones 189 y 204 *m/z* a sus iones totales.

## RESULTADOS

**[0117]** Para maximizar la producción de amorfadieno, se realizó una estrategia por etapas con la sucesiva integración de constructos en el genoma de *S. cerevisiae*.

**[0118] Producción de amorfadieno.** Se modificó una célula huésped plataforma, *S. cerevisiae*, para una producción a nivel elevado de isoprenoides. *S. cerevisiae* dirige toda su producción de isoprenoides a través de isopentenil difosfato (IPP), y la mayoría de ésta a través entonces de farnesil difosfato (FPP). Los niveles de IPP y FPP se incrementaron en la cepa huésped. IPP y FPP se metabolizan hasta una variedad de productos nativos. En lugar de medir los niveles de FPP, se midió el nivel de amorfadieno, un producto directo de FPP que no se metabolizará o degradará durante la evolución con el tiempo del crecimiento. Se expresó la amorfadieno sintasa (*ADS*) en *S. cerevisiae* para la ciclación enzimática de FPP en el sesquiterpeno amorfadieno. El amorfadieno también se cuantifica fácilmente mediante GCMS.

**[0119]** Se expresó *ADS* en plásmidos de 2 micras pRS425ADS bajo el control inducible del promotor *GAL1*. Los cultivos de *S. cerevisiae* se desarrollaron durante seis días en galactosa para la expresión de *ADS*, y se midieron los niveles de amorfadieno cada 24 horas. *S. cerevisiae* se modificó únicamente mediante la introducción de pRS425ADS alcanzando una producción máxima de amorfadieno de 4,6  $\mu$ g de amorfadieno por ml después de cuatro días (figura 3A).

**[0120]** Los experimentos de control previos que consistían en medio adicionado con amorfadieno puro mostraron la pérdida rápida del sesquiterpeno de la fase líquida. Se añadió una fase de dodecano equivalente al 10% del volumen del medio a cada matraz de agitación para secuestrar el amorfadieno del cultivo. La adición de esta fase orgánica asegura una medición precisa de la cantidad total de amorfadieno producido evitando la pérdida en el aire. La volatilización de amorfadieno es un problema particular durante las evoluciones extendidas de varios días como los utilizados en este estudio.

**[0121] Sobreexpresión de HMG-CoA reductasa.** La importancia médica de la biosíntesis del colesterol y la facilidad experimental del análisis en *S. cerevisiae* lo ha convertido en un organismo ideal para el estudio de la regulación del mecanismo del mevalonato durante las últimas décadas (Szkopinska et al. (2000) *Biochemical and Biophysical Research Communications*, 267 (1):473-477; Dimster-Denk et al. (1999) *J. Lipid Res.*, 40(5):850-860).

[0122] Estos estudios han elucidado un sistema complejo de regulación, con 3-hidroxi-3-metilglutaril-coenzima A reductasa (HNIGR) como el principal punto de control regulador del mecanismo. Dos isozimas de HMGR, Hmg1p y Hmg2p, están presentes en la levadura, siendo Hmg1p la más estable de las dos (Hampton et al. (1996) Trends in Biochemical Sciences, 21(4):140-145). Hmg1p es una proteína integral unida a membrana que contiene una región N-terminal responsable del anclaje de la proteína a la membrana del ER (Liscum et al. (1985) J. Biol. Chem. 260(1):522-530). Para la expresión de una forma soluble de la enzima (Donald et al. (1997) Appl. Environ. Microbiol. 63(9):3341-44) se extrajo el extremo N-terminal unido a membrana de Hmg1p y se expresó sólo el dominio catalítico. En nuestro estudio, esta forma truncada de HMGR (tHMGR) en un plásmido de 2 micras se expresó bajo el control del promotor *GAL1*. Cuando se expresó conjuntamente con ADS, *S. cerevisiae* alcanzó una producción máxima de 11,2 µg de amorfadieno por ml después de cuatro días (figura 3A.).

[0123] **Sobreexpresión de factores de transcripción implicados en esteroles.** En otra estrategia para incrementar el amorfadieno, se utilizaron dos factores de transcripción de *S. cerevisiae* identificados previamente por su importancia en la regulación de la biosíntesis de esteroles. Los mutantes *upc2-1* de *S. cerevisiae* se identificaron originalmente por su capacidad única de captar esteroides bajo condiciones aeróbicas (Lewis et al. (1988) Yeast, 4(2):93-106). La caracterización posterior mostró que estos mutantes habían incrementado las capacidades de síntesis del esteroles (Lewis et al. (1988) Yeast, 4(2):93-106). La mutación responsable para estas características es una única transición de guanina a adenina en el gen *UPC2*; esta mutación puntual da lugar a un cambio de residuos de glicina a aspartato en el aminoácido 888 próximo al extremo carboxilo (Crowley et al. (1998) J. Bacteriol., 180(16):4177-83). Se identificó posteriormente un homólogo a este gen, *ECM22*, con una identidad en la secuencia de aminoácidos del 45% (Shianna et al. (2001) J. Bacteriol., 183(3):830-834). Se conservan completamente 36 aminoácidos entre *UPC2* y *ECM22* en el locus de la mutación puntual *upc2-1* (Shianna et al. (2001) J. Bacteriol., 183(3):830-834). La mutación puntual *upc2-1* se introdujo en el alelo *ECM22* +de tipo salvaje dando lugar a una cepa con un fenotipo similar al del mutante *upc2-1* (Shianna et al. (2001) J. Bacteriol., 183(3):830-834).

[0124] Vik y Rine identificaron *ERG2* y *ERG3* como dianas para la regulación génica por Ecm22p y Upc2p. Se identificó un elemento regulador de esteroles de 7 pares de bases como la localización de unión necesaria para estos factores de transcripción. Este elemento de secuencia de 7 pares de bases se halla en los promotores de muchos otros genes del mecanismo del esteroles, incluyendo *ERGB*, *ERG12*, y *ERG13* (Vik et al. (2001) Mol. Cell. Biol., 21(19):6395-6405.). Los productos enzimáticos para cada uno de estos tres genes están implicados en la síntesis de isoprenoides cadena arriba de FPP (véase la figura 1).

[0125] Se realizó la hipótesis de que la coexpresión de los alelos mutantes para *UPC2* y *ECM22* con ADS incrementaría la producción de amorfadieno al incrementar el flujo metabólico a través del mecanismo del mevalonato. Los alelos mutantes *upc2-1* de *UPC2* y *EMC22* se expresaron cada uno bajo el control del promotor *GAL1* en un plásmido de 2 micras en una cepa que ya albergaba pRS425ADS. La producción absoluta de amorfadieno en los cultivos se incrementó únicamente de forma mínima para la expresión de *UPC2* y *EMC22*, en parte debido a la disminución de las densidades celulares. Sin embargo, la producción normalizada para la densidad celular aumentó un 76% y un 53% para la expresión de *UPC2* y *EMC22*, respectivamente (figura 3B).

[0126] Este incremento relativamente pequeño en la producción de amorfadieno en comparación con la sobreexpresión de tHMGR apoya el hecho de que la actividad de HMGR es el principal cuello de botella limitante del mecanismo del mevalonato. Incluso es improbable que una expresión a nivel elevado de *ERG 8*, *ERG12*, y *ERG13* aumente considerablemente el flujo a través del mecanismo si la HMGR permanece en un nivel de expresión basal. La menor densidad celular observada para la sobreexpresión de *UPC2* y *EMC22* es improbable que sea debida al incremento del flujo a través del mecanismo del mevalonato hacia FPP. En cambio, es probable que esté causada por un cambio desfavorable en la regulación transcripcional para otro u otros genes controlados por *UPC2* y *EMC22*.

[0127] **Coexpresión de tHMGR y *upc2-1*.** La sobreexpresión de tHMGR y *upc2-1* aumentó cada uno el rendimiento final de amorfadieno en los cultivos celulares. Para analizar la posibilidad de un efecto sinérgico de la sobreexpresión de estos genes juntos, se integraron los cassettes de expresión de manera secuencial en el genoma de *S. cerevisiae*. Se utilizó el plásmido p $\delta$ -UB (Lee et al. (1997) Biotechnol Prog., 13(4):368-373) para la construcción de los plásmidos de integración. Este plásmido contiene un Cassette Blaster URA3 reutilizable que permite reciclar el marcador URA3. Adicionalmente, se integra en una secuencia  $\delta$  (hallada en las repeticiones largas terminales de los sitios de Tytransposon), de la cual existen aproximadamente 425 dispersadas a lo largo del genoma (Dujon (1996) Trends in Genetics, 12(7):263-270).

[0128] Se integró tHMGR en el cromosoma de una cepa que albergaba pRS425ADS utilizando p $\delta$ -HMGR. El nivel de producción de amorfadieno de 13,8 µg de amorfadieno por ml era comparable en esta cepa con la cepa EP203 que contenía tHMGR en un plásmido de copia elevada (figura 5). Después de reciclar el marcador URA3 mediante el emplacado en 5-FOA, se integró *upc2-1* en el cromosoma utilizando el plásmido p $\delta$ -UPC2. Se combinaron los efectos de la sobreexpresión de tHMGR y *upc2-1* para aumentar la producción de amorfadieno

hasta 16,2 µg de amorfadieno por ml (figura 5). Aunque la expresión de *upc2-1* en combinación con *tHMGR* aumentó la producción absoluta de amorfadieno en un 17%, este incremento sólo es comparable con el observado cuando *upc2-1* se expresa con *ADS* solo. Con la eliminación del cuello de botella de *HMGR*, se esperó un impacto más significativo de la expresión de *upc2-1*. Los aumentos potenciales en la producción de amorfadieno se podrían evitar debido al direccionamiento de FPP a otros metabolitos.

**[0129] Regulación por descenso de escualeno sintasa.** Los incrementos observados en la producción de amorfadieno sugirieron un aumento del grupo de precursores de FPP. FPP es fundamental en la síntesis de un conjunto de compuestos de *S. cerevisiae* incluyendo esteroides, dolicoles y poliprenoles y proteínas preniladas. Aunque un incremento del flujo a través del mecanismo del mevalonato, conducía a una mayor producción de amorfadieno, un grupo de otras enzimas también competían por incremento del grupo de FPP, de manera más destacada la escualeno sintasa codificada por *ERG9*. La síntesis de escualeno es el punto de ramificación de la FPP que conduce a ergosterol. En una cepa que expresaba el dominio catalítico de *HMGR* y que contenía una delección de *ERG9*, se observó que se acumulaba FPP (Song (2003) *Analytical Biochemistry*, 317(2):180-185). Con el objetivo de direccionar el FPP fuera de la producción de esterol y hacia la producción de amorfadieno, la reducción en la actividad de escualeno sintasa sería útil. Sin embargo, una delección de *ERG9* es letal sin la suplementación exógena de esteroides.

**[0130]** Utilizando una estrategia alternativa, se reguló *ERG9* por descenso de manera transcripcional mediante la sustitución de su promotor nativo por un promotor reprimible de metionina,  $P_{MET3}$  (Cherest et al. (1985) *Gene*, 34(2-3):269-281). Gardner et al. utilizaron previamente dicho constructo de fusión  $P_{MET3}$ -*ERG9* para el estudio de las señales de degradación de *HMGR* (Gardner et al. (1999) *J. Biol. Chem.* 274(44):31671-31678; Gardner et al. (2001) *J. Biol. Chem.*, 276(12):8681-8694). Se construyó el plásmido pRS-*ERG9* para utilizar la misma estrategia que Gardner en la sustitución del promotor nativo *ERG9* por el promotor *MET3*. La utilidad de la fusión  $P_{MET3}$ -*ERG9* se destaca por el estrecho control regulador entre 0 y 100 µM de concentraciones extracelulares de metionina (Mao et al. (2002) *Current Microbiology*, 45(1):37-40). En presencia de concentraciones extracelulares elevadas de metionina, la expresión a partir del promotor *MET3* es muy baja. Después de la integración de pRS-*ERG9* en el locus de *ERG9*, se pudo ajustar la expresión de escualeno sintasa en base a la suplementación de metionina al medio.

**[0131]** Se integró pRS-*ERG9* en la cepa EPY209, y se midió la producción de amorfadieno con un intervalo de 0 a 1 mM en el medio. Se muestran los puntos de tiempo de 64 y 87 horas después de la inoculación (figura 4). Los datos sugieren que la expresión mínima de *ERG9* (concentraciones de metionina por encima de 0,5 mM) maximiza la producción de amorfadieno. A medida que se incrementa la densidad celular en los cultivos de *S. cerevisiae* y se metabolizan los nutrientes en el medio, la concentración de metionina probablemente desciende, explicando el porqué los cultivos proporcionados con metionina 0,1 mM en medio presentan rendimientos menores de amorfadieno. Se seleccionó metionina 1 mM para futuros experimentos para asegurar concentraciones extracelulares elevadas a lo largo de las evoluciones extendidas con el tiempo.

**[0132]** Se desarrolló en cultivo la cepa EPU212 que contenía una copia integrada de *tHMGR* y *upc2-1*, así como el alelo reprimible de metionina de *ERG9*, y se midió la producción de amorfadieno durante seis días (figura 5). La limitación del FPP incorporado en escualeno tenía un gran impacto en la producción de amorfadieno, incrementándolo cuatro veces hasta 61 µg de amorfadieno por ml sobre la cepa EPY209 que contenía el alelo *ERG9* de tipo salvaje. Aunque limitada en su capacidad de producir ergosterol, la EPY212 aún crecía hasta una DO final de ~75% de la de EPY209.

**[0133] Sobreexpresión de FPP Sintasa.** La FPP Sintasa (FPPS), codificada por *ERG20*, fue reconocida como la próxima diana para la sobreexpresión con la esperanza de incrementar más las producciones de sesquiterpenos. Un incremento de seis veces en la actividad de FPPS se ha correlacionado con un incremento del 80% y el 32% en dolicol y ergosterol, respectivamente (Szkopinska et al. (2000) *Biochemical and Biophysical Research Communications*, 267(1):473-477). De forma similar a los estudios que sobreexpresan *HMGR* y *upc2-1*, se clonó primero *ERG20* detrás del promotor *GAL1* en un plásmido de copias elevadas para crear pRS-*ERG20*. La coexpresión de *ERG20* en este plásmido con pRS425ADS en realidad disminuyó la productividad absoluta de amorfadieno en un 60%. Es posible que un incremento en la actividad de FPPS incremente sólo el contenido de otros productos derivados de FPP, tales como ergosterol. Otra posibilidad es que la sobreexpresión de FPPS incremente la concentración intracelular de FPP, la principal señal para la degradación de *HMGR* (Gardner et al. (1999) *J. Biol. Chem.* 274(44):31671-31678). Sin la sobreexpresión de una forma no regulada de la reductasa, las concentraciones de FPP incrementadas podrían actuar para limitar el flujo a través del mecanismo del mevalonato y disminuir la producción de amorfadieno.

**[0134]** A continuación, se construyó p $\delta$ -*ERG20* para la integración y expresión de *ERG20* en nuestro mayor productor de amorfadieno. Se recicló el marcador *URA3* y se integró p $\delta$ -*ERG20* en el cromosoma para crear la cepa EPY214. Esta cepa que sobreexpresaba FPPS, incrementó adicionalmente la producción de amorfadieno hasta 73 µg de amorfadieno por ml (figura 5). Anteriormente se había observado un descenso del 60% en la producción de amorfadieno en la cepa EPY206 que sobreexpresaba *ERG20* con *ADS*. Sin embargo, ahora en

una cepa que expresaba *tHMGR* y *upc2-1* y con una escualeno sintasa regulada, la producción de amorfadieno se incrementó en un 20% con la sobreexpresión de ERG20.

5 [0135] En las cepas EPY206 y EPY214 que expresaban cada una *ERG20*, se observó una disminución de la densidad celular. Este descenso en el crecimiento celular se podría explicar por la toxicidad causada directamente por ERG20p. Alternativamente, pudo surgir un efecto a partir de una acumulación o agotamiento de un intermedio del mecanismo debido al flujo modificado a través de la FPP sintasa.

LISTADO DE SECUENCIAS

10 [0136]  
 <110> THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 <120> CÉLULAS HUÉSPED MODIFICADAS GENÉTICAMENTE Y USO DE LAS MISMAS PARA PRODUCIR  
 COMPUESTOS ISOPRENOIDES  
 15 <130> AHB/FP6424808  
 <140> 05775484.82005-07-21  
 <150> PCT/US2005/026190  
 <151> 2005-07-21  
 <150> 60/592,009  
 20 <151> 2004-07-27  
 <160> 12  
 <170> FastSEQ for Windows Version 4.0  
 <210> 1  
 <211> 1509  
 25 <212> ADN  
 <213> *Saccharomyces cerevisiae*  
 <220>  
 <221> CDS  
 <222> (1)...(1509)  
 30 <400> 1

35	atg gtt tta acc aat aaa aca gtc att tct gga tcg aaa gtc aaa agt	48
	Met Val Leu Thr Asn Lys Thr Val Ile Ser Gly Ser Lys Val Lys Ser	
	1 5 10 15	
40	tta tca tct gcg caa tcg agc tca tca gga cct tca tca tct agt gag	96
	Leu Ser Ser Ala Gln Ser Ser Ser Ser Gly Pro Ser Ser Ser Ser Glu	
	20 25 30	
45	gaa gat gat tcc cgc gat att gaa agc ttg gat aag aaa ata cgt cct	144
	Glu Asp Asp Ser Arg Asp Ile Glu Ser Leu Asp Lys Lys Ile Arg Pro	
	35 40 45	
50	tta gaa gaa tta gaa gca tta tta agt agt gga aat aca aaa caa ttg	192
	Leu Glu Glu Leu Glu Ala Leu Leu Ser Ser Gly Asn Thr Lys Gln Leu	
	50 55 60	
55	aag aac aaa gag gtc gct gcc ttg gtt att cac ggt aag tta cct ttg	240
	Lys Asn Lys Glu Val Ala Ala Leu Val Ile His Gly Lys Leu Pro Leu	
	65 70 75 80	
60	tac gct ttg gag aaa aaa tta ggt gat act acg aga gcg gtt gcg gta	288
	Tyr Ala Leu Glu Lys Lys Leu Gly Asp Thr Thr Arg Ala Val Ala Val	
	85 90 95	
65	cgt agg aag gct ctt tca att ttg gca gaa gct cct gta tta gca tct	336
	Arg Arg Lys Ala Leu Ser Ile Leu Ala Glu Ala Pro Val Leu Ala Ser	

ES 2 386 359 T3

	100	105	110		
5	gat cgt tta cca tat aaa aat tat Asp Arg Leu Pro Tyr Lys Asn Tyr 115	gac tac gac cgc gta ttt ggc gct Asp Tyr Asp Arg Val Phe Gly Ala 120		384	
10	tgt tgt gaa aat gtt ata ggt tac atg cct ttg ccc gtt ggt gtt ata Cys Cys Glu Asn Val Ile Gly Tyr Met Pro Leu Pro Val Gly Val Ile 130		140	432	
15	ggc ccc ttg gtt atc gat ggt aca tct tat cat ata cca atg gca act Gly Pro Leu Val Ile Asp Gly Thr Ser Tyr His Ile Pro Met Ala Thr 145	150	155	160	480
20	aca gag ggt tgt ttg gta gct tct gcc atg cgt ggc tgt aag gca atc Thr Glu Gly Cys Leu Val Ala Ser Ala Met Arg Gly Cys Lys Ala Ile 165	170	175	528	
25	aat gct ggc ggt ggt gca aca act gtt tta act aag gat ggt atg aca Asn Ala Gly Gly Gly Ala Thr Thr Val Leu Thr Lys Asp Gly Met Thr 180	185	190	576	
30	aga ggc cca gta gtc cgt ttc cca act ttg aaa aga tct ggt gcc tgt Arg Gly Pro Val Val Arg Phe Pro Thr Leu Lys Arg Ser Gly Ala Cys 195	200	205	624	
35	aag ata tgg tta gac tca gaa gag gga caa aac gca att aaa aaa gct Lys Ile Trp Leu Asp Ser Glu Glu Gly Gln Asn Ala Ile Lys Lys Ala 210	215	220	672	
40	ttt aac tct aca tca aga ttt gca cgt ctg caa cat att caa act tgt Phe Asn Ser Thr Ser Arg Phe Ala Arg Leu Gln His Ile Gln Thr Cys 225	230	235	240	720
45	cta gca gga gat tta ctc ttc atg aga ttt aga aca act act ggt gac Leu Ala Gly Asp Leu Leu Phe Met Arg Phe Arg Thr Thr Thr Gly Asp 245	250	255	768	
50	gca atg ggt atg aat atg att tct aaa ggt gtc gaa tac tca tta aag Ala Met Gly Met Asn Met Ile Ser Lys Gly Val Glu Tyr Ser Leu Lys 260	265	270	816	
55	caa atg gta gaa gag tat ggc tgg gaa gat atg gag gtt gtc tcc gtt Gln Met Val Glu Glu Tyr Gly Trp Glu Asp Met Glu Val Val Ser Val 275	280	285	864	
60	tct ggt aac tac tgt acc gac aaa aaa cca gct gcc atc aac tgg atc Ser Gly Asn Tyr Cys Thr Asp Lys Lys Pro Ala Ala Ile Asn Trp Ile 290	295	300	912	
65	gaa ggt cgt ggt aag agt gtc gtc gca gaa gct act att cct ggt gat Glu Gly Arg Gly Lys Ser Val Val Ala Glu Ala Thr Ile Pro Gly Asp 305	310	315	320	960
70	ggt gtc aga aaa gtg tta aaa agt gat gtt tcc gca ttg gtt gag ttg Val Val Arg Lys Val Leu Lys Ser Asp Val Ser Ala Leu Val Glu Leu 325	330	335	1008	
75	aac att gct aag aat ttg gtt gga tct gca atg gct ggg tct gtt ggt			1056	



ES 2 386 359 T3

	Asn Ile Ala Lys Asn Leu Val Gly Ser Ala Met Ala Gly Ser Val Gly	
	340 345 350	
5	gga ttt aac gca cat gca gct aat tta gtg aca gct gtt ttc ttg gca Gly Phe Asn Ala His Ala Ala Asn Leu Val Thr Ala Val Phe Leu Ala	1104
	355 360 365	
10	tta gga caa gat cct gca caa aat gtt gaa agt tcc aac tgt ata aca Leu Gly Gln Asp Pro Ala Gln Asn Val Glu Ser Ser Asn Cys Ile Thr	1152
	370 375 380	
15	ttg atg aaa gaa gtg gac ggt gat ttg aga att tcc gta tcc atg cca Leu Met Lys Glu Val Asp Gly Asp Leu Arg Ile Ser Val Ser Met Pro	1200
	385 390 395 400	
20	tcc atc gaa gta ggt acc atc ggt ggt ggt act gtt cta gaa cca caa Ser Ile Glu Val Gly Thr Ile Gly Gly Gly Thr Val Leu Glu Pro Gln	1248
	405 410 415	
25	ggt gcc atg ttg gac tta tta ggt gta aga ggc ccg cat gct acc gct Gly Ala Met Leu Asp Leu Leu Gly Val Arg Gly Pro His Ala Thr Ala	1296
	420 425 430	
30	cct ggt acc aac gca cgt caa tta gca aga ata gtt gcc tgt gcc gtc Pro Gly Thr Asn Ala Arg Gln Leu Ala Arg Ile Val Ala Cys Ala Val	1344
	435 440 445	
35	ttg gca ggt gaa tta tcc tta tgt gct gcc cta gca gcc ggc cat ttg Leu Ala Gly Glu Leu Ser Leu Cys Ala Ala Leu Ala Ala Gly His Leu	1392
	450 455 460	
40	ggt caa agt cat atg acc cac aac agg aaa cct gct gaa cca aca aaa Val Gln Ser His Met Thr His Asn Arg Lys Pro Ala Glu Pro Thr Lys	1440
	465 470 475 480	
45	cct aac aat ttg gac gcc act gat ata aat cgt ttg aaa gat ggg tcc Pro Asn Asn Leu Asp Ala Thr Asp Ile Asn Arg Leu Lys Asp Gly Ser	1488
	485 490 495	
50	gtc acc tgc att aaa tcc taa Val Thr Cys Ile Lys Ser *	1509
	500	
	<210> 2	
	<211> 502	
	<212> PRT	
50	<213> Saccharomyces cerevisiae	
	<400> 2	
55	Met Val Leu Thr Asn Lys Thr Val Ile Ser Gly Ser Lys Val Lys Ser 1 5 10 15	
	Leu Ser Ser Ala Gln Ser Ser Ser Ser Gly Pro Ser Ser Ser Ser Glu 20 25 30	
60	Glu Asp Asp Ser Arg Asp Ile Glu Ser Leu Asp Lys Lys Ile Arg Pro 35 40 45	
	Leu Glu Glu Leu Glu Ala Leu Leu Ser Ser Gly Asn Thr Lys Gln Leu 50 55 60	
65	Lys Asn Lys Glu Val Ala Ala Leu Val Ile His Gly Lys Leu Pro Leu	

ES 2 386 359 T3

	65					70					75					80
	Tyr	Ala	Leu	Glu	Lys	Lys	Leu	Gly	Asp	Thr	Thr	Arg	Ala	Val	Ala	Val
					85					90					95	
5	Arg	Arg	Lys	Ala	Leu	Ser	Ile	Leu	Ala	Glu	Ala	Pro	Val	Leu	Ala	Ser
				100					105					110		
	Asp	Arg	Leu	Pro	Tyr	Lys	Asn	Tyr	Asp	Tyr	Asp	Arg	Val	Phe	Gly	Ala
				115				120					125			
10	Cys	Cys	Glu	Asn	Val	Ile	Gly	Tyr	Met	Pro	Leu	Pro	Val	Gly	Val	Ile
				130			135					140				
	Gly	Pro	Leu	Val	Ile	Asp	Gly	Thr	Ser	Tyr	His	Ile	Pro	Met	Ala	Thr
	145					150					155					160
15	Thr	Glu	Gly	Cys	Leu	Val	Ala	Ser	Ala	Met	Arg	Gly	Cys	Lys	Ala	Ile
				165						170					175	
	Asn	Ala	Gly	Gly	Gly	Ala	Thr	Thr	Val	Leu	Thr	Lys	Asp	Gly	Met	Thr
				180					185					190		
20	Arg	Gly	Pro	Val	Val	Arg	Phe	Pro	Thr	Leu	Lys	Arg	Ser	Gly	Ala	Cys
				195				200					205			
	Lys	Ile	Trp	Leu	Asp	Ser	Glu	Glu	Gly	Gln	Asn	Ala	Ile	Lys	Lys	Ala
				210			215					220				
25	Phe	Asn	Ser	Thr	Ser	Arg	Phe	Ala	Arg	Leu	Gln	His	Ile	Gln	Thr	Cys
	225					230					235					240
	Leu	Ala	Gly	Asp	Leu	Leu	Phe	Met	Arg	Phe	Arg	Thr	Thr	Thr	Gly	Asp
				245						250					255	
30	Ala	Met	Gly	Met	Asn	Met	Ile	Ser	Lys	Gly	Val	Glu	Tyr	Ser	Leu	Lys
				260					265					270		
	Gln	Met	Val	Glu	Glu	Tyr	Gly	Trp	Glu	Asp	Met	Glu	Val	Val	Ser	Val
				275				280					285			
35	Ser	Gly	Asn	Tyr	Cys	Thr	Asp	Lys	Lys	Pro	Ala	Ala	Ile	Asn	Trp	Ile
				290			295					300				
	Glu	Gly	Arg	Gly	Lys	Ser	Val	Val	Ala	Glu	Ala	Thr	Ile	Pro	Gly	Asp
	305					310					315					320
40	Val	Val	Arg	Lys	Val	Leu	Lys	Ser	Asp	Val	Ser	Ala	Leu	Val	Glu	Leu
				325						330					335	
	Asn	Ile	Ala	Lys	Asn	Leu	Val	Gly	Ser	Ala	Met	Ala	Gly	Ser	Val	Gly
				340					345					350		
45	Gly	Phe	Asn	Ala	His	Ala	Ala	Asn	Leu	Val	Thr	Ala	Val	Phe	Leu	Ala
				355				360					365			
	Leu	Gly	Gln	Asp	Pro	Ala	Gln	Asn	Val	Glu	Ser	Ser	Asn	Cys	Ile	Thr
				370			375					380				
50	Leu	Met	Lys	Glu	Val	Asp	Gly	Asp	Leu	Arg	Ile	Ser	Val	Ser	Met	Pro
	385					390						395				400
	Ser	Ile	Glu	Val	Gly	Thr	Ile	Gly	Gly	Gly	Thr	Val	Leu	Glu	Pro	Gln
				405						410					415	
55	Gly	Ala	Met	Leu	Asp	Leu	Leu	Gly	Val	Arg	Gly	Pro	His	Ala	Thr	Ala
				420					425					430		
	Pro	Gly	Thr	Asn	Ala	Arg	Gln	Leu	Ala	Arg	Ile	Val	Ala	Cys	Ala	Val
				435				440					445			
	Leu	Ala	Gly	Glu	Leu	Ser	Leu	Cys	Ala	Ala	Leu	Ala	Ala	Gly	His	Leu
				450			455					460				
60	Val	Gln	Ser	His	Met	Thr	His	Asn	Arg	Lys	Pro	Ala	Glu	Pro	Thr	Lys
	465					470					475					480
	Pro	Asn	Asn	Leu	Asp	Ala	Thr	Asp	Ile	Asn	Arg	Leu	Lys	Asp	Gly	Ser
				485						490					495	
65	Val	Thr	Cys	Ile	Lys	Ser										
				500												

<210> 3  
 <211> 32  
 <212> ADN  
 <213> Secuencia artificial  
 5 <220>  
 <223> cebador  
 <400> 3  
 ggactagtaa aacaatggcc ctgaccgaag ag 32  
 <210> 4  
 10 <211> 29  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 15 <400> 4  
 ccaagctttc agatggacat cgggtaaac 29  
 <210> 5  
 <211> 35  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 <400> 5  
 cgggatccaa aacaatggct gcagaccaat tggg 35  
 25 <210> 6  
 <211> 30  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> cebador  
 <400> 6  
 gcgtcgactt aggatataat gcaggtgacg 30  
 <210> 7  
 <211> 29  
 35 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 <400> 7  
 40 ctgccgcggg gccgcaaatt aaagccttc 29  
 <210> 8  
 <211> 29  
 <212> ADN  
 <213> Secuencia artificial  
 45 <220>  
 <223> cebador  
 <400> 8  
 ctgccgcggt agtacggatt agaagccgc 29  
 <210> 9  
 50 <211> 35  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 55 <400> 9  
 cgggatccaa aacaatgagc gaagtcggtg tacag 35  
 <210> 10  
 <211> 34  
 <212> ADN  
 60 <213> Secuencia artificial  
 <220>  
 <223> cebador  
 <400> 10  
 gcgtcgactc ataacgaaa atcagagaaa ttg 34  
 65 <210> 11  
 <211> 36

## ES 2 386 359 T3

<212> ADN  
<213> Secuencia artificial  
<220>  
<223> cebador  
5 <400> 11  
cgggatccaa aacaatgaca tccgatgatg ggaatg 36  
<210> 12  
<211> 34  
<212> ADN  
10 <213> Secuencia artificial  
<220>  
<223> cebador  
<400> 12  
gcgtcgactt acataaaagc tgaaaagttt gtag 34  
15

**REIVINDICACIONES**

1. Célula huésped de *Saccharomyces cerevisiae* que produce un compuesto precursor de isoprenoide o compuesto isoprenoide a través del mecanismo del mevalonato, en la que dicha célula es modificada genéticamente para comprender:
- 5 a) un ácido nucleico heterólogo integrado en el cromosoma de la célula huésped que codifica una 3-hidroxi-3-metilglutaril coenzima-A reductasa (HMGR) truncada que carece de un dominio que comprende la membrana y que mantiene su dominio catalítico C-terminal;
- 10 b) una farnesil difosfato sintasa (FPPS) heteróloga integrada en el cromosoma de la célula huésped para incrementar el nivel de actividad de dicha FPPS; y
- c) un ácido nucleico heterólogo integrado en el cromosoma de la célula huésped para disminuir el nivel de actividad de la escualeno sintasa;
- 15 en la que la célula huésped comprende además una sesquiterpeno sintasa heteróloga y es capaz de producir un isoprenoide que deriva de la acción de dicha sesquiterpeno sintasa, y
- en la que las modificaciones genéticas proporcionan la producción del isoprenoide derivado de la acción de dicha sesquiterpeno sintasa a un nivel que es por lo menos el 50% superior al nivel del isoprenoide en una célula de control que no comprende las modificaciones genéticas.
2. Célula huésped según la reivindicación 1, en la que la terpeno sintasa se selecciona entre amorfa-4,11-dieno sintasa; betacariofileno sintasa; germacreno A sintasa; 8-epicedrol sintasa; valenceno sintasa; (+)-delta-cadineno sintasa; germacreno C sintasa; (E) beta-farneseno sintasa; vetispiradieno sintasa; 5-epiaristolqueno sintasa; aristolqueno sintasa; alfa-humuleno sintasa; (E,E)-alfa-farneseno sintasa; E-alfa-bisaboleno sintasa; (E)-gamma-bisaboleno sintasa; longifoleno sintasa, gamma-humuleno sintasa, Delta-selineno sintasa, epi-cedrol sintasa; alfa-zingibereno sintasa; guaiadieno sintasa; cascariladieno sintasa; cis-muroladieno sintasa; pachulol sintasa.
- 25 3. Método de producción de un compuesto precursor de isoprenoide o compuesto isoprenoide, cuyo método comprende cultivar una célula huésped, según cualquiera de las reivindicaciones anteriores, y aislar el compuesto precursor de isoprenoide o compuesto isoprenoide del medio de cultivo.

30

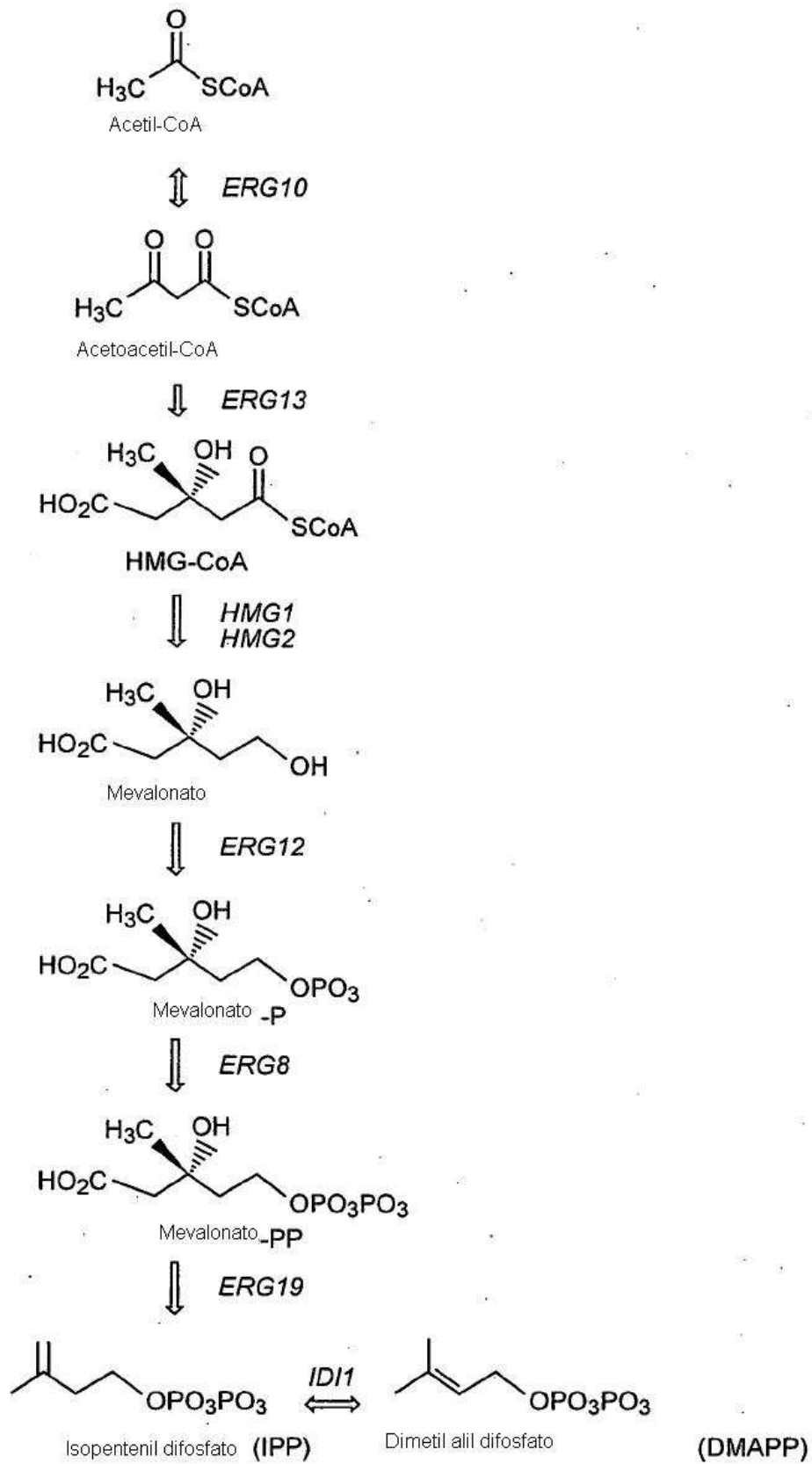


FIG. 1

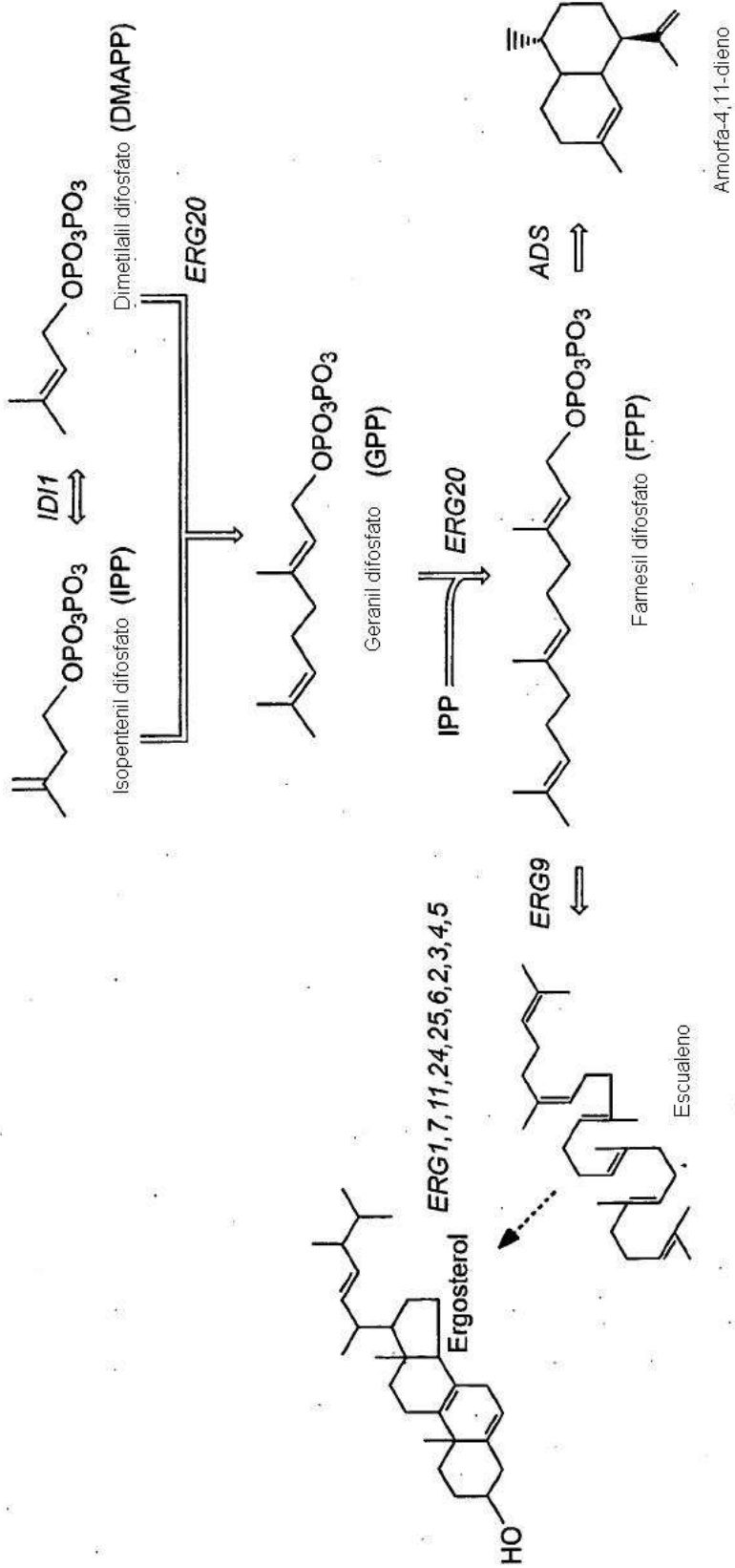


FIG. 2

FIG. 3B

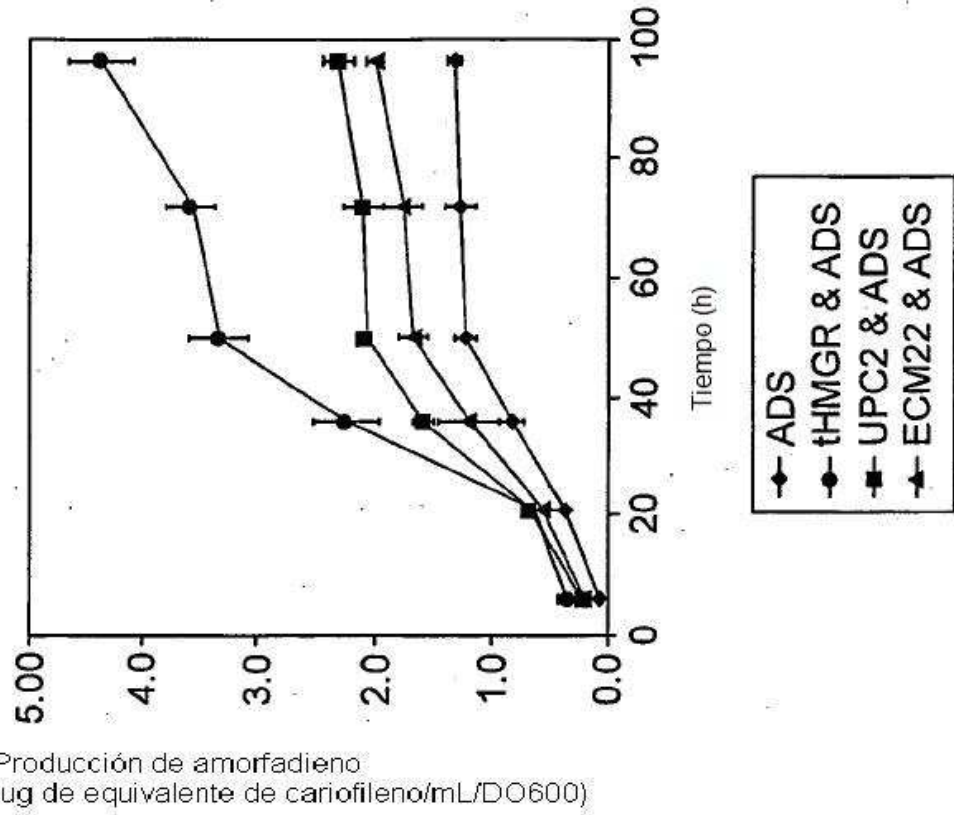
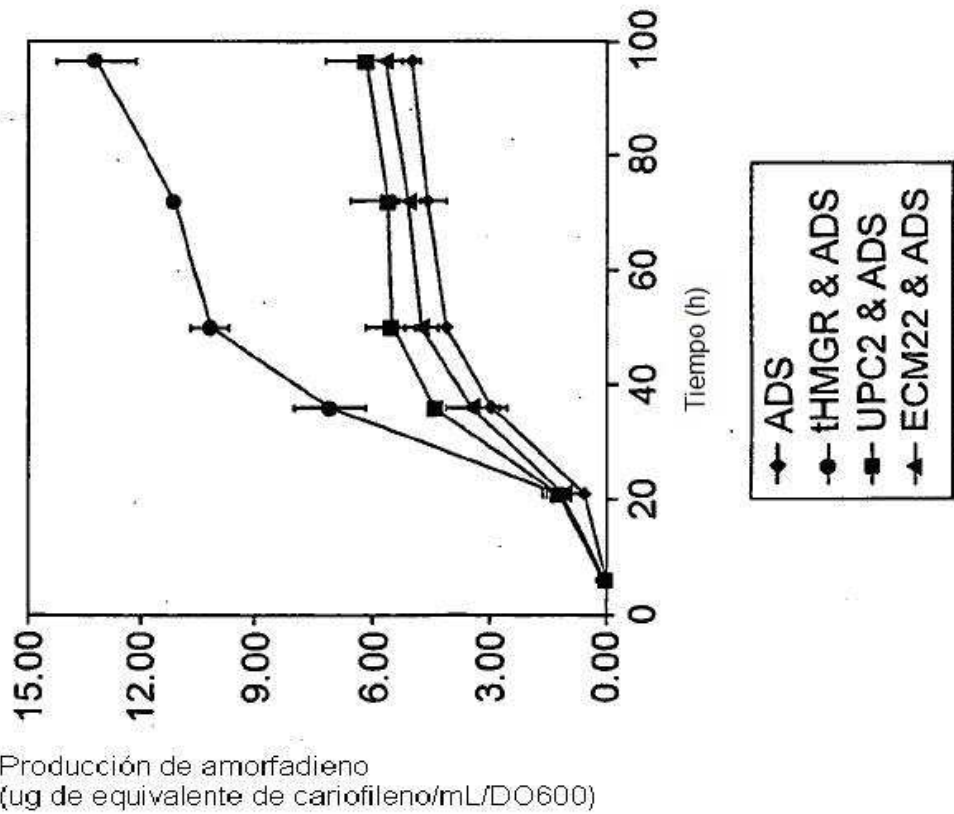


FIG. 3A





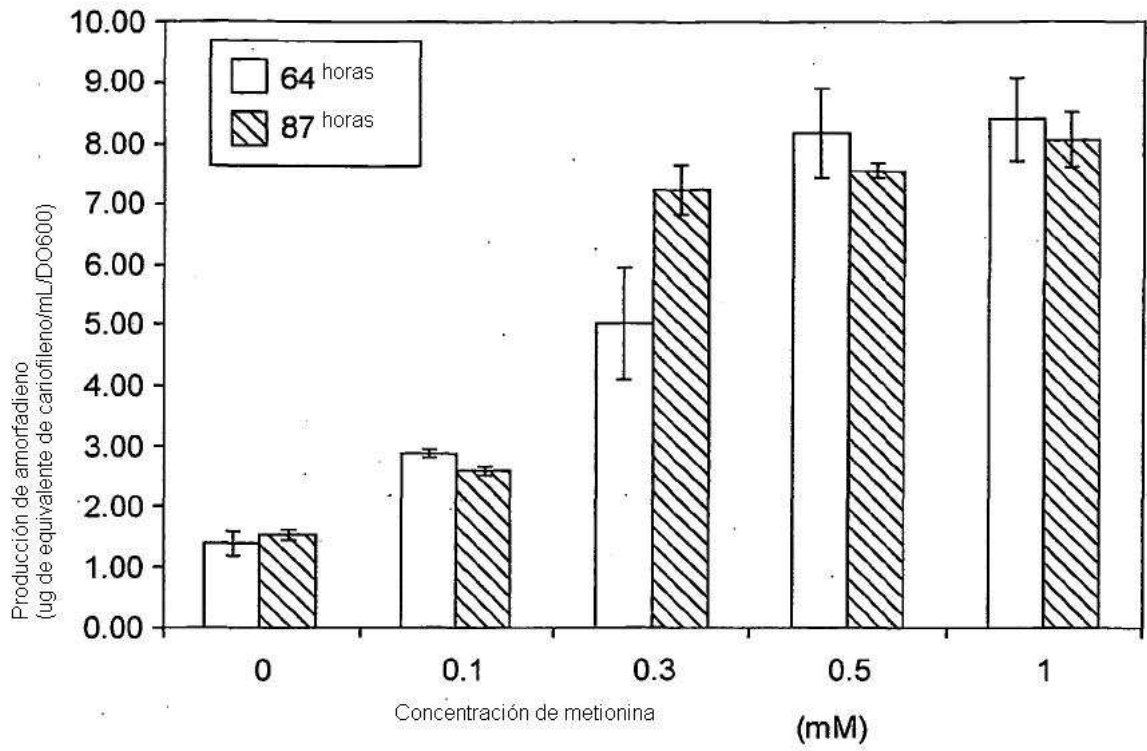


FIG. 4

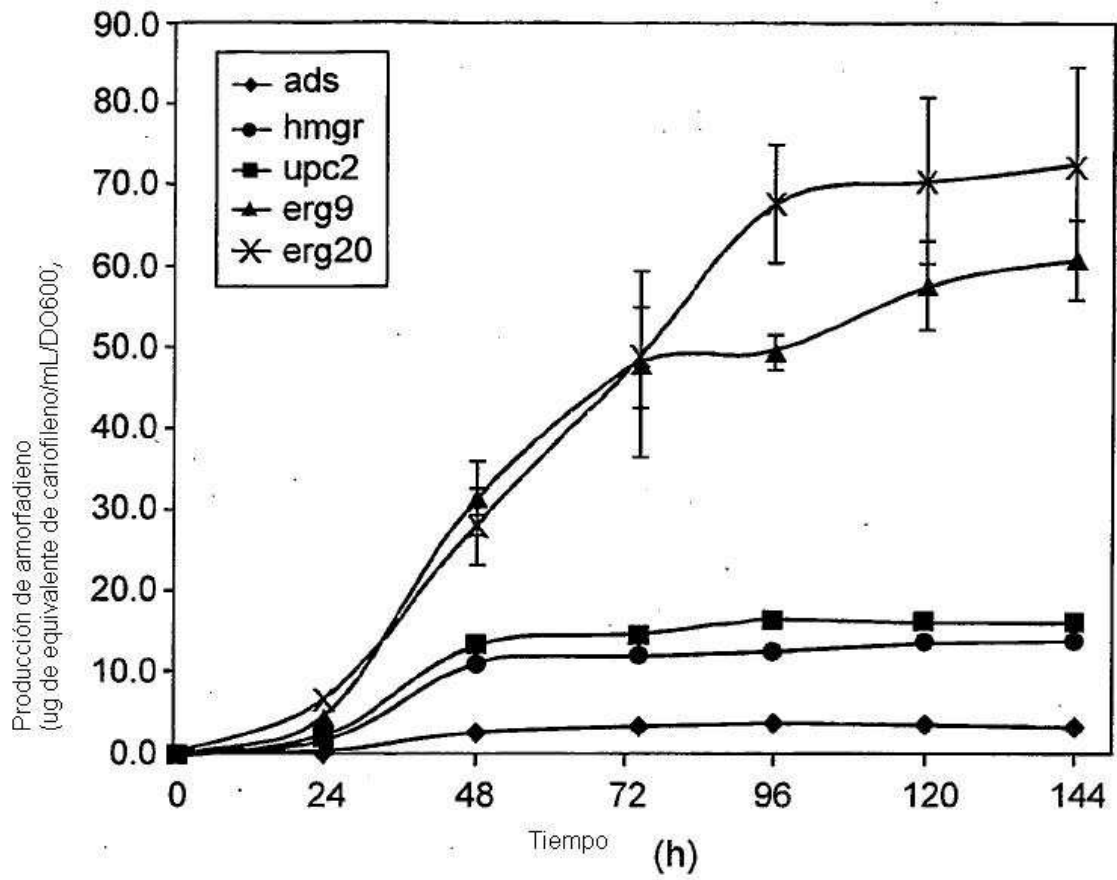


FIG. 5

## FIG. 6

Secuencia codificante de HMGR truncada

ATGGTTTTAACCAATAAAACAGTCATTTCTGGATCGAAAGTCAAAGTTTATCATCTGCGCAATCGAGCTC  
 ATCAGGACCTTCATCATCTAGTGAGGAAGATGATTCCCGCGATATTGAAAGCTTGGATAAGAAAATACGTC  
 CTTTAGAAGAATTAGAAGCATTATTAAGTAGTGGAAATACAAAACAATTGAAGAACAAGAGGTCGCTGCC  
 TTGGTTATTCACGGTAAGTTACCTTTGTACGCTTTGGAGAAAAAATTAGGTGATACACGAGAGCGGTTGC  
 GGTACGTAGGAAGGCTCTTTCAATTTTGGCAGAAGCTCCGTGATTAGCATCTGATCGTTTACCAATATAAAA  
 ATTATGACTACGACCGCGTATTTGGCGCTTGTGTGAAAATGTTATAGGTTACATGCCTTTGCCCGTTGGT  
 GTTATAGGCCCTTGGTTATCGATGGTACATCTTATCATATACCAATGGCAACTACAGAGGGTGTGTTGGT  
 AGCTTCTGCCATGCGTGGCTGTAAGGCAATCAATGCTGGCGGTGGTGCAACAACGTTTTTAACTAAGGATG  
 GTATGACAAGAGGCCAGTAGTCCGTTCCCAACTTTGAAAAGATCTGGTGCCGTGTAAGATATGGTTAGAC  
 TCAGAAGAGGGACAAAACGCAATTA AAAAAGCTTTTAACTCTACATCAAGATTTGCACGCTCGCAACATAT  
 TCAAACCTTGCTAGCAGGAGATTTACTCTTCATGAGATTTAGAACAAC TACTGGTGACGCAATGGGTATGA  
 ATATGATTTCTAAAGGTGTGTAATACTCATTAAAGCAAATGGTAGAAGAGTATGGCTGGGAAGATATGGAG  
 GTTGTCTCCGTTTCTGGTAACTACTGTACCGACAAAAACCAGCTGCCATCAACTGGATCGAAGGTCGTGG  
 TAAGAGTGTGTCGCGAGAAGCTACTATTCCTGGTGATGTTGTCAGAAAAGTGTAAAAAGTGATGTTCCG  
 CATTGGTTGAGTTGAACATTCGTAAGAATTTGGTTGGATCTGCAATGGCTGGGTCTGTTGGTGGATTTAAC  
 GCACATGCAGCTAATTTAGTGACAGCTGTTTTCTTGGCATTAGGACAAGATCCTGCACAAAATGTTGAAAG  
 TTCCAACGTATAACATTGATGAAAGAAGTGGACGGTGATTTGAGAATTTCCGTATCCATGCCATCCATCG  
 AAGTAGGTACCATCGGTGGTGGTACTGTTCTAGAACCAAGGTGCCATGTTGGACTTATTAGGTGTAAGA  
 GGCCCGCATGCTACCGCTCCTGGTACCAACGCACGTCAATTAGCAAGAATAGTTGCCCTGTGCCGCTTTGGC  
 AGGTGAATTATCCTTATGTGCTGCCCTAGCAGCCGGCCATTTGGTTCAAAGTCATATGACCCACAACAGGA  
 AACCTGCTGAACCAACAAAACCTAACAAATTTGGACGCCACTGATATAAAATCGTTTGAAAGATGGGTCCGTC  
 ACCTGCATTAAATCCTAA (SEQ ID NO: 1)

## FIG. 7A

Met	Val	Leu	Thr	Asn	Lys	Thr	Val	Ile	Ser	Gly	Ser	Lys	Val	Lys	Ser
1				5					10					15	
Leu	Ser	Ser	Ala	Gln	Ser	Ser	Ser	Ser	Gly	Pro	Ser	Ser	Ser	Ser	Glu
			20					25					30		
Glu	Asp	Asp	Ser	Arg	Asp	Ile	Glu	Ser	Leu	Asp	Lys	Lys	Ile	Arg	Pro
	35						40					45			
Leu	Glu	Glu	Leu	Glu	Ala	Leu	Leu	Ser	Ser	Gly	Asn	Thr	Lys	Gln	Leu
	50					55					60				
Lys	Asn	Lys	Glu	Val	Ala	Ala	Leu	Val	Ile	His	Gly	Lys	Leu	Pro	Leu
65					70					75					80
Tyr	Ala	Leu	Glu	Lys	Lys	Leu	Gly	Asp	Thr	Thr	Arg	Ala	Val	Ala	Val
				85					90					95	
Arg	Arg	Lys	Ala	Leu	Ser	Ile	Leu	Ala	Glu	Ala	Pro	Val	Leu	Ala	Ser
			100					105					110		
Asp	Arg	Leu	Pro	Tyr	Lys	Asn	Tyr	Asp	Tyr	Asp	Arg	Val	Phe	Gly	Ala
		115					120					125			
Cys	Cys	Glu	Asn	Val	Ile	Gly	Tyr	Met	Pro	Leu	Pro	Val	Gly	Val	Ile
	130					135						140			
Gly	Pro	Leu	Val	Ile	Asp	Gly	Thr	Ser	Tyr	His	Ile	Pro	Met	Ala	Thr
145					150					155					160
Thr	Glu	Gly	Cys	Leu	Val	Ala	Ser	Ala	Met	Arg	Gly	Cys	Lys	Ala	Ile
				165					170					175	
Asn	Ala	Gly	Gly	Gly	Ala	Thr	Thr	Val	Leu	Thr	Lys	Asp	Gly	Met	Thr
			180					185					190		
Arg	Gly	Pro	Val	Val	Arg	Phe	Pro	Thr	Leu	Lys	Arg	Ser	Gly	Ala	Cys
		195					200					205			
Lys	Ile	Trp	Leu	Asp	Ser	Glu	Glu	Gly	Gln	Asn	Ala	Ile	Lys	Lys	Ala
	210					215					220				
Phe	Asn	Ser	Thr	Ser	Arg	Phe	Ala	Arg	Leu	Gln	His	Ile	Gln	Thr	Cys
225					230					235					240
Leu	Ala	Gly	Asp	Leu	Leu	Phe	Met	Arg	Phe	Arg	Thr	Thr	Thr	Gly	Asp
				245					250					255	
Ala	Met	Gly	Met	Asn	Met	Ile	Ser	Lys	Gly	Val	Glu	Tyr	Ser	Leu	Lys
			260					265					270		
Gln	Met	Val	Glu	Glu	Tyr	Gly	Trp	Glu	Asp	Met	Glu	Val	Val	Ser	Val
		275					280					285			
Ser	Gly	Asn	Tyr	Cys	Thr	Asp	Lys	Lys	Pro	Ala	Ala	Ile	Asn	Trp	Ile
	290					295					300				
Glu	Gly	Arg	Gly	Lys	Ser	Val	Val	Ala	Glu	Ala	Thr	Ile	Pro	Gly	Asp
305					310					315					320
Val	Val	Arg	Lys	Val	Leu	Lys	Ser	Asp	Val	Ser	Ala	Leu	Val	Glu	Leu
				325					330					335	
Asn	Ile	Ala	Lys	Asn	Leu	Val	Gly	Ser	Ala	Met	Ala	Gly	Ser	Val	Gly
			340					345					350		
Gly	Phe	Asn	Ala	His	Ala	Ala	Asn	Leu	Val	Thr	Ala	Val	Phe	Leu	Ala
		355					360					365			
Leu	Gly	Gln	Asp	Pro	Ala	Gln	Asn	Val	Glu	Ser	Ser	Asn	Cys	Ile	Thr
	370					375					380				
Leu	Met	Lys	Glu	Val	Asp	Gly	Asp	Leu	Arg	Ile	Ser	Val	Ser	Met	Pro
385					390					395					400

FIG. 7B

Ser	Ile	Glu	Val	Gly	Thr	Ile	Gly	Gly	Gly	Thr	Val	Leu	Glu	Pro	Gln
				405					410					415	
Gly	Ala	Met	Leu	Asp	Leu	Leu	Gly	Val	Arg	Gly	Pro	His	Ala	Thr	Ala
			420					425					430		
Pro	Gly	Thr	Asn	Ala	Arg	Gln	Leu	Ala	Arg	Ile	Val	Ala	Cys	Ala	Val
		435					440					445			
Leu	Ala	Gly	Glu	Leu	Ser	Leu	Cys	Ala	Ala	Leu	Ala	Ala	Gly	His	Leu
	450					455					460				
Val	Gln	Ser	His	Met	Thr	His	Asn	Arg	Lys	Pro	Ala	Glu	Pro	Thr	Lys
465					470					475					480
Pro	Asn	Asn	Leu	Asp	Ala	Thr	Asp	Ile	Asn	Arg	Leu	Lys	Asp	Gly	Ser
				485					490					495	
Val	Thr	Cys	Ile	Lys	Ser	(SEQ ID NO: 2)									
			500												