

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 575 790**

51 Int. Cl.:

C12P 15/00 (2006.01)

C12P 17/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.11.2010 E 13150186 (8)**

97 Fecha y número de publicación de la concesión europea: **16.03.2016 EP 2594648**

54 Título: **Ingeniería microbiana para la preparación de productos químicos y farmacéuticos a partir de la vía isoprenoide**

30 Prioridad:

10.11.2009 US 280877 P

30.09.2010 US 388543 P

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

01.07.2016

73 Titular/es:

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY
(50.0%)**

**77 Massachusetts Avenue
Cambridge, MA 02139, US y**

NATIONAL UNIVERSITY OF SINGAPORE (50.0%)

72 Inventor/es:

**AJIKUMAR, PARAYIL K.;
STEPHANOPOULOS, GREGORY y
TOO, HENG PHON**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 575 790 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

Ingeniería microbiana para la preparación de productos químicos y farmacéuticos a partir de la vía isoprenoide

Campo de la invención

La invención se refiere a la producción de uno o más terpenoides a través de la ingeniería microbiana.

5 Antecedentes de la Invención

Taxol y sus análogos estructurales han sido reconocidos como los medicamentos más potentes contra el cáncer y con éxito comercial introducidos en la última década. ¹Taxol fue aislado por vez primera de la corteza del tejo del Pacífico² y los métodos de producción en fase inicial requerían sacrificar de dos a cuatro árboles completamente crecidos para el suministro de una dosis suficiente para un paciente³. La complejidad estructural de Taxol necesitaba una compleja vía de síntesis química que requería 35-51 etapas con un rendimiento más alto de 0,4%^{4,5,6}. Sin embargo, se concibió una vía semi-sintética por la que el compuesto intermedio biosintético bacatina III fue aislado a partir de fuentes vegetales y fue convertido subsiguientemente en Taxol⁷. Si bien este enfoque y los subsiguientes esfuerzos de producción basados en cultivo de células vegetales han disminuido la necesidad de recolectar el tejo, la producción sigue dependiendo de procesos basados en plantas⁸, con las limitaciones acompañantes de la productividad y capacidad de ampliación, y las limitaciones en el número de derivados de Taxol que se pueden sintetizar en búsqueda de fármacos más eficaces^{9,10}.

Sumario de la Invención

Desarrollos recientes en la ingeniería metabólica y la biología sintética ofrecen nuevas posibilidades para la sobreproducción de productos naturales complejos a través de huéspedes microbianos más técnicamente asequibles^{11,12}. Aunque se ha hecho un progreso emocionante en la elucidación del mecanismo de biosíntesis de Taxol en *Taxus*¹³⁻¹⁶, cepas productoras de Taxol comercialmente relevantes han eludido los intentos anteriores destinados a la transferencia de esta compleja maquinaria biosintética en un huésped microbiano^{17,18}. Sin embargo, al igual que con otros productos naturales, la producción microbiana a través de cepas tratadas metabólicamente, ofrece una economía atractiva y un gran potencial para la síntesis de una diversa gama de nuevos compuestos con actividad anti-cáncer y otra actividad farmacéutica^{19,20}.

La vía metabólica para Taxol y sus análogos consiste en una vía *isoprenoide* aguas arriba que es nativa de *E. coli*, y una vía terpenoide aguas abajo heteróloga (Fig. 6). Las vías ácido mevalónico (MVA) o fosfato de metiliteritrol (MEP) aguas arriba pueden producir los dos bloques de construcción comunes, isopentenil pirofosfato (IPP) y dimetilalil pirofosfato (DMAPP), a partir de los cuales se forman Taxol y otros compuestos isoprenoides¹². Estudios recientes han puesto de manifiesto el diseño de las vías de aguas arriba anteriores para apoyar la biosíntesis de isoprenoides heterólogos tales como licopeno y ácido artemisinico²¹⁻²³. La vía taxadieno aguas abajo ha sido reconstruida en *E. coli* pero, hasta la fecha, los títulos no han excedido de 1,3 mg/L²⁴.

Los enfoques de ingeniería metabólica racional anteriores se centraron en la vía terpenoide aguas arriba (MVA o MEP) o aguas abajo, asumiendo implícitamente que las modificaciones son aditivas, es decir, un comportamiento lineal²⁵⁻²⁷. Si bien este enfoque puede producir incrementos moderados en el flujo, ignora generalmente efectos no específicos tales como la toxicidad de los metabolitos intermedios, efectos celulares de los vectores utilizados para la expresión y las vías desconocidas ocultas que pueden competir con la vía principal y desviar el flujo lejos de la diana deseada. Enfoques combinatorios pueden evitar este tipo de problemas, ya que ofrecen la oportunidad de muestrear adecuadamente el espacio de parámetros y elucidar estas interacciones no lineales complejas^{21,28,29,30}. Sin embargo, requieren una pantalla de alto rendimiento, que a menudo no está disponible para muchos productos naturales deseables³¹. Todavía otra clase de métodos de optimización de la vía ha explorado el espacio combinatorio de diferentes fuentes de los genes heterólogos que comprenden la vía de interés³². Dependientes todavía de un ensayo de alto rendimiento, estos métodos generalmente ignoran la necesidad de determinar un nivel óptimo de expresión para los genes de la vía individual y, como tales, han demostrado ser menos eficaces en la estructuración de una vía óptima.

En el presente trabajo, los autores del mismo se centran en el equilibrio óptimo entre vía de formación de IPP aguas arriba con la vía terpenoide aguas abajo de la síntesis de taxadieno. Esto se logra mediante la agrupación de la vía

de nueve enzimas en dos módulos - un módulo de cuatro genes, aguas arriba, nativo (MEP) y una vía de dos genes, aguas abajo, heteróloga para taxadieno (Fig. 1). Utilizando esta configuración básica, se evalúan parámetros tales como el efecto del número de copias del plásmido en la fisiología celular, el orden de genes y la fuerza del promotor en una casete de expresión y la integración cromosómica con respecto a su efecto sobre la producción de taxadieno. Este enfoque combinatorio modular y multivariable permite a los autores de la invención muestrear de manera eficiente los principales parámetros que influyen en el flujo de la vía, sin necesidad de una pantalla de alto rendimiento. La búsqueda multivariada a través de múltiples promotores y números de copias para cada uno de los módulos de la vía revela un panorama de flujo de taxadieno altamente no lineal con un máximo global que muestra un aumento de 15.000 veces en la producción de taxadieno sobre el control, proporcionando una producción de 300 mg/L de taxadieno en fermentaciones a pequeña escala. Además, los autores de la invención han diseñado la química de oxidación basada en P450 en la biosíntesis de Taxol en *E. coli*, con cepas modificadas por ellos que mejoran la producción de taxadien-5 α -ol 2400 veces sobre el estado de la técnica. Estas mejoras desbloquean el potencial para la producción a gran escala de miles de terpenoides valiosos por parte de sistemas microbianos bien establecidos.

La invención se refiere a un método para aumentar la producción de terpenoides en una célula de *E. coli* que produce uno o más terpenoides, que comprende controlar la acumulación de indol en la célula o en un cultivo de las células a menos de 100 mg/L y sobre-expresar en la célula uno o más componentes de la vía no mevalonato (MEP), aumentando con ello la producción de terpenoides en la célula, en el que la etapa de controlar la acumulación de indol en la célula o en un cultivo de las células comprende (i) equilibrar una vía MEP aguas arriba con una vía de síntesis de terpenoides aguas abajo y/o modificar o regular una vía indol, o (ii) separar del cultivo celular el indol acumulado a través de métodos químicos.

Aspectos de la invención se refieren a métodos que implican expresar de forma recombinante una enzima taxadieno sintasa y una enzima geranylgeranyl difosfato sintasa (GGPPS) en una célula de *E. coli* que sobre-expresa uno o más componentes de la vía no mevalonato (MEP).

En algunas realizaciones, la enzima taxadieno sintasa es una enzima de *Taxus* tal como una enzima de *Taxus brevifolia*. En algunas realizaciones, la enzima GGPPS es una enzima de *Taxus* tal como una enzima de *Taxus canadensis*. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y/o el gen que codifica la enzima GGPPS y/o los genes que codifican uno o más componentes de la vía MEP se expresa a partir de uno o más plásmidos. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y/o el gen que codifica la enzima GGPPS y/o los genes que codifican el uno o más componentes del MEP se incorpora en el genoma de la célula.

En algunas realizaciones, uno o más componentes de la vía no mevalonato (MEP) se seleccionan del grupo que consiste en *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. En determinadas realizaciones se sobre-expresan *dxs*, *idi*, *ispD* e *ispF*. Por ejemplo, *dxs*, *idi*, *ispD* e *ispF* pueden ser sobre-expresados en el operón *dxs-idi-idpDF*. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y el gen que codifica la enzima GGPPS se expresan juntos en un operón.

En algunas realizaciones, la célula expresa, además, una taxadieno 5 α -hidroxilasa (T5 α OH) o una parte catalíticamente activa de la misma. En determinadas realizaciones, la enzima T5 α OH o una parte catalíticamente activa de la misma se fusiona a una enzima citocromo P450 reductasa o a una parte catalíticamente activa de la misma. Por ejemplo, la enzima T5 α OH puede ser At24T5 α OH-tTCPR.

La expresión de la enzima taxadieno sintasa, la enzima GGPPS y el uno o más componentes de la vía MEP puede ser equilibrada para maximizar la producción del taxadieno. Métodos asociados con la invención pueden abarcar, además, cultivar una célula para producir taxadieno o taxadieno-5 α -ol. En algunas realizaciones se producen al menos 10 mg L⁻¹ de taxadieno. En determinadas realizaciones, se producen al menos 250 mg L⁻¹ de taxadieno. En algunas realizaciones se producen al menos 10 mg L⁻¹ de taxadieno-5 α -ol. En determinadas realizaciones se producen al menos 50 mg L⁻¹ de taxadieno-5 α -ol. En algunas realizaciones, el porcentaje de conversión de taxadieno en taxadieno-5 α -ol y el subproducto 5(12)-Oxa-3(11)-ciclotaxano es al menos 50%, al menos 75% o al menos 95%.

Métodos asociados con la invención pueden comprender, además, recuperar el taxadieno o taxadieno-5 α -ol a partir del cultivo celular. En algunas realizaciones, el taxadieno o taxadieno-5 α -ol se recupera de la fase de gas, mientras que en otras realizaciones, se añade una capa orgánica al cultivo celular, y el taxadieno o taxadieno-5 α -ol se recupera de la capa orgánica.

Aspectos de la invención se refieren a células de *E. coli* que sobre-expresan uno o más componentes de la vía no mevalonato (MEP), y que expresan de manera recombinante una enzima sintasa taxadieno y una enzima geranylgeranyl difosfato sintasa (GGPPS).

5 En algunas realizaciones, la enzima taxadieno sintasa es una enzima de *Taxus* tal como una enzima de *Taxus brevifolia*. En algunas realizaciones, la enzima GGPPS es una enzima de *Taxus* tal como una enzima de *Taxus canadensis*. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y/o el gen que codifica la enzima GGPPS y/o los genes que codifican el uno o más componentes de la vía MEP se expresa a partir de uno o más plásmidos. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y/o el gen que codifica la enzima GGPPS y/o los genes que codifican el uno o más componentes del MEP se incorpora en el genoma de la
10 célula.

En algunas realizaciones, el uno o más componentes de la vía no mevalonato (MEP) se selecciona del grupo que consiste en *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. En determinadas realizaciones se sobre-expresan *dxs*, *idi*, *ispD* e *ispF*. Por ejemplo, *dxs*, *idi*, *ispD* e *ispF* pueden ser sobre-expresados en el operón *dxs-idi-idpDF*. En algunas realizaciones, el gen que codifica la enzima taxadieno sintasa y el gen que codifica la enzima GGPPS se expresan juntos en un operón. En algunas realizaciones, la expresión de la enzima taxadieno sintasa, la enzima GGPPS y el uno o más componentes de la vía MEP están equilibrados para maximizar la producción del taxadieno.
15

En algunas realizaciones, la célula expresa, además, una taxadieno 5 α -hidroxilasa (T5 α OH) o una parte catalíticamente activa de la misma. En determinadas realizaciones, la enzima T5 α OH o una parte catalíticamente activa de la misma se fusiona a una enzima citocromo P450 reductasa o a una parte catalíticamente activa de la misma. Por ejemplo, la enzima T5 α OH puede ser At24T5 α OH-tTCPR. En algunas realizaciones, la célula produce taxadieno y/o taxadieno-5 α -ol.
20

Se describen métodos para la selección de una célula que exhiben una producción incrementada de un terpenoide, incluyendo la crear u obtener una célula que sobre-expresa uno o más componentes de la vía no mevalonato (MEP), producir terpenoide a partir de la célula, comparar la cantidad de terpenoide producido a partir de la célula con la cantidad de terpenoide producido en una célula control y seleccionar una primera célula mejorada que produce una mayor cantidad de terpenoide que una célula control, en los que una primera célula mejorada que produce una mayor cantidad de terpenoide que la célula control es una célula que exhibe una producción mejorada de terpenoides.
25

La célula utilizada en estos métodos puede expresar de manera recombinante una enzima terpenoide sintasa y/o una enzima geranylgeranyl difosfato sintasa (GGPPS). Los métodos pueden comprender, además, alterar el nivel de expresión de uno o más de los componentes de la vía no mevalonato (MEP), la enzima terpenoide sintasa y/o la enzima geranylgeranyl difosfato sintasa (GGPPS) en la primera célula mejorada para producir una segunda célula mejorada y comparar la cantidad de terpenoide producido a partir de la segunda célula mejorada con la cantidad de terpenoide producido en la primera célula mejorada, en los que una segunda célula mejorada que produce una mayor cantidad de terpenoide que la primera célula mejorada es una célula que exhibe una producción mejorada de terpenoide. La enzima terpenoide sintasa puede ser una enzima taxadieno sintasa. La célula puede expresar de forma recombinante, además, cualquiera de los polipéptidos asociados con la invención.
30
35

Se describen también polipéptidos aislados que comprenden una enzima taxadieno 5 α -hidroxilasa (T5 α OH) o una parte catalíticamente activa de la misma, fusionados a una enzima citocromo P450 reductasa o una parte catalíticamente activa de la misma. En algunas realizaciones, la enzima citocromo P450 reductasa es una citocromo P450 reductasa de *Taxus* (TCPR). En determinadas realizaciones, la taxadieno 5 α -hidroxilasa y TCPR están unidas por un enlazador tal como GSTGS (SEQ ID NO: 50). En algunas realizaciones, la taxadieno 5 α -hidroxilasa y/o TCPR están truncadas para separar la totalidad o parte de la región de transmembrana. En determinadas realizaciones, están truncados 8, 24 ó 42 aminoácidos N-terminales de taxadieno 5 α -hidroxilasa. En determinadas realizaciones, están truncados 74 aminoácidos de TCPR. En algunas realizaciones, un péptido adicional está fusionado a taxadieno 5 α -hidroxilasa. En determinadas realizaciones, el péptido adicional es 17 α hidroxilasa bovina. En determinadas realizaciones, el péptido es MALLLAVF (SEQ ID NO: 51). En determinadas realizaciones, el polipéptido aislado es At24T5 α OH-tTCPR. También se describen moléculas de ácido nucleico que codifican cualquiera de los polipéptidos asociados con la invención y células que expresan de forma recombinante a cualquiera de los polipéptidos asociados con la invención.
40
45
50

Aspectos de la invención se refieren a métodos para aumentar la producción de terpenoides en una célula de *E. coli* que produce uno o más terpenoides. Los métodos incluyen controlar la acumulación de indol en la célula o en un cultivo de las células, aumentando con ello la producción de terpenoides en una célula.

5 En algunas realizaciones, la etapa de controlar la acumulación de indol en la célula o en un cultivo de las células incluye equilibrar la vía isoprenoide no mevalonato aguas arriba con las vías de la síntesis de producto aguas abajo y/o modificar o regular la vía de indol. En otras realizaciones, la etapa de controlar la acumulación de indol en la célula o en un cultivo de las células incluye, o incluye además, separar el indol acumulado de la fermentación a través de métodos químicos, tales como mediante el uso de absorbentes o agentes depuradores.

10 El uno o más terpenoides producidos por la o las células o en el cultivo pueden ser un monoterpenoide, un sesquiterpenoide, un diterpenoide, un triterpenoide o un tetraterpenoide. En determinadas realizaciones, los terpenoides son taxadieno o cualquier precursor de taxol.

15 Aspectos de la invención se refieren a métodos que incluyen medir la cantidad o concentración de indol en una célula que produce uno o más terpenoides o en un cultivo de las células que producen uno o más terpenoides. Los métodos pueden incluir medir la cantidad o concentración de indol dos o más veces. En algunas realizaciones, la cantidad o la concentración medida de indol se utiliza para guiar un proceso de producción de uno o más terpenoides. En algunas realizaciones, la cantidad o concentración medida de indol se utiliza para guiar la construcción de la cepa.

Estos y otros aspectos de la invención, así como diversas formas de realización de la misma, resultarán más evidentes con referencia a los dibujos y la descripción detallada de la invención.

20 **Breve descripción de los dibujos**

Los dibujos adjuntos no están destinados a ser dibujados a escala. En los dibujos, cada uno de los componentes idénticos o casi idénticos que se ilustra en diversas figuras se representa mediante un número similar. Por motivos de claridad, no todos los componentes pueden ser marcados en cada uno de los dibujos. En los dibujos:

25 **Figura 1. La ingeniería genética de la vía isoprenoide multivariable modular revela respuesta no lineal fuerte en la acumulación de terpenos.** Para aumentar el flujo a través de la vía MEP aguas arriba, los autores de la invención fijaron como objetivo etapas enzimáticas de cuello de botella (dxs, idi, ispD e ispF) para la sobre-expresión por parte de un operón (dxs-idi-ispDF)²⁸. Para canalizar el flujo de rebose de los precursores isoprenoides universales, IPP y DMAPP, hacia la biosíntesis de Taxol, se construyó un operón sintético de genes aguas abajo GGPP sintasa (G) y taxadieno sintasa (T)¹⁶. Las vías isoprenoide aguas arriba y taxadieno sintético aguas abajo se colocaron bajo el control de promotores inducibles para controlar su expresión génica relativa. (a) Esquema de los dos módulos, la vía isoprenoide MEP aguas arriba nativa (izquierda) y vía taxadieno sintético (derecha). En la red de biosíntesis de *E. coli*, la vía de isoprenoide MEP se inicia mediante la condensación de los precursores gliceraldehídos-3 fosfato (G3P) y piruvato (PYR) de la glucólisis. La bifurcación de la vía Taxol parte de los precursores isoprenoides universales IPP y DMAPP para formar primero el precursor geranylgeranyl difosfato "lineal", y luego el taxadieno "cíclico", un compuesto intermedio comprometido y clave para Taxol. La olefina taxadieno cíclica se somete a múltiples rondas de oxidaciones, acilaciones y benzoilación estereoespecíficas con el conjunto de la cadena lateral para formar, en última instancia, Taxol. (b) Esquema del enfoque de la ingeniería genética de la vía isoprenoide multivariable-modular para sondear la respuesta no lineal en la acumulación de terpenoide de células tratadas mediante ingeniería de la vía de aguas arriba y aguas abajo. La expresión de las vías aguas arriba y aguas abajo se modula variando la potencia del promotor (Trc, T5 y T7) o aumentando el número de copias utilizando diferentes plásmidos. La variación de la expresión de la vía de aguas arriba y aguas abajo da diferentes máximos en la acumulación de taxadieno.

45 **Figura 2. Optimización de la producción de taxadieno mediante la regulación de la expresión de las vías modulares aguas arriba y aguas abajo.** (a) Respuesta en la acumulación de taxadieno al aumento de las potencias de la vía aguas arriba para valores constantes de la vía aguas abajo. (b) La dependencia de la vía aguas abajo para aumentos constantes en la potencia de la vía aguas arriba. Los máximos locales múltiples observados en respuesta a taxadieno dependen del aumento de la intensidad de expresión de la vía aguas arriba o aguas abajo. (c) La respuesta a taxadieno por parte de las cepas tratadas mediante ingeniería (17-24) con altas sobre-expresiones de la vía aguas arriba (20-100) con dos expresiones aguas abajo diferentes (~ 30 y ~ 60) para identificar la respuesta a taxadieno con expresiones equilibradas. La expresión de la vía aguas abajo a partir del plásmido con bajo número de copias (p5 y p10) bajo un fuerte operón T7TG promotor se utiliza para modular estas expresiones.

Obsérvese que tanto la vía aguas arriba como la vía aguas abajo expresadas a partir de diferentes plásmidos con diferentes promotores pueden imponer una carga metabólica portada por el plásmido. (d) La modulación de la vía aguas arriba con el aumento de la fuerza del promotor del cromosoma con dos expresiones aguas abajo diferentes ~30 y ~60) para identificar el espacio de búsqueda que falta con efectos tóxicos reducidos (cepas **25-32**). (e) Datos genéticos de las cepas productoras de taxadieno. Los números correspondientes a diferentes cepas y su correspondiente genotipo, E- K12mG1655 Δ recA Δ endA de *E. coli*, EDE3- K12mG1655 Δ recA Δ endA de *E. coli* con la construcción DE3 de T7 ARN polimerasa en el cromosoma, MEP - operón dxs-idi-ispDF, operón GT- GPPS-TS, operón TG-TS-GPPS, Ch1 - 1 copia en el cromosoma, Trc - promotor Trc, T5 - promotor T5, T7 - promotor T7, p5, p10, p20 – plásmido de ~ 5 (SC101), ~ 10 (p15) y ~ 20 (pBR322) copias.

Figura 3. El metabolito se correlaciona inversamente con la producción de taxadieno. (a) Espectro de masas de metabolito que se detectó para correlacionarse inversamente con la producción de taxadieno en las construcciones de cepa de la Fig. 2. Los picos característicos observados del metabolito son 233, 207, 178, 117, 89 y 62. (b) La correlación entre el subproducto de isoprenoides de la Fig. 3a y taxadieno. Las cepas **26-29** y **30-32**, todas con expresión vía aguas arriba cromosómicamente integrada fueron elegidas para una comparación consistente. En las cepas **26-29** y **30-32**, la expresión aguas arriba aumentó cambiando los promotores de Trc a T5 y T7, respectivamente. Los dos conjuntos de cepas difieren sólo en la expresión de la vía de aguas abajo con el segundo conjunto (**30-32**) que tiene dos veces el nivel de expresión de la primera. Con el primer conjunto, el equilibrio óptimo se logra con la cepa **26**, que utiliza el promotor Trc para la expresión de la vía de aguas arriba y también muestra la acumulación de metabolitos más baja. Con las cepas **30-32**, la cepa **31** muestra la acumulación de metabolito más bajo y la más alta producción de taxadieno. Los datos demuestran la correlación inversa observada entre el metabolito desconocido y la producción de taxadieno.

Figura 4. Niveles de expresión génica transcripcional de la vía de aguas arriba y aguas abajo y cambios en la fisiología celular de cepas tratadas mediante ingeniería. La expresión relativa de los primeros genes en el operón de la vía de aguas arriba (DXS) y de aguas abajo (TS) se cuantifica mediante qPCR. Se observaron perfiles de expresión similares con los genes aguas abajo de los operones. Los números de cepa correspondientes se muestran en el gráfico. (a) Expresión génica DXS del nivel de transcrito relativa cuantificada a partir de diferentes expresiones aguas arriba moduladas utilizando promotores y plásmidos bajo dos expresiones diferentes aguas abajo. (b) Expresión génica TS del nivel de transcrito relativa cuantificada a partir de dos diferentes expresiones aguas arriba moduladas utilizando plásmidos p5T7 y p10T7 bajo expresiones diferentes aguas arriba. El análisis de expresión génica de los autores de la invención soportaba directamente la hipótesis, con un aumento en el número de copias del plásmido (5, 10 y 20) y la fuerza del promotor (Trc, T5 y T7), la expresión de las vías de aguas arriba y aguas abajo puede ser modulada. (c) Crecimiento celular de las cepas 25-29 tratadas mediante ingeniería. El fenotipo de crecimiento se vio afectado por la activación del metabolismo de isoprenoides (cepa 26), la expresión de proteína recombinante (cepa 25) y la carga metabólica portada por el plásmido (cepas control frente a tratadas mediante ingeniería) y (d) fenotipos de crecimiento de las cepas 17, 22, 25-32. Las líneas de color negro son las cepas tratadas mediante ingeniería productoras de taxadieno y de las líneas de color gris son las cepas de control sin expresión aguas abajo que portan un plásmido vacío con promotor y sitios de clonación múltiple. El crecimiento fue correlacionado con la activación del metabolismo de terpenoides, la carga metabólica portada por el plásmido, así como la expresión de la proteína recombinante.

Figura 5. Ingeniería genética de la química de oxidación de Taxol p450 en *E. coli*. (a) Esquema de la conversión de taxadieno en taxadieno 5 α -ol en Taxol. (b) La ingeniería genética y la construcción de la transmembrana de la proteína quimera de un componente a partir de taxadieno 5 α -ol hidroxilasa (T5 α OH) y la reductasa citocromo P450 de Taxus (TCPR). **1** y **2** representan las proteínas de longitud completa de T5 α OH y TCPR identificadas con regiones TM de 42 y 74 aminoácidos, respectivamente, **3** - enzimas quimera generadas a partir de las tres diferentes construcciones de T5 α OH manipuladas genéticamente con TM, (At8T5 α OH, At24T5 α OH y At42T5 α OH) construidas mediante la fusión de péptido sintético de 8 residuos (**A**) a T5 α OH truncado con 8, 24 y 42 AA) a través de una fusión translacional con TCPR truncado (tTCPR) de 74 AA utilizando péptido enlazador GSTGS de 5 residuos. (c) Actividad funcional de las construcciones At8T5 α OH-tTCPR, At24T5 α OH-tTCPR y At42T5 α OH-tTCPR transformadas en la cepa 18 productora de taxadieno. (d) Perfil del curso en el tiempo de acumulación de taxadien-5 α -ol y perfil de crecimiento de la cepa 18-At24T5 α OH-tTCPR fermentada en un biorreactor de 1 L.

Figura 6. Esquema biosintético para la producción de taxol en *E. coli*. Esquemas de los dos módulos, vía isoprenoide nativa aguas arriba (izquierda) y vía Taxol sintética (derecha). En la red de biosíntesis de *E. coli*, la divergencia de la vía MEP isoprenoide se inicia a partir de los precursores de gliceraldehído-3 fosfato (G3P) y Piruvato (PYR) de la glucólisis (I-V). La bifurcación de la vía Taxol parte del precursor isoprenoide IPP y DMAPP de *E. coli* a precursor Geranylgeranyl difosfato (VIII) "lineal", taxadieno (IX) "cíclico", taxadieno 5 α -ol (X) "oxidado" a múltiples rondas de oxidaciones, acilaciones, benzoylaciones y epoxidación estereoespecíficas de precursor

temprano Bacatina III (XII) y finalmente con el ensamblaje de la cadena lateral a Taxol (XIII). DXP - 1-desoxi-D-xilulosa-5-fosfato, MEP-2C-metil-D-eritritol-4-fosfato, CDP-ME - 4-difosfocitidil-2C-metil-D-eritritol, CDP-MEP - 4-difosfocitidil- 2C-metil-D-eritritol-2-fosfato, ME-cPP - 2C-metil-D-eritritol-2,4-ciclodifosfato, IPP- isopentenil difosfato, DMAPP - dimetilalil difosfato. Los genes implicaban vías biosintéticas de G3P y PYR a Taxol. DXS-1-desoxi-D-xilulosa-5-fosfato sintasa, ispC-1-desoxi-D-xilulosa-5-fosfato reductoisomerasa, IspD-4-difosfocitidil-2C-metil-D-eritritol sintasa, IspE-4-difosfocitidil 2-C-metil- D-eritritol quinasa, IspF-2C-metil-D-eritritol-2,4-ciclodifosfato Sintasa, IspG-1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintasa, IspH-4-hidroxi-3-metil-2-(E)-butenil-4-difosfato reductasa, IDI-isopentenil-difosfato isomerasa, GGPPS-geranil geranildifosfato sintasa, Taxadieno sintasa, Taxoide 5 α -hidroxilasa, Taxoide-5 α -O-acetiltransferasa, Taxoide 13 α -hidroxilasa, Taxoide 10 β -hidroxilasa, Taxoide 2 α -hidroxilasa, Taxoide 2-O-benzoiltransferasa, Taxoide 7 β -hidroxilasa, Taxoide 10-O-acetiltransferasa, Taxoide 1 β -hidroxilasa*, Taxoide 9 α -hidroxilasa, Taxoide 9-ceto-oxidasa*, Taxoide C4,C20- β -epoxidasa*, Fenilalanina aminomutasa, la cadena lateral CoA-ligasa*, Taxoide 13O-fenilpropanoiltransferasa, Taxoide 2'-hidroxilasa*, Taxoide 3'-N-benzoiltransferasa^{216,219*}. Los genes marcados se han de identificar o caracterizar todavía.

Figura 7. Mejoras múltiples en la producción de taxadieno a partir de la búsqueda de la expresión de la vía modular. (a) Respuesta de taxadieno en múltiples mejoras de todos los máximos observados de las Figuras 2a, b, y c en comparación con la cepa 1. Las diferencias de 2,5 veces entre los dos máximos más altos (cepas **17** y **26**) y de 23 veces (cepas **26** y **10**) con el máximo más bajo indica que la falta una respuesta óptima resulta en títulos significativamente más bajos.

Figura 8. Metabolito (a) Correlación entre la acumulación de taxadieno y metabolitos. La acumulación de metabolitos de la cepa manipulada genéticamente está anti-proporcionalmente relacionada de una manera exponencial con la producción de taxadieno. Se determinó que el coeficiente de correlación para esta relación era 0,92. (b) Perfil GC representativo de las cepas 26-28 para demostrar el cambio en la acumulación de taxadieno y metabolito. Los números en el cromatograma 1 y 2 corresponden al pico de metabolito y de taxadieno, respectivamente. (c) El perfil GC-MS de metabolito (1) y taxadieno (2), respectivamente. Los picos característicos observados del metabolito son 233, 207, 178, 117, 89 y 62. m/z de ion característica de taxa-4(20),11,12-dieno 272 (P⁺), 257 (P⁺-CH₃), 229 (P⁺-C₃H₇); 121, 122, 123 (racimo de fragmento del anillo C).⁶⁰ El pico marcado con una estrella es el cariofileno patrón interno.

Figura 9. Perfiles GC-MS y producción de taxadieno/taxadien-5 α -ol a partir de la enzima quimera artificial manipulada genéticamente en la cepa 26. (a) El perfil GC del extracto de hexano:éter (8:2) de tres construcciones (A-At8T5 α OH-tTCPR, t24T5 α OH-tTCPR y At42T5 α OH-tTCPR) transferidas a la cepa 26 y fermentadas durante 5 días. Las marcas 1, 2 y 3 en los picos corresponden al taxadieno, taxadien-5 α -ol y 5(12)-Oxa-3(11)-ciclotaxano (OCT), respectivamente. (b) La producción de taxa-4(20),11,12-dien-5 α -ol y OCT cuantificada a partir de las tres cepas. (c) y (d) El perfil GC-MS de taxa-4(20),11,12-dien-5 α -ol y OCT y los picos correspondientes a la fragmentación se comparó con los patrones auténticos e informes previos^{42,47} El análisis por GC-MS confirmó la identidad del espectro de masas al taxa-4(20),11,12-dien-5 α -ol auténtico con m/z de ion característica 288(P⁺), 273 (P⁺H₂O), 255 (P⁺-H₂O-CH₃).

La **Figura 10** presenta un esquema que representa la vía biosintética terpenoide y productos naturales producidos por esta vía.

La **Figura 11** presenta un esquema que representa la modulación de la vía aguas arriba para amplificar la producción de taxadieno.

La **Figura 12** presenta un esquema que representa la modulación de la vía aguas abajo para amplificar la producción de taxadieno.

La **Figura 13** presenta un esquema que indica que la desviación de la vía recientemente identificada no es la característica de vía de síntesis aguas abajo.

Figura 14. La intensidad de la vía se correlaciona con los niveles de expresión génica transcripcional. (c) Expresión relativa de genes idi, ispD e ispF con el aumento de la intensidad de la vía aguas arriba y la intensidad de la vía aguas abajo en 31 unidades arbitrarias, y (d) expresión relativa de los genes idi, ispD e ispF con el aumento de la intensidad de la vía aguas arriba y la intensidad de la vía aguas abajo a 61 unidades arbitrarias. Como era de esperar la expresión génica aumentado a medida que aumentaba la intensidad de la vía aguas arriba. Los números de cepa correspondientes se indican en el gráfico de barras. La expresión relativa se cuantificó utilizando la expresión del gen de referencia *RRSA*. Los datos son la media +/- SD de cuatro repeticiones.

Figura 15. Impacto de la acumulación de metabolito subproducto indol en la producción y el crecimiento de taxadieno. (a) Correlación inversa entre taxadieno e indol. Las cepas 26 a 28 y 30 a 32, todas con expresión cromosómicamente integrada de la vía de aguas arriba, fueron elegidas para la comparación consistente. Los dos conjuntos de cepas difieren sólo en la expresión de la vía de aguas abajo, teniendo el segundo conjunto (30 a 32) dos veces el nivel de expresión del primero. En las cepas 26 a 28 y 30 a 32, la expresión aguas arriba aumentaba al cambiar los promotores de Trc a T5 y T7, respectivamente. Con el primer conjunto, el equilibrio óptimo se logra con la cepa 26, que utiliza el promotor Trc para la expresión de la vía de aguas arriba y también muestra la acumulación de indol más baja. Con las cepas 30 a 32, la cepa 31 muestra la acumulación de indol más baja y la producción de taxadieno más alta. Las mejoras múltiples son en relación con las cepas 25 y 29, respectivamente, para los dos conjuntos. (b) Efecto de indol introducido externamente sobre la producción de taxadieno para la cepa 26 de alta producción. Diferentes concentraciones de indol se introdujeron en cultivos de células cultivadas en medio mínimo con extracto de levadura al 0,5%. La producción de taxadieno se redujo significativamente a medida que la concentración de indol aumentaba de 50 mg/L a 100 mg/L. (c) Efecto de indol introducido externamente sobre el crecimiento celular para las cepas manipuladas genéticamente de *E. coli*. Los datos son la media +/- SD de tres repeticiones. Se seleccionaron cepas carentes de la vía de aguas abajo y con diferentes intensidades de la vía de aguas arriba (1, 2, 6, 21, 40 y 100). La cepa 26, el alto productor de taxadieno, exhibe la inhibición más fuerte.

Figura 16. Metabolito desconocido identificado como indol. (A) y (a) Cromatograma de gases y espectro de masas del metabolito desconocido extraído utilizando hexano del cultivo celular. (B) y (b) corresponden al cromatograma de gases y al espectro de masas de indol puro disuelto en hexano. Además de para confirmar la identidad química, el metabolito se extrajo a partir del caldo de fermentación utilizando la extracción con hexano y se purificó por cromatografía en columna de sílice utilizando hexano:acetato de etilo (8:2) como eluyente. La pureza del compuesto se confirmó por TLC y GC-MS. Los espectros ¹H-RMN y ¹³C-RMN confirmaron la identidad química del metabolito como indol. (c) El espectro de ¹H-RMN de indol extraído de cultivo celular (CDCl₃, 400 MHz) δ: 6,56 (d, 1H, Ar C-H), 7,16 (m, 3H, Ar C-H), 7,38 (d, 1H, Ar C-H), 7,66 (d, 1H, Ar C-H), 8,05 (b, 1H, indol NH). (d) ¹³C-RMN δ: 135,7, 127,8, 124,2, 122, 120,7, 119,8, 111, 102,6. (e) es el espectro de ¹H-RMN de indol puro.

Figura 17. Cultivo alimentado discontinuo de cepas manipuladas genéticamente en biorreactor de 1L. Cursos en el tiempo de la acumulación de taxadieno (a), el crecimiento celular (b), la acumulación de ácido acético (c) y la adición total del sustrato (glicerol) (d) para las cepas 22, 17 y 26 durante 5 días de cultivo alimentado discontinuo en el biorreactor en recipientes del biorreactor de 1 L bajo condiciones controladas de pH y oxígeno con medios mínimos y extracto de levadura al 0,5%. Después de que el glicerol se agota a ~ 0,5 a 1 g/L en el fermentador, se introdujeron en el biorreactor 3 g/L de glicerol durante la fermentación. Los datos son la media de dos biorreactores de réplica.

Descripción Detallada de la Invención

Taxol es un potente fármaco anti-cancerígeno aislado por primera vez como un producto natural del tejo del Pacífico *Taxus brevifolia*. Sin embargo, la producción fiable y rentable de Taxol o análogos de Taxol por rutas tradicionales de producción de extractos vegetales es limitada. En esta memoria, los autores de la invención presentan un enfoque multivariado-modular para la ingeniería genética de la vía metabólica para amplificar ~15000 veces la producción de Taxadieno en *Escherichia coli* manipulada genéticamente. Taxadieno, el primer compuesto intermedio de Taxol comprometido, es el producto de biosíntesis de la vía no mevalonato en *E. coli* que comprende dos módulos: la vía de aguas arriba nativa formando Isopentenil Pirofosfato (IPP) y una vía de formación de terpenoide de aguas abajo heteróloga. La búsqueda sistemática multivariada identificó condiciones que equilibran de manera óptima los dos módulos de la vía para minimizar la acumulación de productos intermedios inhibitorios y el desvío del flujo a productos secundarios. Los autores de la invención también manipularon genéticamente la siguiente etapa, después de taxadieno, en la biosíntesis de Taxol, una etapa de oxidación basada en P450, que proporcionó > 98% de conversión del sustrato y presentan el primer ejemplo de la producción *in vivo* de cualesquiera productos intermedios de Taxol funcionalizados en *E. coli*. El enfoque de la ingeniería genética de la vía modular no sólo pone de manifiesto la complejidad de las vías multi-etapa, sino que también permitió la acumulación de altos títulos de taxadieno y taxadien-5α-ol (~300 mg/L y 60 mg/L, respectivamente) en fermentaciones a pequeña escala, ejemplificando así el potencial de la producción microbiana de Taxol y sus derivados.

Esta invención no se limita en su aplicación a los detalles de la construcción y la disposición de componentes recogidos en la siguiente descripción o ilustrados en los dibujos. La invención es capaz de otras realizaciones y de ser puesta en práctica o de ser llevada a cabo de diversas maneras. Además, la fraseología y terminología utilizada en esta memoria es con el propósito de descripción y no deben considerarse como limitantes. El uso de "que incluye", "que comprende" o "que tiene", "que contiene", "que implica", y variaciones de los mismos en esta memoria, pretende abarcar los elementos listados en lo sucesivo, así como elementos adicionales.

- La producción microbiana de terpenoides tales como taxadieno se demuestra en esta memoria. Cuando se expresa en niveles satisfactorios, las vías microbianas reducen drásticamente el costo de producción de compuestos de este tipo. Adicionalmente, se utilizan materias primas baratas, abundantes y renovables (tales como azúcares y otros hidratos de carbono) y pueden ser la fuente para la síntesis de numerosos derivados que pueden exhibir propiedades superiores a las del compuesto original. Un elemento clave en la producción de costo competitivo de compuestos de la vía isoprenoide utilizando una vía microbiana es la amplificación de esta vía con el fin de permitir la sobreproducción de estas moléculas. Se describen en esta memoria métodos que potencian o amplifican el flujo hacia la producción de terpenoides en *Escherichia coli* (*E. coli*). Específicamente, se proporcionan métodos para amplificar el flujo metabólico a la síntesis de pirofosfato de isopentenilo (IPP) (un compuesto intermedio clave para la producción de compuestos isoprenoides), dimetilalil pirofosfato (DMAPP), geranil difosfato (GPP), farnesil difosfato (FPP), geranilgeranil difosfato (GGPP) y farnesil geranil difosfato (FGPP), paclitaxel (Taxol), ginkolidas, geraniol, farnesol, geranilgeraniol, linalool, isopreno, monoterpénoides tales como mentol, carotenoides tales como licopeno, poliisoprenoides tales como poliisopreno o caucho natural, diterpenoides tales como eleuterobina y sesquiterpenoides tales como artemisinina.
- Aspectos de la invención se refieren a la producción de terpenoides. Tal como se utiliza en esta memoria, un terpenoide, al que también se alude como un isoprenoide, es un compuesto químico orgánico derivado de una unidad de isopreno de cinco carbonos. Varios ejemplos no limitativos de terpenoides, clasificados en base al número de unidades de isopreno que contienen, incluyen: hemiterpenoides (1 unidad de isopreno), monoterpénoides (2 unidades de isopreno), sesquiterpenoides (3 unidades de isopreno), diterpenoides (4 unidades de isopreno), sesterterpenoides (5 unidades de isopreno), triterpenoides (6 unidades de isopreno), tetraterpenoides (8 unidades de isopreno) y politerpenoides con un mayor número de unidades de isopreno. En algunas realizaciones, el terpenoide que se produce es taxadieno. En algunas realizaciones, el terpenoide que se produce es Citronellol, Cubebol, Nootkatona, Cineol, Limoneno, Eleuterobina, Sarcodictina, Pseudopterinas, Ginkgolidas, Esteviósido, Rebaudiósido A, esclareol, labdenodiol, levopimaradieno, sandracopimaradieno o isopemaradieno.
- Se describen en esta memoria métodos y composiciones para optimizar la producción de terpenoides en las células mediante el control de la expresión de genes o proteínas que participan en una vía de aguas arriba y vía de aguas abajo. La vía de aguas arriba implica la producción de isopentil pirofosfato (IPP) y dimetilalil pirofosfato (DMAPP), que se puede conseguir por dos vías metabólicas diferentes: la vía ácido mevalónico (MVA) y la vía MEP (2-C-metil-D-eritritol 4-fosfato), también denominada la vía MEP/DOXP (2-C-metil-D-eritritol 4-fosfato/1-desoxi-D-xilulosa 5-fosfato), la vía no mevalonato o la vía independiente del ácido mevalónico.
- La vía de aguas abajo es una vía sintética que conduce a la producción de un terpenoide e implica la expresión recombinante de genes de una enzima terpenoide sintasa (también conocida como terpeno ciclasa) y una enzima geranilgeranil difosfato sintasa (GGPPS). En algunas realizaciones, una enzima terpenoide sintasa es una enzima diterpenoide sintasa. Varios ejemplos no limitativos de enzimas diterpenoide sintasa incluyen casbeno sintasa, taxadieno sintasa, levopimaradieno sintasa, abietadieno sintasa, isopimaradieno sintasa, *ent*-copalil difosfato sintasa, *syn*-stemar-13-eno sintasa, *syn*-stemod-13(17)-eno sintasa, *syn*-pimara-7,15-dieno sintasa, *ent*-sandaracopimaradieno sintasa, *ent*-cassa-12,15-dieno sintasa, *ent*-pimara-8(14),15-dieno sintasa, *ent*-kaur-15-eno sintasa, *ent*-kaur-16-eno sintasa, afidicolan-16 β -ol sintasa, fillocladan-16 α -ol sintasa, fusicocca-2,10(14)-dieno sintasa y terpenotetrienos ciclasa.
- Sorprendentemente, tal como se demuestra en la sección de Ejemplos, la optimización de la síntesis de terpenoides mediante la manipulación de las vías de aguas arriba y aguas abajo que se describen en esta memoria, no era un proceso lineal simple o aditivo. Más bien, a través de análisis combinatorio complejo, la optimización se logró mediante el equilibrio de los componentes de las vías de aguas arriba y aguas abajo. Inesperadamente, tal como se demuestra en las Figuras 1 y 2, la acumulación de taxadieno exhibió una fuerte dependencia no lineal de las intensidades relativas de las vías MEP aguas arriba y taxadieno aguas abajo sintéticas.
- Aspectos de la invención se refieren al control de la expresión de genes y proteínas en la vía MEP para la producción optimizada de un terpenoide tal como taxadieno. La producción optimizada de un terpenoide se refiere a la producción de una mayor cantidad de un terpenoide después de la búsqueda de una estrategia de optimización que se lograría en ausencia de una estrategia de este tipo. Se debe apreciar que cualquier gen y/o proteína dentro de la vía MEP queda abarcado por los métodos y composiciones descritos en esta memoria. En algunas realizaciones, un gen dentro de la vía MEP es uno de los siguientes: *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* o *ispB*. La expresión de uno o más genes y/o proteínas dentro de la vía MEP puede ser regulada al alza y/o regulada a la baja. En determinadas realizaciones, la regulación positiva de uno o más genes y/o proteínas dentro de la vía MEP se puede combinar con la regulación negativa de uno o más genes y/o proteínas dentro de la vía MEP.

Se debe apreciar que los genes y/o proteínas se pueden regular por sí solos o en combinación. Por ejemplo, la expresión de *dxs* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o la regulación a la baja de la expresión de uno o más de *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispC* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispD* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispE* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispF* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispG* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispH*, *idi*, *ispA* e *ispB*. La expresión de *ispH* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *idi*, *ispA* e *ispB*. La expresión de *idi* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *ispA* e *ispB*. La expresión de *ispA* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi* e *ispB*. La expresión de *ispB* se puede regular positivamente o regular negativamente por sí sola o en combinación con la regulación al alza o regulación a la baja de la expresión de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi* e *ispA*. En algunas realizaciones, la expresión del gen y/o la proteína de uno o más de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH* e *idi* se regula positivamente, mientras que la expresión del gen y/o la proteína de *ispA* y/o *ISPB* se regula negativamente.

La expresión de genes dentro de la vía MEP se puede regular en un método modular. Tal como se utiliza en esta memoria, la regulación mediante un método modular se refiere a la regulación de múltiples genes juntos. Por ejemplo, en algunas realizaciones, múltiples genes dentro de la vía MEP se expresan de forma recombinante en una región contigua de ADN, tal como un operón. Debe apreciarse que una célula que expresa un módulo de este tipo también puede expresar uno o más de otros genes dentro de la vía MEP, ya sea de forma recombinante o endógena.

Un ejemplo no limitante de un módulo de genes dentro de la vía MEP es un módulo que contiene los genes *dxs*, *idi*, *ispD* e *ispF*, tal como se presenta en la sección de Ejemplos, y se alude en esta memoria como *dxs-idi-ispDF*. Se debe apreciar que los módulos de genes dentro de la vía MEP, consistente con aspectos de la invención, pueden contener cualquiera de los genes dentro de la vía MEP, en cualquier orden.

La expresión de genes y proteínas dentro de la vía de síntesis de terpenoide sintético aguas abajo también se puede regular con el fin de optimizar la producción de terpenoides. La vía de síntesis aguas abajo de terpenoide sintético implica la expresión recombinante de una enzima terpenoide sintasa y una enzima GGPPS. Cualquier enzima terpenoide sintasa, tal como se comentó anteriormente, se puede expresar con GGPPS dependiendo del producto aguas abajo a producir. Por ejemplo, taxadieno sintasa se utiliza para la producción de taxadieno. La expresión recombinante de la enzima taxadieno sintasa y la enzima GGPPS se puede regular de forma independiente o en conjunto. En algunas realizaciones, las dos enzimas se regulan juntas de una manera modular. Por ejemplo, las dos enzimas se pueden expresar en un operón en cualquier orden (GGPPS-TS, al que se alude como "GT", o TS-GGPPS, al que se alude como "TG").

La manipulación de la expresión de genes y/o proteínas, incluyendo módulos tales como el operón *dxs-idi-ispDF* y el operón TS-GGPPS, se puede lograr a través de métodos conocidos por un experto ordinario en la técnica. Por ejemplo, la expresión de los genes u operones se puede regular a través de la selección de promotores, tales como promotores inducibles, con diferentes fuerzas. Varios ejemplos no limitantes de promotores incluyen *Trc*, *T5* y *T7*. Adicionalmente, la expresión de genes u operones se puede regular a través de la manipulación del número de copias del gen u operón en la célula. Por ejemplo, en determinadas realizaciones, una cepa que contiene una copia adicional del operón *dxs-idi-ispDF* en su cromosoma bajo el control del promotor *Trc* produce una cantidad incrementada de taxadieno con relación a una que sobre-expresa solamente la vía sintética aguas abajo. En algunas realizaciones, la expresión de genes u operones se puede regular a través de la manipulación del orden de los genes dentro de un módulo. Por ejemplo, en determinadas realizaciones, el cambio del orden de los genes en un operón sintético aguas abajo de GT a TG resulta en un aumento de 2-3 veces en la producción de taxadieno. En algunas realizaciones, la expresión de genes u operones se regula a través de la integración de uno o más genes u

operones en un cromosoma. Por ejemplo, en determinadas realizaciones, la integración del operón *dxs-idi-ispDF* aguas arriba en el cromosoma de una célula resulta en una producción incrementada de taxadieno.

Se debe apreciar que los genes asociados con la invención se pueden obtener de una diversidad de fuentes. En algunas realizaciones, los genes dentro de la vía MEP son genes bacterianos tales como genes de *Escherichia coli*.

5 En algunas realizaciones, el gen que codifica GGPPS es un gen vegetal. Por ejemplo, el gen que codifica GGPPS puede ser de una especie de *Taxus* tal como *Taxus canadensis* (*T. canadensis*). En algunas realizaciones, el gen que codifica taxadieno sintasa es un gen vegetal. Por ejemplo, el gen que codifica taxadieno sintasa puede ser de una especie de *Taxus* tal como *Taxus brevifolia* (*T. brevifolia*). Números de Acceso a GenBank representativos para *T. canadensis* GGPPS y *T. brevifolia* taxadieno sintasa se proporcionan por AF081514 y U48796.

10 Como una persona de experiencia ordinaria en la técnica será consciente, genes homólogos para uso en métodos asociados con la invención se pueden obtener de otras especies, y pueden ser identificados mediante búsquedas de homología, por ejemplo mediante una búsqueda de proteínas BLAST, disponible en el sitio de Internet (www.ncbi.nlm.nih.gov) del Centro Nacional de Información sobre Biotecnología (NCBI). Genes y/u operones asociados con la invención pueden ser clonados, por ejemplo, mediante amplificación por PCR y/o digestión de restricción, a partir de ADN de cualquier fuente de ADN que contiene el gen dado. En algunas realizaciones, un gen y/u operón asociado con la invención es sintético. Cualquier medio de obtener un gen y/u operón asociado con la invención es compatible con la presente invención.

20 En algunas realizaciones, una optimización adicional de la producción de terpenoides se consigue mediante la modificación de un gen antes de que sea expresado de forma recombinante en una célula. En algunas realizaciones, la enzima GGPPS tiene una o más de las siguientes mutaciones: A162V, G140C, L182M, F218Y, D160G, C184S, K367R, A151T, M185I, D264Y, E368D, C184R, L331I, G262V, R365S, A114D, S239C, G295D, I276V, K343N, P183S, I172T, D267G, I149V, T234I, E153D y T259A. En algunas realizaciones, la enzima GGPPS tiene una mutación en el residuo S239 y/o residuo G295. En determinadas realizaciones, la enzima GGPPS tiene la mutación S239C y/o G295D.

25 En algunas realizaciones, la modificación de un gen antes de que sea expresado de forma recombinante en una célula implica la optimización de codones para la expresión en una célula bacteriana. A los usos de codones para una diversidad de organismos se puede acceder en la Base de Datos de Uso de Codones (www.kazusa.or.jp/codon/). La optimización de codones, incluyendo la identificación de codones óptimos para una diversidad de organismos, y métodos para alcanzar la optimización de los codones, es familiar para un experto normal en la técnica, y se puede lograr utilizando métodos estándares.

35 En algunas realizaciones, la modificación de un gen antes de que sea expresado de forma recombinante en una célula implica la realización de una o más mutaciones en el gen antes de que sea expresado de forma recombinante en una célula. Por ejemplo, una mutación puede implicar una sustitución o delección de un solo nucleótido o de varios nucleótidos. En algunas realizaciones, una mutación de uno o más nucleótidos en un gen resultará en una mutación en la proteína producida a partir del gen, tal como una sustitución o delección de uno o más aminoácidos.

40 En algunas realizaciones, puede ser ventajoso utilizar una célula que haya sido optimizada para la producción de un terpenoide. Por ejemplo, en algunas realizaciones, una célula que sobre-expresa uno o más componentes de la vía no mevalonato (MEP) se utiliza, al menos en parte, para amplificar isopentil difosfato (IPP) y dimetilalil difosfato (DMAPP), sustratos de GGPPS. En algunas realizaciones, la sobre-expresión de uno o más componentes de la vía no mevalonato (MEP) se consigue aumentando el número de copias de uno o más componentes de la vía no mevalonato (MEP). Por ejemplo, se puede amplificar los números de copias de componentes en etapas limitantes de la velocidad en la vía MEP tales como (*dxs*, *ispD*, *ispF*, *idi*) mediante expresión episomal adicional.

45 En algunas realizaciones, el "diseño racional" está implicado en la construcción de mutaciones específicas en proteínas tales como enzimas. Tal como se utiliza en esta memoria, "diseño racional" se refiere a la incorporación de los conocimientos de la enzima, o enzimas relacionadas, tales como su estructura tridimensional, su sitio o sitios activos, su sustrato o sustratos y/o la interacción entre la enzima y el sustrato, en el diseño de la mutación específica. Sobre la base de un enfoque de diseño racional, se pueden crear mutaciones en una enzima que puede entonces ser rastreada en cuanto a la producción incrementada de un terpenoide con relación a los niveles de control. En algunas realizaciones, las mutaciones se pueden diseñar de forma racional en base al modelado de homología. Tal como se utiliza en esta memoria, "modelado de homología" se refiere al proceso de construir un modelo de resolución atómica de una proteína a partir de su secuencia de aminoácidos y una estructura tridimensional de una proteína homóloga relacionada.

50

- En algunas realizaciones, se pueden hacer mutaciones aleatorias en un gen, tal como un gen que codifica una enzima, y estas mutaciones se pueden rastrear en cuanto a la producción incrementada de un terpenoide con relación a los niveles de control. Por ejemplo, el rastreo de mutaciones en componentes de la vía MEP, o componentes de otras vías, que conducen a una producción incrementada de un terpenoide puede llevarse a cabo a través de un rastreo aleatorio de mutagénesis, o a través del rastreo de mutaciones conocidas. En algunas realizaciones, la clonación aleatoria de fragmentos genómicos se podría utilizar para identificar regiones genómicas que conducen a un incremento en la producción de un terpenoide, a través del rastreo de células u organismos que tienen estos fragmentos para una producción incrementada de un terpenoide. En algunos casos una o más mutaciones se pueden combinar en la misma célula u organismo.
- 5
- 10 En algunas realizaciones, la producción de un terpenoide en una célula se puede aumentar mediante la manipulación de enzimas que actúan en la misma vía que las enzimas asociadas con la invención. Por ejemplo, en algunas realizaciones puede ser ventajoso aumentar la expresión de una enzima u otro factor que actúa aguas arriba de una enzima diana tal como una enzima asociada con la invención. Esto podría lograrse sobre-expresando el factor de aguas arriba utilizando cualquier método estándar.
- 15 La optimización de la expresión de proteínas también se puede lograr a través de la selección de promotores apropiados y sitios de unión al ribosoma. En algunas realizaciones, esto puede incluir la selección de plásmidos de alto número de copias, o plásmidos de bajo o medio número de copias. La etapa de terminación de la transcripción también se puede fijar como objetivo para la regulación de la expresión génica, a través de la introducción o eliminación de estructuras tales como tallo-bucles.
- 20 Aspectos de la invención se refieren a la expresión de genes recombinantes en células de *E. coli*. Se debe apreciar que células de *E. coli* compatibles con la invención pueden expresar una copia endógena de uno o más de los genes asociados con la invención, así como una copia recombinante. En algunas realizaciones, si una célula tiene una copia endógena de uno o más de los genes asociados con la invención, entonces los métodos no requieren necesariamente la adición de una copia recombinante del o de los genes que se expresan de forma endógena. En algunas realizaciones, la célula puede expresar de manera endógena una o más enzimas de las vías descritas en esta memoria y puede expresar de forma recombinante una o más de otras enzimas de las vías descritas en esta memoria para la producción eficiente de un terpenoide.
- 25
- 30 Se describe el rastreo de células o cepas bacterianas que exhiben una producción optimizada de terpenos. Tal como se describió anteriormente, métodos asociados con la invención implican la generación de células que sobre-expresan uno o más genes en la vía MEP. La producción de terpenoide a partir del cultivo de este tipo de células puede ser medida y comparada con una célula control, en donde una célula que exhibe una mayor cantidad de una producción de terpenoides en relación con una célula control se selecciona como una primera célula mejorada. La célula puede ser modificada, además, mediante la expresión recombinante de una enzima terpenoide sintasa y una enzima GGPPS. El nivel de expresión de uno o más de los componentes de la vía no mevalonato (MEP), la enzima terpenoide sintasa y/o la enzima GGPPS en la célula se puede manipular entonces y la producción de terpenoides se puede medir de nuevo, dando lugar a la selección de una segunda célula mejorada que produce mayores cantidades de un terpenoide que la primera célula mejorada. En algunas realizaciones, la enzima terpenoide sintasa es una enzima taxadieno sintasa.
- 35
- 40 Aspectos adicionales de la invención se refieren a la identificación y caracterización (a través de GC-MS) de un metabolito desconocido previamente en células bacterianas de *E. coli* (Figuras 3 y 6). El nivel de acumulación del metabolito recientemente identificado, indol, se puede controlar mediante la ingeniería genética de la vía microbiana por la sobre-expresión, regulación a la baja o mutación de los genes de la vía isoprenoide. El metabolito indol se anti-correlaciona como una variable directa a la producción de taxadieno en cepas manipuladas genéticamente (Figuras 3, 6 y 15). Un control adicional de la acumulación de indol para mejorar el flujo hacia la biosíntesis de terpenoides en sistemas bacterianos (específicamente en células, tales como células de *E. coli*) u otras células, puede lograrse mediante el equilibrio de la vía isoprenoide no mevalonato aguas arriba con las vías de síntesis de productos aguas abajo o mediante modificaciones a o regulación de la vía indol. De este modo, la persona experta puede reducir o controlar la acumulación de indol y, con ello, puede reducir el efecto inhibitor de indol en la producción de taxadieno y otros terpenoides derivados de las vías descritas, tales como: monoterpenoides, sesquiterpenoides (incluyendo amorfadieno), diterpenoides (incluyendo levopimaradieno), triterpenos y tetraterpenos. Otros métodos para reducir o controlar la acumulación de indol incluyen separar el indol acumulado de la fermentación a través de métodos químicos tales como mediante el uso de absorbentes, agentes depuradores, etc.
- 45
- 50

Se proporcionan métodos que incluyen medir la cantidad o concentración de indol en una célula que produce uno o más terpenoides o en un cultivo de las células que producen una o más terpenoides. La cantidad o concentración de indol se puede medir una vez, o dos o más veces, según sea adecuado, utilizando métodos conocidos en la técnica y tal como se describen en esta memoria. Tales métodos se pueden utilizar para guiar los procesos de producir uno o más terpenoides, p. ej., en la mejora de los procesos. Tales métodos se pueden utilizar para guiar la construcción de la cepa, p. ej., para la mejora de cepas.

La identificación de los medios para lograr este equilibrio proporcionó una mejora de 15.000 veces en la sobreproducción de terpenoides tales como taxadieno, en comparación con las células bacterianas de tipo salvaje, expresada con una vía de biosíntesis de taxadieno heterólogo. La producción se incrementó adicionalmente a través de métodos de fermentación modificados que proporcionaron concentraciones de aproximadamente 2 g/L, que son 1500 veces mayores en comparación con cualquier producción previa de taxadieno informada. Como se demuestra en esta memoria, mediante la modificación genética de la vía isoprenoide no mevalonato en *E. coli* se puede controlar ahora la acumulación de este metabolito que regula el flujo hacia la biosíntesis de isoprenoides en células bacterianas de *E. coli*. También se demuestra en esta memoria la canalización adicional de la producción de taxadieno en el siguiente precursor clave de Taxol, taxadien-5 α -ol, logrado a través de la ingeniería de la química de oxidación para la biosíntesis de Taxol. El Ejemplo 5 presenta la primera extensión con éxito de la vía sintética de taxadieno a taxadien-5 α -ol. De manera similar a la mayoría de otros terpenoides, la biosíntesis de Taxol sigue la forma unificada del proceso de biosíntesis de "dos fases", (i) la "fase ciclasa" de acoplamiento lineal de los precursores de prenil (IPP y DMAPP) a GGPP, seguida de la ciclación molecular y el reordenamiento del precursor comprometido taxadieno (Figura 6, VIII-IX)^{57,58}. Después el precursor comprometido, (ii) la "fase de oxidación", la estructura del núcleo de taxadieno de la olefina cíclica es funcionalizada entonces por siete oxigenasas del citocromo P450, junto con sus participantes redox, decorado con dos grupos acetato y un grupo benzoato mediante acil y aroil transferasas dependientes de CoA, con un grupo ceto mediante ceto-oxidasa y con un grupo epóxido mediante epoxidasa conduce al compuesto intermedio tardío baccatina III, al que se une la cadena lateral C13 para el Taxol ((Figura 6, X-XIII)¹⁵. Aunque se predice un orden secuencial aproximado de las reacciones en fase de oxidación temprana, el momento/orden preciso de algunas de las hidroxilaciones, acilaciones y reacciones de benzoilación es incierto. Sin embargo, es claro que la bifurcación temprana comienza a partir de la hidroxilación del núcleo taxadieno mediada por el citocromo p450 en la posición C5, seguida de las hidroxilaciones aguas abajo utilizando una familia homóloga de enzimas del citocromo p450 con alta similitud deducida entre sí (> 70%), pero con semejanza limitada (<30%) con otros p450s de vegetales^{41,59}. Además, la diversidad estructural y funcional con el posible análisis evolutivo implícita que el gen taxadieno-5 α -ol puede ser la secuencia parental de la que evolucionaron los otros genes hidroxilasa en la vía de biosíntesis de Taxol, reflejando el orden de hidroxilaciones¹⁵.

La expresión funcional del citocromo P450 de vegetales ha sido considerada difícil debido a las limitaciones inherentes de las plataformas bacterianas tales como la ausencia de un mecanismo de transferencia de electrones, reductasas del citocromo P450 y la incompatibilidad de traducción de los módulos de señales de la membrana de enzimas P450, debido a la falta de un retículo endoplásmico.

En algunas realizaciones, la taxadieno-5 α -hidroxilasa asociada con métodos de la invención se optimiza a través de manipulación de la transmembrana N-terminal y/o la generación de enzimas quiméricas mediante la fusión traduccional con un participante redox CPR. En algunas realizaciones, el participante CPR redox es una citocromo P450 reductasa de *Taxus* (TCPR; Figura 5b). En determinadas realizaciones, la citocromo P450 taxadieno-5 α -hidroxilasa (T5 α OH) se obtiene de *Taxus cuspidate* (Número de Acceso a GenBank AY289209). En algunas realizaciones, NADPH:citocromo P450 reductasa (TCPR) se obtiene de *Taxus cuspidate* (Número de Acceso a GenBank AY571340).

Taxadieno 5 α -hidroxilasa y TCPR pueden estar unidos por un enlazador tal como GSTGS (SEQ ID NO: 50). En algunas realizaciones, taxadieno 5 α -hidroxilasa y/o TCPR están truncados para separar todo o parte de la región de transmembrana de una o ambas proteínas. Por ejemplo, taxadieno 5 α -hidroxilasa en algunas realizaciones está truncada para separar 8, 24 ó 42 aminoácidos N-terminales. En algunas realizaciones, los 74 aminoácidos N-terminales de TCPR están truncados. Un péptido adicional también se puede fusionar a taxadieno 5 α -hidroxilasa. Por ejemplo, uno o más aminoácidos de 17 α hidroxilasa bovina se pueden añadir a taxadieno 5 α -hidroxilasa. En determinadas realizaciones, el péptido MALLLAVF (SEQ ID NO: 51) se añade a taxadieno 5 α -hidroxilasa. Un ejemplo no limitante de polipéptido que comprende taxadieno 5 α -hidroxilasa fusionado a TCPR es At24T5 α OH-tTCPR.

En algunas realizaciones, la enzima quimérica es capaz de llevar a cabo la primera etapa de oxidación con una conversión de más de 10% de taxadieno en taxadieno-5 α -ol y el subproducto 5(12)-Oxa-3(11)-ciclotaxano. Por ejemplo, el porcentaje de conversión de taxadieno en taxadieno-5 α -ol y el subproducto 5(12)-Oxa-3(11)-ciclotaxano

puede ser al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 98%, aproximadamente 99% o aproximadamente 100%.

5 En determinadas realizaciones, la enzima quimérica es At245 α OH-tTCPR, que se encontró que es capaz de llevar a cabo la primera etapa de oxidación con una conversión de más del 98% de taxadieno en taxadieno-5 α -ol y el subproducto 5(12)-Oxa-3(11)-ciclotaxano (OCT; Figura 9a). La manipulación de la etapa de producción de taxadieno-5 α -ol es crítica en la producción de Taxol y se encontró que era limitante en los esfuerzos previos para la construcción de esta vía en levaduras. La construcción manipulada, desarrollada en esta memoria, demostró más de un 98% de conversión de taxadieno *in vivo* con una mejora de 2400 veces sobre la expresión heteróloga previa en levaduras. Por lo tanto, además de sintetizar cantidades significativamente mayores de compuestos intermedios clave de Taxol, este estudio también proporciona la base para la síntesis de metabolitos subsiguientes en la vía mediante una química similar de P450.

15 Tal como se utilizan en esta memoria, los términos "proteína" y "polipéptido" se utilizan indistintamente y, por lo tanto, el término polipéptido se puede utilizar para referirse a un polipéptido de longitud completa y se puede utilizar también para referirse a un fragmento de un polipéptido de longitud completa. Tal como se utiliza en esta memoria con respecto a polipéptidos, proteínas o fragmentos de los mismos, "aislado" significa separado de su entorno nativo y presente en cantidad suficiente para permitir su identificación o uso. Aislado, cuando se hace referencia a una proteína o polipéptido, significa, por ejemplo: (i) producido selectivamente mediante clonación de expresión o (ii) purificado tal como mediante cromatografía o electroforesis. Proteínas o polipéptidos aislados pueden ser, pero no tienen por qué ser, sustancialmente puros. La expresión "sustancialmente puros" significa que las proteínas o los polipéptidos están esencialmente libres de otras sustancias con las que se pueden encontrar en la producción, la naturaleza, o sistemas *in vivo* en una medida práctica y apropiada para el uso previsto. Polipéptidos sustancialmente puros pueden obtenerse de forma natural o pueden producirse mediante los métodos descritos en esta memoria, y se pueden purificar con técnicas bien conocidas en la técnica. Debido a que una proteína aislada se puede mezclar con otros componentes en una preparación, la proteína puede comprender sólo un pequeño porcentaje en peso de la preparación. No obstante, la proteína está aislada, debido a que ha sido separada de las sustancias con las que puede estar asociada en los sistemas vivos, es decir, aislada de otras proteínas.

25 En algunas realizaciones, uno o más de los genes asociados con la invención se expresa en un vector de expresión recombinante. Tal como se utiliza en esta memoria, un "vector" puede ser cualquiera de un cierto número de ácidos nucleicos en los que se pueden insertar una secuencia o secuencias deseadas mediante restricción y ligamiento para el transporte entre diferentes entornos genéticos o para la expresión en una célula huésped. Típicamente, los vectores están compuestos de ADN, aunque también están disponibles vectores de ARN. Los vectores incluyen, pero no se limitan a: plásmidos, fósidos, fagémidos, genomas de virus y cromosomas artificiales.

35 Un vector de clonación es uno que es capaz de replicarse de forma autónoma o integrada en el genoma en una célula huésped, y que se caracteriza, además, por uno o más sitios de endonucleasas de restricción en los que el vector puede ser cortado de una manera determinable y en los que una secuencia de ADN deseada puede ser ligada de tal manera que el nuevo vector recombinante mantiene su habilidad para replicarse en la célula huésped. En el caso de plásmidos, la replicación de la secuencia deseada puede ocurrir muchas veces, ya que el plásmido aumenta en el número de copias dentro de la célula huésped tal como una bacteria huésped o simplemente una sola vez por huésped antes de que se reproduzca por mitosis. En el caso de fagos, la replicación puede ocurrir de forma activa durante una fase lítica o de forma pasiva durante una fase lisogénica.

40 Un vector de expresión es uno en el que una secuencia de ADN deseada puede insertarse mediante restricción y ligamiento de tal manera que está unida operativamente a secuencias reguladoras y puede ser expresada como un transcrito de ARN. Los vectores pueden contener, además, una o más secuencias de marcador apropiadas para su uso en la identificación de células que han sido o que no han sido transformadas o transfectadas con el vector. Los marcadores incluyen, por ejemplo, genes que codifican proteínas que aumentan o disminuyen la resistencia o sensibilidad a antibióticos u otros compuestos, genes que codifican enzimas, cuyas actividades son detectables mediante ensayos estándares conocidos en la técnica (p. ej., β -galactosidasa, luciferasa o fosfatasa alcalina) y genes que afectan visiblemente al fenotipo de células, huéspedes, colonias o placas transformados o transfectados (p. ej., proteína fluorescente verde). Vectores preferidos son los que son capaces de replicación autónoma y de expresión de los productos génicos estructurales presentes en los segmentos de ADN a los que están unidos operativamente.

Tal como se utiliza en esta memoria, una secuencia codificante y secuencias reguladoras se dice que están unidas "operativamente" cuando están unidas de forma covalente de una manera tal que colocan la expresión o

transcripción de la secuencia codificadora bajo la influencia o el control de las secuencias reguladoras. Si se desea que las secuencias codificantes sean traducidas en una proteína funcional, se dice que dos secuencias de ADN están unidas operativamente si la inducción de un promotor en las secuencias reguladoras 5' resulta en la transcripción de la secuencia codificante y si la naturaleza del enlace entre las dos secuencias de ADN (1) no resulta en la introducción de una mutación de desplazamiento del marco, (2) interfiere con la capacidad de la región del promotor para dirigir la transcripción de las secuencias codificantes o (3) interfiere con la capacidad del transcrito de ARN correspondiente a ser traducido en una proteína. Por lo tanto, una región de promotor se uniría operativamente a una secuencia codificante si la región de promotor fuera capaz de efectuar la transcripción de esa secuencia de ADN de manera que el transcrito resultante se pueda traducir en la proteína o el polipéptido deseado.

5
10 Se puede utilizar una diversidad de secuencias de control de la transcripción (p. ej., secuencias de promotor/potenciador) para dirigir la expresión de la molécula de ácido nucleico que codifica cualquiera de las enzimas asociadas con la invención. El promotor puede ser un promotor nativo, es decir, el promotor del gen en su contexto endógeno, que proporciona una regulación normal de expresión del gen. En algunas realizaciones, el promotor puede ser constitutivo, es decir, el promotor es no regulado, permitiendo una transcripción continua de su gen asociado. Se puede utilizar también una diversidad de promotores condicionales, tales como promotores controlados por la presencia o ausencia de una molécula.

15
20 La naturaleza precisa de las secuencias reguladoras, necesarias para la expresión génica, puede variar entre especies o tipos de células, pero en general incluirá, según sea necesario, secuencias 5' no transcritas y 5' no traducidas implicadas en el inicio de la transcripción y la traducción, respectivamente, tal como una caja TATA, secuencia de casquete, secuencia CAAT, y similares. En particular, secuencias reguladoras 5' no transcritas de este tipo incluirán una región de promotor que incluye una secuencia de promotor para el control transcripcional del gen operativamente unido. Las secuencias reguladoras también pueden incluir secuencias de potenciador o secuencias de activador aguas arriba, según se desee. Los vectores pueden incluir opcionalmente secuencias 5' conductora o de señal. La elección y el diseño de un vector apropiado está dentro de la habilidad y criterio de un experto ordinario en la técnica.

25
30 Están disponibles comercialmente vectores de expresión que contienen todos los elementos necesarios para la expresión y son conocidos para los expertos en la técnica. Véase, p. ej., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Segunda Edición, Cold Spring Harbor Laboratory Press, 1989. Las células son manipuladas genéticamente mediante la introducción en las células de ADN heterólogo (ARN). Ese ADN heterólogo (ARN) es colocado bajo el control operativo de elementos transcripcionales para permitir la expresión del ADN heterólogo en la célula huésped.

35 Una molécula de ácido nucleico que codifica una enzima asociada con la invención se puede introducir en una célula o células de *E.coli* utilizando métodos y técnicas que son estándares en la técnica. Por ejemplo, moléculas de ácido nucleico pueden ser introducidas mediante protocolos estándares tales como transformación, incluyendo la transformación química y la electroporación, transducción, bombardeo de partículas, etc. La expresión de la molécula de ácido nucleico que codifica las enzimas asociadas con la invención también puede lograrse mediante la integración de la molécula de ácido nucleico en el genoma.

40 Células de *E. coli* de acuerdo con la invención pueden cultivarse en medios de cualquier tipo (rico o mínimo) y de cualquier composición. Como se entenderá por un experto normal en la técnica, la optimización de rutina permitiría el uso de una diversidad de tipos de medios. El medio seleccionado puede suplementarse con diversos componentes adicionales. Algunos ejemplos no limitativos de componentes suplementarios incluyen glucosa, antibióticos, IPTG para la inducción de genes, Suplemento de minerales traza ATCC, y glicolato. De manera similar, otros aspectos del medio, y las condiciones de crecimiento de las células pueden ser optimizados mediante experimentación rutinaria. Por ejemplo, el pH y la temperatura son ejemplos no limitativos de factores que pueden ser optimizados. En algunas realizaciones, factores tales como la elección de los medios, los suplementos de los medios y la temperatura pueden influir en los niveles de producción de terpenoides, tales como taxadieno. En algunas realizaciones, se pueden optimizar la concentración y la cantidad de un componente suplementario. En algunas realizaciones, se optimiza la frecuencia con la que los medios se suplementan con uno o más componentes suplementarios, y la cantidad de tiempo que los medios se cultivan antes de recoger un terpenoide, tal como taxadieno.

45
50 De acuerdo con aspectos de la invención, títulos elevados de un terpenoide tal como taxadieno se producen a través de la expresión recombinante de genes asociados con la invención, en una célula de *E. coli*. Tal como se utiliza en esta memoria "título elevado" se refiere a un título a escala en miligramos por litro (mg L⁻¹). El título producido para

un producto dado se verá influenciado por múltiples factores, incluyendo la elección de los medios. En algunas realizaciones, el título total de taxadieno es de al menos 1 mg L⁻¹. En algunas realizaciones, el título total de taxadieno es de al menos 10 mg L⁻¹. En algunas realizaciones, el título total de taxadieno es de al menos 250 mg L⁻¹. Por ejemplo, el título total de taxadieno puede ser al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900 o más de 900 mg L⁻¹, incluyendo todos los valores intermedios. En algunas realizaciones, el título total de taxadieno puede ser al menos 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2,0, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8, 2,9, 3,0, 3,1, 3,2, 3,3, 3,4, 3,5, 3,6, 3,7, 3,8, 3,9, 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8,4,9, 5,0, o más de 5,0 g L⁻¹ incluyendo cualesquiera valores intermedios.

En algunas realizaciones, el título total de taxadieno 5 α -ol es al menos 1 mg L⁻¹. En algunas realizaciones, el título total de taxadieno 5 α -ol es al menos 10 mg L⁻¹. En algunas realizaciones, el título total de taxadieno 5 α -ol es al menos 50 mg L⁻¹. Por ejemplo, el título total de taxadieno 5 α -ol puede ser de al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, o más de 70 mg L⁻¹, incluyendo cualesquiera valores intermedios.

Los cultivos líquidos utilizados para cultivar las células de *E. coli* asociadas con la invención pueden alojarse en cualquiera de los recipientes de cultivo conocidos y utilizados en la técnica. En algunas realizaciones de producción a gran escala en un recipiente de reacción ventilado tal como un reactor de tanque agitado puede utilizarse para producir grandes cantidades de terpenoides tales como taxadieno, que se pueden recuperar del cultivo celular. En algunas realizaciones, el terpenoide se recupera de la fase gaseosa del cultivo de células, por ejemplo mediante la adición de una capa orgánica, tal como dodecano, al cultivo celular y la recuperación del terpenoide de la capa orgánica.

Terpenoides tales como taxadieno, producidos a través de métodos descritos en esta memoria tienen amplias aplicaciones, incluyendo productos farmacéuticos tales como paclitaxel (Taxol), artemisinina, ginkolidas, eleuterobina y pseudopterosinas, y muchos otros compuestos farmacéuticos potenciales. Aplicaciones adicionales incluyen compuestos utilizados en sabores y cosméticos tales como geraniol, farnesol, geranligeraniol, linalool, limoneno, pineno, cineol e isopreno. Aplicaciones adicionales incluyen compuestos para uso como biocombustibles tales como alcoholes de 5, 10 y 15 átomos de carbono de longitud. Se señala que los compuestos anteriores se producen actualmente como extractos de diversas plantas. Métodos basados en extractos de plantas son tediosos, proporcionan cantidades muy pequeñas y están limitados en cuanto a las moléculas reales que se pueden obtener de esta manera, a saber, no permiten la fácil producción de derivados que pueden poseer propiedades muy superiores a las de los compuestos originales.

35 Ejemplos

MÉTODOS

Cepas, plásmidos, oligonucleótidos y genes

La cepa K12 MG1655 de *E. coli* se utilizó como cepa huésped de toda la construcción de la cepa taxadieno. Las cepas K12MG1655 Δ (recA, endA) de *E. coli* y K12MG1655 Δ cepas (recA, endA) ED3 de *E. coli* fueron proporcionados por el laboratorio del Profesor Kristala Prather en el MIT (Cambridge, MA). Un detalle de todos los plásmidos construidos para el estudio se muestra en la Tabla 2. Todos los oligonucleótidos utilizados en este estudio están contenidos en la Tabla 3.

Las secuencias de geranilgeranil pirofosfato sintasa (GGPPS)⁵⁰, Taxadieno sintasa (TS)⁵¹, Citocromo P450 Taxadieno 5 α -hidroxilasa (T5 α OH) y NADPH:citocromo P450 reductasa de *Taxus* (TCPR)⁴⁶ se obtuvieron de *Taxus canadensis*, *Taxus brevifolia*, *Taxus cuspidate* (Códigos de acceso a GenBank: AF081514, U48796, AY289209 y AY571340). Los genes fueron sintetizados por encargo utilizando los plásmidos y protocolos reseñados por Kodumal et al.⁵² (Detalles complementarios, Apéndice 1) para incorporar el codón de traducción de *E. coli* y separar sitios de restricción para fines de clonación. Los nucleótidos que corresponden a los 98 y 60 aminoácidos N-terminales de GGPPS y TS (péptido de tránsito de plastidio) se separaron y se insertó la secuencia de inserción de la traducción Met.¹⁷

Construcción de la vía MEP (operón dxs-idi-ispDF).

El operón dxs-idi-ispDF se construyó inicialmente clonando cada uno de los genes del genoma de *E. coli* K12 MG1655 utilizando los dxs(s) cebadores, dxs(a), idi(s), idi(a), ispDF(s) e ispDFI(a) bajo pET21C + plásmido con el promotor T7 (p20T7MEP)⁵³. Utilizando los cebadores dxsidiispDFNcoI (s) y dxsidiispDFKpnI (a), se sub-clonó el operón dxs-idi-ispDF y el plásmido pTrcHis2B (Invitrogen), después de digerido con NcoI y KpnI para el plásmido pTrcMEP (p20TrcMEP). El plásmido p20TrcMEP digerido con MluI y PmeI y clonado en MluI y el plásmido pACYC184-melA(P2A) digerido con PmeI para construir el plásmido p10TrcMEP. El plásmido pTrcMEP digerido con BstZ17I y ScaI y clonado en el plásmido pCL1920 digerido con PvuII para construir el plásmido p5TrcMEP. Para la construcción de plásmido p20T5MEP, inicialmente el operón dxs-idi-ispDF fue clonado en el plásmido pQE con el promotor T5 (pQE-MEP) utilizando los cebadores dxsidiispDFNcoI (s) y dxsidiispDFXhoI (a). Una fracción del ADN del operón con el promotor T5 se amplificó utilizando los cebadores T5AgeI(s) y T5NheI(a) a partir del plásmido pQEMEP. El fragmento de ADN fue digerido con AgeI/NheI y clonado en el plásmido p20T7MEP digerido con enzimas SgrAI/NheI.

Construcción de la vía Taxadieno (operones GT y TG).

Las vías taxadieno aguas abajo (operón GT y TG) se construyeron mediante clonación de fragmentos de PCR de GGPS y TS en los sitios *NcoI* - *EcoRI* y *EcoRI* - *SalI* del plásmido pTrcHIS2B para crear p20TrcGT y p20TrcTG utilizando los cebadores GGPPSNcoI(s), GGPPSEcoRI(a), TSEcoRI(s), TSSalI(a), TSNcoI(s) TSEcoRI(a) GGPPSEcoRI(s) y GGPPSSalI (a). Para la construcción de p20T5GT, inicialmente el operón se amplificó con los cebadores GGPPSNcoI(s) y TSXhoI(a) y se clonó en un plásmido pQE bajo el promotor T5 digerido con NcoI/XhoI. Además, la secuencia se digirió con XbaI y XhoI y se clonó en la cadena principal del plásmido pTrc amplificado utilizando los cebadores pTrcSal(s) y pTrcXba(a). p10T7TG se construyó mediante subclonación del operón TG digerido con NcoI/SalI de p20TrcTG en el plásmido pACYC-Duet1 digerido con NcoI/SalI. p5T7TG se construyó clonando el fragmento digerido con BspEI/XbaI al ADN digerido con XbaI/BspEI amplificado a partir del plásmido pCL1920 utilizando los cebadores pCLBspEI(s) y pCLXbaI(a).

Construcción de plásmidos de la vía MEP de integración cromosómica

Para la construcción de los plásmidos con la casete FRP-Km-FRP para la amplificación de la secuencia de integración, p20T7MEP y p20T5MEP fueron digeridos con XhoI/ScaI. La casete FRP-Km-FRP se amplificó a partir de la casete Km con la secuencia FRP del plásmido pKD13 utilizando los cebadores KmFRPXhoI(s) y KmFRPScaI(a). El ADN amplificado fue digerido con XhoI/ScaI y clonado en el plásmido p20T7MEP y p20T5MEP digerido con XhoI/ScaI (p20T7MEPKmFRP y p20T5MEPKmFRP). De manera similar, el plásmido p20TrcMEP se digirió con SacI/ScaI y el ADN amplificado utilizando los cebadores KmFRPSacI(s) y KmFRPScaI(a) se digirió, clonado en el plásmido p20TrcMEP (p20TrcMEPKm-FRP).

Integración cromosómica de la casete de la vía de MEP (LacIq-MEP-FRP-Km-FRP)

Las vías MEP construidos bajo los promotores T7, T5 y Trc se localizaron en la región del operón ara en el cromosoma con el marcador Kan. Los fragmentos de PCR fueron amplificados a partir de p20T7MEPKmFRP, p20T5MEPKmFRP y p20TrcMEPKm-FRP utilizando los cebadores IntT7T5(s), IntTrc(s) e Int(a) y luego se electroporaron en células MG1655 recA-end- de *E. coli* y células MG1655 recA-end-EDE3 de *E. coli* para la integración cromosómica a través de la técnica de recombinación λ Red⁵⁴. El sitio de localización específica se confirmó y el marcador Km se separó mediante la acción de la FLP recombinasa después de la integración con éxito de genes.

Construcción de la vía Taxadieno 5 α -ol

La región de transmembrana (TM) de la taxadieno 5 α -ol hidroxilasa (T5 α OH) y Citocromo P450 reductasa de *Taxus* (TCPR) se identificó utilizando el software PredictProtein (www.predictprotein.org)⁵⁵. Para la manipulación de la transmembrana se realizó un truncamiento selectivo en 8, 24 y 42 residuos aminoácidos en la región de la transmembrana N-terminal de taxadieno hidroxilasa 5 α -ol (T5 α OH) y la región de 74 aminoácidos en la TCPR. La separación de los 8, 24 y 42 residuos aminoácidos N-terminales de taxadieno hidroxilasa 5 α -ol (T5 α OH), la incorporación del péptido MALLLAVF (SEQ ID NO: 51) de 8 residuos N-terminales de 17 α hidroxilasa bovina sustituido con un aminoácido a las secuencias de T5 α OH⁴⁴ N-terminal truncadas y el enlazador peptídico GSTGS se llevó a cabo utilizando los cebadores CYP17At8AANdel(s), CYP17At24AANdel(s), CYP17At42AANdel(s) y

CYPLinkBamHI(a). Utilizando estos cebadores se amplificó cada uno de los ADN modificados, se digirió NdeI/BamHI y se clonó en el plásmido pACYC DUET1 digerido con NdeI/BamHI para construir los plásmidos p10At8T5αOH, p10At24T5αOH y p10At42T5αOH. La secuencia TCPR truncada (tTCPR) de 74 aminoácidos se amplificó utilizando los cebadores CPRBamHI(s) y CPRSall(a). La secuencia tTCPR amplificada y los plásmidos p10At8T5αOH, p10At24T5αOH y p10At42T5αOH, se digirieron con BamHI/Sall y se clonaron para construir los plásmidos p10At8T5αOH-tTCPR, p10At24T5αOH-tTCPR y p10At42T5αOH-tTCPR.

Crecimiento del cultivo para el rastreo de taxadieno y análisis de taxadieno-5α-ol

Transformantes individuales de cepas de *E. coli* pre-manipuladas, que albergan el plásmido apropiado con la vía (MEP) aguas arriba, la vía taxadieno aguas abajo y taxadieno 5α ol-se cultivaron durante 18 h a 30°C en medio de Luria-Bertani (LB) (suplementado con antibióticos apropiados, 100 mg/mL de carbenicilina, 34 mg/mL de cloranfenicol, 25 mg/L de kanamicina o 50 mg/L de espectinomicina). Para los cultivos a pequeña escala para rastrear las cepas manipuladas, éste pre-inóculo se utilizó para sembrar medios ricos de 2 mL recientes (5 g/L de extracto de levaduras, 10 g/L de triptona, 15 g/L de glucosa, 10 g/L de NaCl, HEPS 100 mM, 3 mL/L de Antiespumante B, pH 7,6, 100 ug/mL de carbenicilina y 34 ug/mL de cloranfenicol), a una A₆₀₀ de partida de 0,1. El cultivo se mantuvo con los antibióticos apropiados e IPTG 100 mM para la inducción génica a 22°C durante 5 días.

Experimentos de biorreactores para la cepa productora de taxadieno 5a-ol.

El biorreactor Bioflo de 3-L (New Brunswick) fue montado según las instrucciones del fabricante. Un litro de medio rico con glicerol al 1% (v/v) se inoculó con 50 mL de cultivo de 8 h (A₆₀₀ de ~2,2) de la cepa 26-At24T5αOH-tTCPR cultivada en medio LB que contiene los antibióticos (100 mg/mL de carbenicilina, 34 mg/mL de cloranfenicol) a las mismas concentraciones. Biorreactores de 1 L con fermentación bifásica líquido-líquido utilizando dodecano al 20% v/v. El oxígeno se suministró como aire filtrado a 0,5 v/v/m y la agitación se ajustó para mantener los niveles de oxígeno disuelto por encima del 50%. El pH del cultivo se controló a 7,0 utilizando NaOH al 10%. La temperatura del cultivo en el fermentador se controló a 30°C hasta que las células fueron cultivadas en una densidad óptica de aproximadamente 0,8, según se mide a una longitud de onda de 600 nm (OD600). La temperatura del fermentador se redujo a 22°C y las células se indujeron con IPTG 0,1 mM. Dodecano se añadió en condiciones asépticas al 20% (v/v) del volumen del medio. Durante el curso de la fermentación se vigiló la concentración de glicerol y la acumulación de acetato a intervalos de tiempo constantes. Durante la fermentación, a medida que la concentración de glicerol se agotaba por debajo de 0,5 g/L, se introdujo glicerol (3 g/L) en el biorreactor.

La fermentación se optimizó adicionalmente utilizando un cultivo alimentado discontinuo con un medio de alimentación definido que contiene extracto de levadura al 0,5% y dodecano al 20% (v/v) (13,3 g/L de KH₂PO₄, 4 g/L de (NH₄)₂HPO₄, 1,7 g/L de ácido cítrico, 0,0084 g/L de EDTA, 0,0025 g/L de CoCl₂, 0,015 g/L de MnCl₂, 0,0015 g/L de CuCl₂, 0,003 g/L de H₃BO₃, 0,0025 g/L de Na₂MoO₄, 0,008 g/L de Zn(CH₃COO)₂, 0,06 g/L de citrato de Fe(III), 0,0045 g/L de tiamina, 1,3 g/L de MgSO₄, 10 g/L de glicerol, 5 g/L de extracto de levadura, pH 7,0). La misma composición del medio se utilizó para la fermentación de las cepas **17** y **26** con los antibióticos apropiados (cepa 17: 100 µg/mL de carbenicilina y 50 µg/mL de espectinomicina; cepa 26: 50 µg/mL de espectinomicina).

Para la cepa productora de taxadieno-5α-ol, un litro de medio complejo con glicerol al 1% (v/v) se inoculó con 50 mL de un cultivo de 8 h (OD de ~ 2,2) de la cepa **26**-At24T5αOH-tTCPR cultivada en medio LB que contiene 50 µg/mL de espectinomicina y 34 µg/mL de cloranfenicol). El oxígeno fue suministrado como aire filtrado a 0,5 (vvm) y la agitación se ajustó para mantener los niveles de oxígeno disuelto por encima del 30%. El pH del cultivo se controló a 7,0 utilizando NaOH al 10%. La temperatura del cultivo en el fermentador se controló a 30°C hasta que las células fueron cultivadas en una densidad óptica de aproximadamente 0,8, según se mide a una longitud de onda de 600 nm (OD600). La temperatura del fermentador se redujo a 22°C y la vía se indujo con IPTG 0,1 mM. Se añadió dodecano en condiciones asépticas al 20% (v/v) del volumen del medio. Durante el curso de la fermentación, la concentración de glicerol y la acumulación de acetato se controlaron a intervalos de tiempo constantes. Durante la fermentación a medida que la concentración de glicerol se agotaba por debajo de 0,5-1 g/L, se introdujeron 3 g/L de glicerol en el biorreactor.

Análisis GC-MS de taxadieno y taxadieno-5α-ol

Para el análisis de la acumulación de taxadieno de cultivo a pequeña escala, 1,5 mL del cultivo se sometieron a vórtice con 1 mL de hexano durante 30 min. La mezcla se centrifuga para separar la capa orgánica. Para el biorreactor 1 uL de la capa de dodecano se diluyó a 200 uL utilizando hexano. 1 uL de la capa de hexano se analizó mediante GC-MS (Varian saturn 3800 GC unido a un Varian 2000 MS). La muestra se inyectó en una columna

HP5ms (30 m x 250 μ m x 0,25 μ m de espesor) (Agilent Technologies USA). Helio (ultra pureza) a un caudal de 1,0 ml/min se utilizó como gas portador. La temperatura de la estufa se mantuvo primero constante a 50°C durante 1 min, y luego se aumentó a 220°C en un incremento de 10°C/min, y finalmente se mantuvo a esta temperatura durante 10 min. Las temperaturas del inyector y de la tubería de transferencia se fijaron en 200°C y 250°C, respectivamente.

No estaban comercialmente disponibles compuestos estándares de fuentes biológicas o sintéticas para taxadieno y taxadieno 5 α -ol. Por lo tanto, los autores de la invención realizaron fermentaciones de *E. coli* productora de taxadieno en un biorreactor de 2L para extraer material puro. Taxadieno se extrajo mediante extracción con disolvente utilizando hexano, seguido de múltiples rondas de cromatografía en columna de sílice para obtener el material puro para construir una curva estándar para el análisis de GC-MS. Los autores de la invención compararon los perfiles GC y MS del taxadieno puro con la bibliografía reseñada para confirmar la autenticidad del compuesto⁶⁰. Con el fin de comprobar la pureza, realizaron 1H-RMN de taxadieno. Dado que la acumulación de taxadieno-5 α -ol era de un nivel muy bajo, utilizaron axadieno como medida para cuantificar la producción de esta molécula y las características de la fragmentación del espectro de masas auténticas de los informes anteriores⁴².

15 Mediciones por qPCR para el análisis transcripcional de cepas manipuladas

Los niveles de expresión génica transcripcional de cada uno de los genes se detectaron mediante qPCR en ARNm aislado de las cepas apropiadas. Para evitar la degradación, el ARN se estabilizó antes de la lisis celular utilizando reactivo bacteriano RNeasy Protect (Qiagen). Subsecuentemente, el ARN total fue aislado utilizando el mini kit RNeasy (Qiagen), combinado con la separación de contaminantes de ADN genómico basada en nucleasa. ADNc se amplificó utilizando el kit de síntesis de ADNc iScript (Biorad). Se llevó a cabo una qPCR en un Bio-Rad iCycler utilizando el iQ SYBR Green Supermix (Biorad). El nivel de expresión del gen *rrsA*, que no está sujeto a una expresión variable, se utilizó para la normalización de los valores de qPCR⁵⁶. La Tabla 3 tiene cebadores utilizados para la qPCR. Para cada par de cebadores, se construyó una curva estándar con ARNm de *E. coli* como molde.

25 Ejemplo 1: La acumulación de taxadieno exhibe una fuerte dependencia no lineal de las intensidades relativas de las vías MEP aguas arriba y taxadieno sintéticas aguas abajo.

La Fig. 1b representa las diversas maneras en que se combinaron los promotores y los números de copias de genes para modular el flujo relativo (o la resistencia) a través de las vías de aguas arriba y aguas abajo de la síntesis de taxadieno. Se construyó un total de 16 cepas con el fin de estrechar la vía MEP, así como equilibrar de manera óptima la vía de taxadieno aguas abajo. Las Figs. 2a, b resumen los resultados de la acumulación de taxadieno en cada una de estas cepas, acentuando la Fig. 2a la dependencia de la acumulación de taxadieno en la vía de aguas arriba para valores constantes de la vía de aguas abajo, y la Fig. 2b la dependencia en la vía de aguas abajo para una intensidad constante de la vía de aguas arriba (véase también la Tabla 1 para el cálculo de la expresión de la vía de aguas arriba y de aguas abajo a partir de las fuerzas del promotor reseñado y de los números de copias del plásmido³³⁻³⁶). Claramente, se exhiben máximos con respecto tanto a la expresión de la vía de aguas arriba y de aguas abajo. Para la expresión de la vía de aguas abajo constante (Fig. 2a), dado que la expresión de la vía de aguas arriba aumenta desde niveles muy bajos, la producción de taxadieno aumenta inicialmente debido al suministro incrementado de precursores a la vía global. Sin embargo, después de un valor intermedio, incrementos adicionales de la vía de aguas arriba no pueden ser acomodados por la capacidad de la vía de aguas abajo. Este desequilibrio de las vías conduce a la acumulación de un compuesto intermedio (véase más adelante) que puede ser o bien inhibitoria de las células o bien puede indicar simplemente la desviación de flujo de una vía competitiva, resultando en última instancia en la reducción de la acumulación de taxadieno.

Para la expresión de la vía aguas arriba constante (Fig. 2b), se observa de manera similar un máximo con respecto al nivel de expresión de la vía de aguas abajo. Esto se atribuye a una limitación inicial de la producción de taxadieno por los bajos niveles de expresión de la vía de aguas abajo que, por lo tanto, limita la velocidad con respecto a la producción de taxadieno. A niveles de expresión altos de la vía de aguas abajo probablemente los autores de la invención están viendo el efecto negativo de un alto número de copias en la fisiología de la célula, por lo tanto, existe un máximo con respecto a la expresión de la vía de aguas abajo. Estos resultados demuestran que se pueden obtener cambios drásticos en la acumulación de taxadieno a partir de los cambios dentro de una estrecha ventana de niveles de expresión de las vías de aguas arriba y de aguas abajo. Por ejemplo, una cepa que contiene una copia adicional de la vía de aguas arriba en su cromosoma bajo el control del promotor Trc (Cepa 8, Fig. 2a) producía 2000 veces más taxadieno que una que sobre-expresa sólo la vía de aguas abajo sintética (Cepa 1, Fig. 2a). Además, cambiando el orden de los genes en el operón sintético aguas abajo de GT (GPPS-TS) a TG (TS-GPPS) resultó en un aumento de 2-3 veces (cepas 1-4 en comparación con 5, 8, 11 y 14). Los resultados observados

demuestran que la clave para la sobreproducción de taxadieno es una amplia capacidad de la vía de aguas abajo y el equilibrio cuidadoso entre la vía del precursor de aguas arriba con la vía de taxadieno sintético aguas abajo. En conjunto, las cepas manipuladas establecieron que el flujo de la vía MEP puede ser sustancial, si se rastrea simultáneamente una amplia gama de niveles de expresión de la vía de aguas arriba endógena y de aguas abajo sintética.

Ejemplo 2: La integración cromosómica y la sintonización de las vías de aguas arriba y de aguas abajo potencia adicionalmente la producción de taxadieno.

Para proporcionar una amplia intensidad de la vía de aguas abajo al tiempo que se minimiza la carga metabólica portada por el plásmido³⁷, se manipularon dos nuevos conjuntos de 4 cepas cada uno (cepas 25-28 y 29-32), en que la vía de aguas abajo se colocó bajo el control de un promotor fuerte (T7) mientras que se mantiene un número relativamente bajo de 5 y 10 copias, respectivamente. Se puede ver (Fig. 2c) que, mientras que el máximo de taxadieno se mantiene a una alta intensidad de aguas abajo (cepas 21-24), se obtiene una respuesta monotónica en la baja intensidad de la vía de aguas abajo (cepas 17-20, Fig. 2c). Esta observación condujo a la construcción de dos conjuntos adicionales de 4 cepas cada uno, que mantenía el mismo nivel de intensidad de la vía de aguas abajo como antes, pero expresaban niveles muy bajos de la vía de aguas arriba (cepas 25-28 y 29-32, Fig. 2d). Adicionalmente, el operón de la vía de aguas arriba del último conjunto de cepas fue integrado cromosómicamente. Se puede observar que no sólo se recupera el máximo de taxadieno, aunque a niveles muy bajos de la vía de aguas arriba, sino que se alcanza un máximo de taxadieno mucho mayor (300 mg/L). Los autores de la invención creen que este aumento significativo se puede atribuir a una disminución de la carga metabólica de la célula. Esto se logró 1) eliminando la dependencia del plásmido a través de la integración de la vía en el cromosoma y 2) logrando un buen equilibrio entre la expresión de la vía de aguas arriba y de aguas abajo.

Las 32 construcciones recombinantes les permitió sondear adecuadamente el espacio de expresión de la vía modular y amplificar ~15000 veces la mejora en la producción de taxadieno. Esto es, con mucho, la más alta producción de terpenoides a partir de la vía isoprenoide MEP de *E. coli* reseñada (Fig. 3a). Adicionalmente, las mejoras múltiples observadas en la producción de terpenoides son significativamente más altas que las de los enfoques de manipulación metabólica combinatoria reseñada que buscó un extenso espacio genético que comprende hasta un billón de variantes combinatorias de la vía isoprenoide³⁰. Esto sugiere que la optimización de la vía depende mucho más del equilibrado preciso de la expresión de los módulos de la vía que la optimización gen combinatoria multi-fuente. Los máximos de múltiples exhibidos en el panorama fenotípico de la Fig.1 subraya la importancia de sondear el espacio de expresión a una resolución suficiente para identificar la región del comportamiento óptimo global de la vía. La Fig. 7 representa las mejoras múltiples en la producción de taxadieno a partir de la búsqueda de la expresión de la vía modular.

Ejemplo 3: El metabolito se correlaciona inversamente con la producción de taxadieno y la identificación de metabolitos.

El análisis metabolómico de las cepas manipuladas previas identificó un subproducto metabolito, hasta ahora, desconocido, que se correlacionaba fuertemente con los niveles de expresión de la vía y la producción de taxadieno (Fig. 3 y Fig. 8). Aunque la identidad química del metabolito era desconocida, la hipótesis de los autores de la invención es que un producto secundario isoprenoide, que resulta de la desviación de la vía y ha sido anti-correlacionado como una variable directa a la producción de taxadieno (Fig. 3 y Fig. 8) a partir de las cepas manipuladas. Un atributo crítico de las cepas óptimas de los autores de la invención es el reparto preciso que alivia la acumulación de este metabolito, resultando en una mayor producción de taxadieno. Este equilibrio puede ser modulado en diferentes niveles, a partir del cromosoma, o plásmidos de diferentes números de copias, utilizando diferentes promotores, con una acumulación de taxadieno significativamente diferente.

Posteriormente, el pico correspondiente en el cromatograma de cromatografía de gases-espectrometría de masas (GC-MS) fue identificado como indol por GC-MS, estudios de espectroscopia por resonancia magnética nuclear (RMN) ¹H y ¹³C (Fig. 16). Los autores de la invención encontraron que la síntesis de taxadieno por la cepa 26 se ve gravemente inhibida por indol exógeno a niveles de indol superiores a ~100 mg/L (Fig. 15b). El aumento adicional de la concentración de indol también inhibía el crecimiento celular, siendo el nivel de inhibición muy dependiente de la cepa (Fig. 15c). Aunque actualmente no está claro el mecanismo bioquímico de la interacción de indol con la vía isoprenoide, los resultados en la Fig. 15 sugieren un posible efecto sinérgico entre los compuestos de indol y terpenoides de la vía isoprenoide en la inhibición del crecimiento celular. Sin conocer el mecanismo específico, parece que la cepa 26 ha mitigado el efecto del indol, lo cual fue llevado adelante para un estudio adicional.

Ejemplo 4: Cultivo de cepas manipuladas.

Con el fin de explorar el potencial de producción de taxadieno bajo condiciones controladas para las cepas manipuladas, se llevaron a cabo cultivos alimentados discontinuos de las tres cepas acumuladoras de taxadieno más altas (~ 60 mg/L de la cepa 22; ~125 mg/L de la cepa 17; ~ 300 mg/L de la cepa 26) en biorreactores de 1 L (Fig. 17). Los estudios de alimentados discontinuos se llevaron a cabo como una fermentación de dos fases líquido-líquido utilizando una capa superior de dodecano al 20% (v/v). El disolvente orgánico se introdujo para evitar el arrastre por el aire de taxadieno secretado del medio de fermentación, tal como se indica por los hallazgos preliminares. En medios definidos con alimentación de glicerol controlada, la productividad de taxadieno aumentó a 174 ± 5 mg/L (SD), 210 ± 7 mg/L (SD) y 1020 ± 80 mg/L (SD), respectivamente para las cepas 22, 17 y 26 (Fig. 17a). Adicionalmente, la producción de taxadieno afectaba significativamente al fenotipo de crecimiento, la acumulación de acetato y el consumo de glicerol (Figs. 17b-17d).

La Fig. 17c muestra que el acetato se acumula inicialmente en todas las cepas, pero después de ~60 h el acetato disminuye en las cepas 17 y 26, mientras que continúa aumentando en la cepa 22. Este fenómeno pone de relieve las diferencias en el metabolismo central de carbono entre las cepas de alto flujo MEP (26 y 17) y la cepa de bajo flujo MEP (22). Adicionalmente, esta observación es otra ilustración de la buena fisiología que caracteriza a una cepa de funcionamiento bien equilibrado. El ácido acético, como producto del metabolismo de sobreflujo, es producido inicialmente por todas las cepas debido a las altas concentraciones de glicerol iniciales utilizados en estas fermentaciones y el correspondiente alto flujo de la vía de glicerol. Este flujo es suficiente para suministrar a la célula también la vía MEP, así como las otras vías metabólicas.

A las ~ 48 h, el glicerol inicial se agota, y el cultivo cambia a un modo alimentado discontinuo, durante el cual se mantienen niveles bajos pero constantes de glicerol. Esto resulta en un bajo flujo global de glicerol, el cual, para las cepas con alto flujo de MEP (cepas 26 y 17), se desvía en su mayoría a la vía MEP al tiempo que minimiza el metabolismo de sobreflujo. Como resultado, se reduce o incluso se elimina totalmente la producción de ácido acético. En cuanto a la disminución de la concentración de ácido acético, es posible que pueda haberse producido, en cierta medida, la asimilación de ácido acético, aunque esto no se investigó ulteriormente desde un punto de vista del análisis del flujo. Disminuye algo de la evaporación y dilución debido a que la alimentación de glicerol está contribuyendo adicionalmente a la disminución de la concentración de ácido acético observada. Por el contrario, para las cepas con bajo flujo MEP (cepa 22), la desviación de flujo a la vía MEP no es muy significativo, de modo que el flujo de glicerol todavía suministra todos los requisitos de carbono y energía necesarios. Continúa produciéndose un metabolismo de sobreflujo, lo cual conduce a la secreción de acetato.

Es evidente que la alta productividad y el crecimiento más robusto de la cepa 26 permitieron una acumulación muy alta de taxadieno. Deberían ser posibles mejoras adicionales a través de la optimización de las condiciones en el biorreactor, el equilibrio de los nutrientes en el medio de crecimiento y la optimización del suministro de carbono.

Ejemplo 5: Los niveles de expresión de la vía de aguas arriba y de aguas abajo y el crecimiento celular revelan la complejidad subyacente.

Para una comprensión más detallada del equilibrio manipulado en la expresión de las vías, los autores de la invención cuantificaron los niveles de expresión génica de la transcripción de *dxs* (vía de aguas arriba) y *TS* (vía de aguas abajo) para las cepas más altas productoras de taxadieno y cepas vecinas de las Figs. 2c y d (cepas 17, 22 y 25-32) (Fig. 4a, b). Como hipotetizaron, la expresión de la vía de aguas arriba aumentaba monotónicamente con la fuerza del promotor y el número de copias para el vector de MEP de: promotor nativo, Trc, T5, T7, y plásmidos de 10 copias y 20 copias, como se ve en la expresión de *DXS* (Fig. 4a). Así, encontraron que el nivel de expresión de *dxs* se correlaciona bien con la intensidad de la vía aguas arriba. Encontraron correlaciones similares para los otros genes de la vía de aguas arriba, *idi*, *ispD* e *ispF* (Fig. 14a, b). En la expresión de genes aguas abajo, se cuantificó una mejora de ~ 2 veces después de transferir la vía del plásmido de 5 a 10 copias (serie 25-28 y serie 29-32) (Fig. 4b).

Mientras que los efectos del promotor y del número de copias influyeron en las expresiones de genes, también eran destacados efectos secundarios sobre la expresión de la otra vía. La Fig. 4a muestra que para las mismas casetes de expresión de *dxs*, al aumentar el número de copias del plásmido *TS* de 5 a 10, aumentó la expresión de *dxs*. Curiosamente, el plásmido *TS* de 5 copias (serie de cepas 25-28) contenía rendimientos de taxadieno sustancialmente más altos (Fig. 2d) y un menor crecimiento (Figs. 4c, d) que el plásmido *TS* de 10 copias. Plásmidos de control que no contenían la vía heteróloga de taxadieno crecieron densidades dos veces más altas, implicando que la inhibición del crecimiento en las cepas de la serie 25-28 está directamente relacionada con la vía

metabólica de taxadieno y la acumulación de taxadieno y sus compuestos intermedios directos (Fig. 4c). Sin embargo, la serie de cepas 29-32 sólo mostró un modesto incremento en el rendimiento del crecimiento cuando se comparan los plásmidos de control vacíos con las cepas que expresan taxadieno (Fig. 4d). Esta interacción entre el crecimiento, la producción de taxadieno y el nivel de expresión también se puede ver con los vectores de expresión de aguas arriba basados en plásmidos (cepas 17 y 22). La inhibición del crecimiento era mucho mayor en la cepa de 10 copias, de alta producción de taxadieno (cepa 17) en comparación con la cepa de 20 copias, de menor producción de taxadieno (cepa 22) (Fig. 4d). Por lo tanto la toxicidad del producto y la desviación de carbono a la vía heteróloga es probable que impidan el crecimiento, en lugar del mantenimiento del plásmido.

También inesperado era el profundo efecto del vector de expresión aguas arriba sobre la expresión aguas abajo. La Fig. 4b tendría dos líneas rectas, si no hubiera interferencias entre las vías. Sin embargo, se observan ~ 3 veces cambios en la expresión de TS para los diferentes vectores de expresión MEP. Esto se debe, probablemente, a una competencia significativa por los recursos (materias primas y energía) que se retiran del metabolismo del huésped para la sobreexpresión tanto de los cuatro genes de aguas arriba como de los dos genes de aguas abajo³⁸. En comparación con la cepa de control 25c, se observó una inhibición del crecimiento de 4 veces con la cepa 25, indicando que la alta sobreexpresión de la vía de taxadieno sintética inducía toxicidad, alterando el fenotipo de crecimiento en comparación con la sobre-expresión de la vía nativo (Fig. 4c). Sin embargo, aumentó la expresión aguas arriba, se redujo la expresión aguas abajo, de manera inadvertida en el caso de los autores de la invención, a niveles deseables para equilibrar las vías de aguas arriba y aguas abajo, reduciendo al mínimo la inhibición del crecimiento (cepa 26).

En el extremo de la sobre-expresión de proteínas, la vía MEP impulsada por el promotor T7 resultó en una inhibición severa del crecimiento, debido a la síntesis de cuatro proteínas de alto nivel (cepas 28 y 32). La expresión de los genes TS por parte de T7 no parece tener un efecto tan drástico por sí mismo. Las altas tasas de síntesis de proteínas a partir de la expresión inducida por T7 (Fig. 4ab) podría conducir a la regulación a la baja de la maquinaria de síntesis de proteínas, incluyendo los componentes de genes de referencia de la fase temprana de crecimiento, perjudica el crecimiento celular y reduce el aumento de la biomasa^{39,40}. Los autores de la invención hipotetizaron que los fenotipos de crecimiento complejos observados son efectos acumulativos de (1) toxicidad inducida por la activación del metabolismo de isoprenoide/taxadieno, y (2) los efectos de la alta expresión de proteína recombinante. En conjunto el enfoque de manipulación de la vía multivariante-modular de los autores de la invención generó una diversidad inesperada en el metabolismo de terpenoide y su correlación con la expresión de la vía y la fisiología celular. Un diseño racional de microbios para la producción de metabolitos secundarios requerirá una comprensión de la expresión de la vía que va más allá de una comprensión lineal/independiente de las fuerzas del promotor y de los números de copias. Sin embargo, enfoques multivariados simples, tal como se emplean aquí, pueden introducir la diversidad necesaria tanto para (1) encontrar altos productores como para (2) proporcionar un panorama para la investigación sistemática de efectos de mayor orden que son dominantes, pero poco apreciados, en la manipulación de la vía metabólica.

Ejemplo 6: Ingeniería de la química de oxidación basada en Taxol P450 en *E. coli*.

Una característica central en la biosíntesis de Taxol es la oxigenación en múltiples posiciones de la estructura del núcleo de taxano, reacciones que se consideran que son mediadas por monooxigenasas dependientes del citocromo P450⁴¹. Después de completarse la etapa de ciclación comprometida de la vía, la olefina parental, taxa-4(5),11(12)-dieno, se hidroxila seguidamente en la posición C5 por una enzima citocromo P450, lo que representa la primera de las ocho etapas de oxigenación (del núcleo de taxano) en la ruta a Taxol (Fig. 6)⁴². Por lo tanto, una etapa clave hacia la manipulación de microbios productores de Taxol es el desarrollo de la química de oxidación basada en P450 *in vivo*. La primera etapa de oxigenación es catalizada por un citocromo P450, taxadieno 5 α -hidroxilasa, una monooxigenasa inusual que cataliza la reacción de hidroxilación junto con la migración del doble enlace en el taxadieno precursor de diterpeno (Fig. 5a). Los autores de la invención informan sobre la primera ampliación con éxito de la vía sintética de taxadieno a taxadien-5 α -ol y presentan los primeros ejemplos de la producción *in vivo* de cualquier producto intermedio de Taxol funcionalizado en *E. coli*.

En general, la expresión funcional del citocromo P450 de vegetales es un reto⁴³, debido a las limitaciones inherentes de las plataformas bacterianas tales como la ausencia de la maquinaria de transferencia de electrones, reductasas del citocromo P450 y la incompatibilidad de traducción de los módulos de señales de la membrana de enzimas P450, debido a la falta de un retículo endoplásmico. Recientemente, a través de la manipulación de transmembrana (TM) y de la generación de enzimas quimeras de P450 y reductasas de CPR, algunos P450s de vegetales han sido expresados en *E. coli* para la biosíntesis de moléculas funcionales^{22,44}. Sin embargo, cada citocromo p450 de vegetales es único en su secuencia señal de transmembrana y características de transferencia de electrones de su homólogo reductasa⁴⁵. Los estudios iniciales de los autores de la invención se centraron en la optimización de la

expresión de taxadieno 5 α -hidroxilasa sintética optimizada en codones por manipulación N-terminal de la transmembrana y la generación de enzimas quimeras a través de fusión traduccional con el participante CPR redox de las especies de *Taxus*, citocromo P450 reductasa de *Taxus* (TCPR) (Fig. 5b)^{42,44,46}. Una de las enzimas quimeras generadas, At24T5 α OH-tTCPR, era altamente eficaz en llevar a cabo la primera etapa de oxidación con más de 98% de conversión de taxadieno en taxadien-5 α -ol y el subproducto 5(12)-Oxa-3(11)-ciclotaxano (OCT) (Fig. 9a).

Comparado con los otros P450s quiméricos, At24T5 α OH-tTCPR proporcionó una producción dos veces mayor (21 mg/L) de taxadien-5 α -ol. Asimismo, la actividad más débil de At8T5 α OH-tTCPR y At24T5 α OH-tTCPR resultó en la acumulación de un subproducto recientemente caracterizado, una reorganización estructural compleja de taxadieno en el éter cíclico 5(12)-Oxa-3(11)-ciclotaxano (OCT) (Fig. 9)⁴⁷. El subproducto se acumulaba aproximadamente en cantidades iguales que el producto deseado taxadien-5 α -ol. La formación de OCT estaba mediada por una secuencia de reacción del citocromo P450 de *Taxus* sin precedentes que implicaba la oxidación y ciclaciones subsiguientes⁴⁷. Por lo tanto, parece probable que mediante manipulación de proteínas de las taxadieno 5 α -hidroxilasas, la conclusión de la reacción antes de la ciclación evitará la acumulación de tal subproducto indeseable y se podría lograr canalizar el flujo de taxadien-5 α -ol.

La productividad de la cepa 26-At24T5 α OH-tTCPR se redujo significativamente con relación a la de la producción taxadieno por parte de la cepa parental 26 (~ 300 mg/L) con un aumento concomitante en la acumulación del metabolito no caracterizado anteriormente descrito. No se observó acumulación de taxadieno. Aparentemente, la introducción de un plásmido de copias medias adicional (10 copias, p10T7) que porta la construcción At24T5 α OH-tTCPR alteró el equilibrio cuidadosamente manipulado en la vía de aguas arriba y de aguas abajo de la cepa 26. Se llevaron a cabo fermentaciones a pequeña escala en biorreactores para cuantificar la producción de alcohol por parte de la cepa 26-At24T5 α OH-tTCPR. El perfil de evolución en el tiempo de la acumulación de taxadien-5 α -ol (Fig. 5d) indica una producción de alcohol de hasta 58 \pm 3 mg/L con una cantidad igual del subproducto OCT producido. La producción de alcohol observada era ~2400 veces mayor que la producción previa en *S. cerevisiae*¹⁷. Aumentos adicionales de la producción de taxadien-5 α -ol son probablemente posibles a través de la optimización de la vía y la manipulación de proteínas.

El enfoque multivariado-modular de la optimización de la vía ha proporcionado muy altas cepas productoras de un precursor de Taxol crítico. Además de ello, las construcciones recombinantes han sido igualmente eficaces en la reorientación del flujo hacia la síntesis de otros compuestos farmacéuticos complejos tales como productos de mono-, sesqui- y di-terpenos (geraniol, linalool, amorfadieno y levopimaradieno) manipulados a partir de la misma vía (resultados no publicados). Por lo tanto, la manipulación de las vías realizada por los autores de la invención abre nuevas vías para bio-sintetizar productos naturales, especialmente en el contexto de terpenoides derivados microbianamente, para su uso como productos químicos y combustibles a partir de recursos renovables. Al centrarse en los precursores terpenoides universales IPP y DMAPP, era posible, primero, definir los módulos de vía críticos y, a continuación, modular la expresión tal como equilibrar de manera óptima los módulos de las vías para la conversión de precursor sin fisuras y la acumulación mínima de compuestos intermedios. Este enfoque parece ser más eficaz que las búsquedas combinatorias de grandes espacios genéticos y tampoco depende de un rastreo de alto rendimiento.

La vía MEP está energéticamente equilibrada y, por lo tanto, en conjunto es más eficiente en la conversión de glucosa o glicerol en isoprenoides. Sin embargo, durante los últimos 10 años, muchos intentos de manipular la vía MEP en *E. coli* para aumentar el suministro de los precursores claves IPP y DMAPP para la sobreproducción de carotenoide^{28,47}, sesquiterpenoide²³ y diterpenoide⁶¹ se encontraron con un éxito limitado. Esta ineficiencia se atribuyó a efectos reguladores desconocidos asociados específicamente con la expresión de la vía MEP en *E. coli*²³. Aquí, los autores de la invención proporcionan evidencia de que tales limitaciones se correlacionan con la acumulación del metabolito indol, debido a la expresión no óptima de la vía, que inhibe la actividad de la vía isoprenoide. La sobreproducción de taxadieno (en condiciones de supresión de la formación de indol), establece la vía MEP como una vía muy eficiente para la biosíntesis de productos farmacéuticos y químicos de la familia de isoprenoides. Simplemente se tienen que equilibrar cuidadosamente las vías modulares según lo sugerido por el enfoque de los autores de la invención de manipulación de la vía multivariante-modular.

Para la producción microbiana con éxito de Taxol, la demostración de la decoración química del núcleo de taxadieno mediante la química de oxidación basada en P450 es esencial⁴¹. Monooxigenasas del citocromo P450 constituyen aproximadamente la mitad de las 19 etapas enzimáticas distintas en la vía biosintética de Taxol. De forma característica, estos genes muestran una inusual alta similitud de la secuencia entre sí (> 70%), pero una baja similitud (<30%) con otros P450s de vegetales¹⁴. Debido a la similitud aparente entre monooxigenasas de Taxol, que expresan la actividad adecuada para llevar a cabo la química de oxidación específica de P450 fue un desafío particular. A través de la manipulación TM y la construcción de una enzima quimera artificial con el participante

redox (TCPR), el citocromo P450 de Taxol, taxadieno 5 α -hidroxilasa, se expresaba funcionalmente en *E. coli* y mostró convertir de manera eficiente taxadieno en el correspondiente producto alcohol *in vivo*. Estudios previos *in vitro* han descrito el mecanismo de convertir taxadieno en taxadien-5 α -ol por parte de la enzima taxadieno 5 α -hidroxilasa nativa, pero no han discutido la misma conversión *in vivo*⁴². Esta reacción de oxigenación y transposición implica la abstracción de hidrógeno de la posición C20 del taxadieno para formar un intermedio radical alílico, seguido de la inserción regio-y estereo-específica de oxígeno en la posición C5 para proporcionar el derivado de alcohol (Fig. 5a). La modesta abundancia observada de la enzima en células de *Taxus* y los bajos valores k_{cat} sugirieron que la etapa de 5 α -hidroxilación de la biosíntesis de Taxol es lenta con relación a los oxigenaciones aguas abajo y acilaciones en la vía Taxol⁴¹. Por lo tanto, la manipulación de esta etapa es clave para la síntesis de Taxol, especialmente en el contexto de la manipulación funcional de P450s de Taxol en un huésped procarionta tal como *E. coli*. Además, esta etapa era limitante en los esfuerzos previos de la construcción de la vía en levaduras¹⁷. La construcción manipulada en este estudio demostró una conversión > 98% de taxadieno *in vivo* con la acumulación de producto a ~ 60 mg/L, una mejora de 2400 veces sobre la expresión previa heteróloga en levaduras. Por lo tanto, este estudio ha tenido éxito no sólo en la síntesis de cantidades significativamente mayores de productos intermedios de Taxol claves, sino que también proporciona la base para la síntesis de subsiguientes metabolitos en la vía mediante una química de P450 similar.

Estudios previos sobre la relación estructura-actividad en Taxol han demostrado que las alteraciones hechas por eliminación o adición de algunos de sus grupos funcionales no cambiaron materialmente la actividad de Taxol^{1,48}. Estos estudios, sin embargo, fueron limitados debido a la capacidad limitada de introducir cambios mediante síntesis química. La disponibilidad de una ruta microbiana para la síntesis de Taxol ampliará drásticamente el espacio de modificaciones químicas que se pueden examinar, aumentando así la probabilidad de identificar candidatos a fármacos más potentes. Esto ofrece nuevas oportunidades interesantes para el desarrollo de fármacos, especialmente cuando se considera que este tipo de candidatos a fármacos también se asociará con una ruta de producción eficiente.

En las últimas décadas, el Taxol ha generado más interés dentro de la comunidad científica y pública en general que cualquier otro candidato fármaco de producto natural¹⁰. Una de las principales crisis de suministro se predice a partir del incremento proyectado del uso de Taxol o análogos de Taxol para la quimioterapia del cáncer, lo que requiere nuevas vías de producción, tales como la manipulación de la maquinaria biosintética de Taxol en microbios⁸. Aunque se han aislado algunos hongos endófitos de especies de *Taxus* capaces de producir Taxol de forma natural, estos sistemas microbianos todavía tienen que demostrar una idoneidad para la producción sostenible del fármaco⁴⁹. Los resultados reseñados aquí representan una etapa disruptiva hacia un Taxol derivado de microbios o un precursor de Taxol, mediante la eliminación de los cuellos de botella en la vía precursor comprometido. Además de ello, el ensamblaje de una vía sintética ofrece nuevas posibilidades para adaptar análogos de Taxol por manipulación selectiva de la vía, alterando de este modo la estructura de taxano. Estos desarrollos plantean optimismo para una ruta microbiana para la producción rentable de Taxol o precursores de Taxol adecuados.

Tabla 1. Clasificación de la expresión de las vías de aguas arriba y aguas abajo en unidades arbitrarias (a.u.). Los niveles de expresión de la vía MEP y GGPP sintasa/taxadieno sintasa se estimaron utilizando los valores publicados de las fuerzas del promotor y el número de copias. Las fuerzas del promotor se calcularon como $trc = 1$, $T5 = 1,96$, $T7 = 4,97$, basado en Brosius et. al. y Brunner et. al.^{33,34}. El número de copias de genes fue asignado por los números de copias publicados para origen de la replicación para los diferentes plásmidos utilizados, y se utilizó una copia para integraciones³⁵⁻³⁷. La expresión total se calculó como el producto de la fuerza del promotor y el número de copias del gen. A la expresión nativa de la vía MEP se le asignó arbitrariamente un valor de uno, y se asumió un cambio en el orden del operón de GGPP sintasa y taxadieno que afecta a la expresión taxadieno sintasa en un 20%³⁵. Estas estimaciones de los esfuerzos totales guiaron a la manipulación de la expresión. E – *E. coli* K12MG1655 con dos deleciones $\Delta recA\Delta endA$; EDE3 – K12 MG1655 $\Delta recA\Delta endA$ con una T7 ARN polimerasa (DE3) integrada; operón MEP – *dxs-idi-ispDF*; operón GT-GPPS-TS; operón TG – TS-GPPS; Ch1 – 1 copia en el cromosoma; Trc – promotor *trc*; T5 – promotor T5; T7 – promotor T7; p5 – ~ 5 copias del plásmido (pSC101); p10 – ~ 10 copias del plásmido (p15A); y p20 – ~ 20 copias del plásmido (pBR322).

Tabla 1

Cepa n°	Modulación de la vía	MEP aguas arriba				GT o TG aguas abajo				Taxadieno (mq/L)	
		Construcción (además de copia nativa) §	Copias	Promotor	Intensidad de expresión* (a.u.)	Construcción	GT/TG	Promotor	Intensidad de expresión* (a.u.)	Media	SD
1	Ep20TrcGT	N/A	0	0	1 [¶]	pBR322	20	1.00	20	0.02	0.01
2	ECh1TrcMEPp20TrcGT	Chr. §	1	1	2	pBR322	20	1.00	20	16.00	1.59
3	Ep5TrcMEPp20TrcGT	pSC101	5	1	6	pBR322	20	1.00	20	2.55	0.21
4	Ep10TrcMEPp20TrcGT	p15A	10	1	11	pBR322	20	1.00	20	1.93	0.323
5	Ep20TrcTG	N/A	0	0	1	pBR322	20	1.20	24	0.19	0.01
6	Ep20TSGT	N/A	0	0	1	pBR322	20	1.96	39	4.36	0.533
7	Ep20TSGTTrcT	N/A	0	0	1	pBR322	20	2.96	59	1.74	0.265
8	ECh1TrcMEPp20TrcTG	Chr.	1	1	2	pBR322	20	1.20	24	45.44	2.28
9	ECh1TrcMEPp20TSGT	Chr.	1	1	2	pBR322	20	1.96	39	16.52	0.84
10	ECh1TrcMEPp20TSGT-TrcT	Chr.	1	1	2	pBR322	20	2.96	59 [¶]	2.52	0.30
11	Ep5TrcMEPp20TrcTG	pSC101	5	1	6	pBR322	20	1.20	24	7.41	0.53
12	Ep5TrcMEPp20TSGT	pSC101	5	1	6	pBR322	20	1.96	39	21.23	5.86
13	Ep5TrcMEPp20TSIG-TrcT	pSC101	5	1	6	pBR322	20	2.96	59	1.40	0.10
14	Ep10TrcMEPp20TrcTG	p15A	10	1	11	pBR322	20	1.20	24	2.36	0.29
15	Ep10TrcMEPp20TSGT	p15A	10	1	11	pBR322	20	1.96	39	8.91	2.94
16	Ep10TrcMEPp20TSIG-TrcT	p15A	10	1	11	pBR322	20	2.96	59	3.40	0.39

17	EDE3p10TrcMEPp5T7TG	p15A	10	1	11	pSC101	5	5.96	31	125.00	8.37
18	EDE3p20TrcMEPp5T7TG	pBR322	20	1	21	pSC101	5	5.96	31	58.00	3.07
19	EDE3p20T5MEPp5T7TG	pBR322	20	1.96	40	pSC101	5	5.96	31	44.00	2.88
20	EDE3p20T7MEPp5T7TG	pBR322	20	4.97	100	pSC101	5	5.96	31	32.00	6.63
21	EDE3p5TrcMEPp10T7TG	pSC101	5	1	6	p15A	10	5.96	61	7.00	1.40
22	EDE3p20TrcMEPp10T7TG	pBR322	20	1	21	p15A	10	5.96	61	59.00	5.57
23	EDE3p20T5MEPp10T7TG	pBR322	20	1.96	40	p15A	10	5.96	61	58.00	5.68
24	EDE3p20T7MEPp10T7TG	pBR322	20	4.97	100	p15A	10	5.96	61	20.00	0.73
25	EDE3p5T7TG	N/A	0	0	1	pSC101	5	5.96	31	19.00	8.23
26	EDE3Ch1TrcMEPp5T7TG	Chr.	1	1	2	pSC101	5	5.96	31	297.00	10.21
27	EDE3Ch1T5MEPp5T7TG	Chr	1	1.96	3	pSC101	5	5.96	31	163.00	10.84
28	EDE3Ch1T7MEPp5T7TG	Chr	1	4.97	6	pSC101	5	5.96	31	26.00	0.32
29	EDE3p10T7TG	N/A	0	0	1	p15A	10	5.96	61	8.00	0.39
30	EDE3Ch1TrcMEPp10T7TG	Chr	1	1	2	p15A	10	5.96	61	30.00	1.59
31	EDE3Ch1T5MEPp10T7TG	Chr	1	1.96	3	p15A	10	5.96	61	40.00	0.56
32	EDE3Ch1T7MEPp10T7TG	Chr	1	4.97	6	p15A	10	5.96	61	17.00	0.41

* Se dio un valor de 1 para representar las copias nativas de la vía MEP.

‡ La construcción MEP está localizada en el cromosoma.

* p20T5GT-TrcT - Una copia adicional del gen T bajo control separado del promotor (Trc) junto con el operón GT (bajo el promotor T5) en el mismo plásmido. Para el cálculo de la intensidad, se añadió el valor equivalente a dos operones separados (TrcT + T5GT = (20x1.96 + 20x1 = 59)) ya que los estudios demostraron que la expresión de T estaba limitada en comparación con G.

Nº	Plásmido	Origen de replicación	Marcador antibiótico
1	p20T7MEP	pBR322	Amp
2	p20TrcMEP	pBR322	Amp
3	p20T5MEP	pBR322	Amp
4	p20T7MEPKmFRP	pBR322	Km
5	p20T5MEPKmFRP	pBR322	Km
6	p20TrcMEPKm-FRP	pBR322	Km
7	p10TrcMEP	p15A	Cm
8	p5TrcMEP	SC101	Spect
9	p20TrcGT	pBR322	Amp
10	p20TrcTG	pBR322	Amp
11	p20T5GT	pBR322	Amp
12	p10T7TG	p15A	Cm
13	p5T7TG	SC101	Spect
14	p10At8T5αOH-tTCPR	p15A	Cm
15	p10At24T5αOH-tTCPR	p15A	Cm
16	p10At42T5αOH-tTCPR	p15A	Cm

Tabla 2. Detalle de todos los plásmidos construidos para el estudio

SEQ ID NO	Nombre del Cebador	Secuencias
1	dxsNdel(s)	CGGCATATGAGTTTTGATATTGCCAAATACCCG
2	dxsNhel(a)	CGGCTAGCTTATGCCAGCCAGGCCTTGATTTTG
3	idiNhel (s)	CGCGGCTAGCGAAGGAGATATACATATGCAAACGGAAC ACGTCATTTTATTG
4	idiEcoRI(a)	CGGAATTCGCTCACAACCCCGCAAATGTCGG
5	ispDFEcoRI(s)	GCGAATTCGAAGGAGATATACATATGGCAACCACTCATT TGGATGTTTG
6	ispDFXhol(a)	GCGCTCGAGTCATTTTGTTGCCTTAATGAGTAGCGCC
7	dxsidiispDFNcol(s)	TAAACCATGGGTTTTGATATTGCCAAATACCCG
8	dxsidiispDFKpnI(a)	CGGGGTACCTCATTTTGTTGCCTTAATGAGTAGCGC
9	dxsidiispDFXhol(a)	CGGCTCGAGTCATTTTGTTGCCTTAATGAGTAGCGC
10	T5AgiI(s)	CGTAACCGGTGCCTCTGCTAACCATGTTTCATGCCTTC
11	T5Nhel(a)	CTCCTTCGCTAGCTTATGCCAGCC
52	GGPPSNcol(s)	

ES 2 575 790 T3

SEQ ID NO	Nombre del Cebador	Secuencias
		CGTACCATGGTTGATTTCAATGAATATATGAAAAGTAAG GC
12	GGPPSEcoRI(a)	CGTAGAATTCACTCACAACCTGACGAAACGCAATGTAATC
13	TXSEcoRI(s)	CGTAGAATTCAGAAGGAGATATACATATGGCTAGCTCTA CGGGTACG
14	TXSsalI(a)	GATGGTCGACTTAGACCTGGATTGGATCGATGTAAAC
15	TXSNcoI(s)	CGTACCATGGCTAGCTCTACGGGTACG
16	TXSEcoRI(a)	CGTAGAATTCCTTAGACCTGGATTGGATCGATGTAAAC
17	GGPPSEcoRI(s)	CGTAGAATTCAGAAGGAGATATACATATGTTTGATTTCA ATGAATATATGAAAAGTAAGGC
18	GGPPSSalI(a)	GATGGTCGACTCACAACCTGACGAAACGCAATGTAATC
19	TSXhoI(a)	GATGCTCGAGTTAGACCTGGATTGGATCGATGTAAAC
20	pTrcSal(s)	GCCGTCGACCATCATCATCATC
21	pTrcXba(a)	GCAGTCTAGAGCCAGAACCGTTATGATGTCGGCGC
22	pCLBspEI(s)	CGTGTCCGGAGCATCTAACGCTTGAGTTAAGCCGC
23	pCLXbal(a)	GCAGTCTAGAGGAAACCTGTCGTGCCAGCTGC
24	KmFRPXhoI(s)	GACGCTCGAGGAGCAATAACTAGCATAACCCCTTGGGGC CTCTAAACGGGTCTTGAGGGGTTTTTTGCTTGTGTAGGCT GGAGCTGCTTCG
25	KmFRPSalI(a)	GACGAGTACTGAACGTCGGAATTGATCCGTCGAC
26	KmFRPSacI(s)	GACGGAGCTCGAGCAATAACTAGCATAACCCCTTGGGGC CTCTAAACGGGTCTTGAGGGGTTTTTTGCTTGTGTAGGCT GGAGCTGCTTCG
27	IntT7T5(s)	ATGACGATTTTTGATAATTATGAAGTGTGGTTTGTCAATTG CATTAAATTGCGTTGCGCTCACTG
28	IntTrc(s)	ATGACGATTTTTGATAATTATGAAGTGTGGTTTGTCAATTG GCATCCGCTTACAGACAAGCTGTG
29	Int(a)	TTAGCGACGAAACCCGTAATACACTTCGTTCCAGCGCAG CCGACGTCGGAATTGATCCGTCGAC
30	CYP17At8AANdel(s)	

SEQ ID NO	Nombre del Cebador	Secuencias
		CGTACATATGGCTCTGTTATTAGCAGTTTTTGTGGCGAAA TTAACGAAGTAACCCAGC
31	CYP17At24AANdel (s)	CGTACATATGGCTCTGTTATTAGCAGTTTTTTTAGCATC GCTTTGAGTGCAATTG
32	CYP17At42AANdel (s)	CGTACATATGGCTCTGTTATTAGCAGTTTTTTTCGCTCG AAACGTCATAGTAGCCTG
33	CYPLinkBamHI(a)	CGCGGGATCCGGTGCTGCCCGGACGAGGGAACAGTTTGA TTGAAAACCC
34	CPRBamHI(s)	CGCGGGATCCCGCCGTGGTGGAAAGTGATACACAG
35	CPRSall(a)	CGCGGTCGACTTACCAAATATCCCGTAAGTAGCGTCCAT C
36	DXS qPCR (s)	ATTCAAAAGCTTCCGGTCCT
37	DXS qPCR (a)	ATCTGGCGACATTCGTTTTC
38	TS qPCR (s)	GACGAACTGTCACCCGATTT
39	TS qPCR (a)	GCTTCGCGGGTAGTAGACAG
40	rrsA qPCR (s)	AGGCCTTCGGGTTGTAAAGT
41	rrsa qPCR (a)	ATTCCGATTAACGCTTGAC

Tabla 3. Detalles del cebador utilizado para la clonación de plásmidos, suministro cromosómico de la vía MEP y mediciones qPCR.

Tabla 4. Secuencias de nucleótidos optimizadas en Proteínas y Codones

GGPP sintasa

MFDNFNEYMKSKAVAVDAALDKAIPLEYPEKIHESMRYSLLAGGKRVRPALCIAACE
 LVGGSQDLAMPACAMEMIHTMSLIHDDLPCMDNDDFRRGKPTNHKVFGEDEVAVL
 AGDALLSFAFEHIAVATSKTVPSDRTLRVISELGKIIGSQGLVGGQVVDITSEGDANV
 DLKTLEWIIHKTAVLLLECSVVS GGILGGATEDEIARIRRYARCVGLLFQVVDILDV
 TKSSSEELGKTAGKDLLTDKATYPKLMGLEKAKEFAAELATRAKEELSSFDQIKAAPL
 LGLADYIAFRQN (SEQ ID NO:42)

ATGTTTGAATTICAATGAATATATGAAAAGTAAGGCTGTTGCGGTAGACGCGGCTC
 TGGATAAAGCGAATCCGCTGGAATATCCCGAGAAGATTCACGAATCGATGCGCT
 ACTCCCTGTTAGCAGGAGGGAAACGCGTTCGTCGGCATTATGCATCGCGGCCCTG
 TGAACTCGTCGGCGGTTACAGGACTTAGCAATGCCAACTGCTTGCACAATGGA
 AATGATTCACACAATGAGCCTGATTCATGATGATTGCTTGCATGGACAACGAT
 GACTTTCGGCGCGGTA AACCTACTAATCATAAGGTTTTTGGCGAAGATACTGCAG
 TGCTGGCGGGCGATGCGCTGCTGTCGTTTGCCTTCGAACATAICGCCGTCGCGAC
 CTCGAAAACCGTCCCCTCGGACCGTACGCTTCGCGTGATTTCCGAGCTGGGAAAG
 ACCATCGGCTCTCAAGGACTCGTGGGTGGTCAGGTAGTTGATATCACGCTTGAGG
 GTGACGCGAACGTGGACCTGAAAACCTGGAGTGGATCCATATTCACAAAACGG
 CCGTGCTGCTCGAATGTAGCGTGGTGT CAGGGGGGATCTTGGGGGGCGCCACGG
 AGGATGAAATCGCGCGTATTCGTCGTTATGCCCGCTGTGTTGGACTGTTATITCA
 GGTGGTGGATGACATCTGGATGTCACAAAATCCAGCGAAGAGCTTGGCAAGAC
 CGCGGGCAAAGACCTTCTGACGGATAAGGCTACATACCCGAAATTGATGGGCTT
 GGAGAAAGCCAAGGAGTTCGCAGCTGAACTTGCCACGCGGGCGAAGGAAGAAC
 TCTCTTCTTTTCGATCAAATCAAAGCCGCGCCACTGCTGGGCCTCGCCGATTACAT
 TGCGTTTCGTCAGAAC (SEQ ID NO:43)

Taxadieno sintasa

MSSSTGTSKVVSETSSITVDDIPRLSANYHGD LWHIHNVIQTLETFFRESSTYQERADE
 LVVKIKDMFNALGDGDISPSAYDTAWVARLATISSDGSEKPRFPQALNWVFNNQLQ
 DGSWGIESHFLCDRLNNTNSVIALSVWKTGHSQVQQAEFIAENLRLLNEEDELSP
 DFQIIFPALLQKAKALGINLPYDLPFIKYLS TTREARLTDVSAADNIPANMLNALEGL
 EEVIDWNKIMRFQSKDGSFLSSPASTACVLMNTGDEKCF TFLNNLLDKFGGCVPCM
 YSIDLLERLSLVDNIEHLGIGRHFKEIKGALDYVYRHW SERGIGWGRDSLVPDLNTT
 ALGLRTL RMHGYNVSSDVLNNFKDENGRFFSSAGQTHVELRSV VNLFRASDLAFP
 ERAMDDARKFAEPLYREALATKISTNTKLFKEIEYVVEY PWHMSIPRLEARSYIDSYD
 DNYVWQRKTLYRMP SLSNSKCLELAKLDFNIVQSLHQEELKLLTRWWKESGMADI
 NFTRHRVAEVYFSSATFEPEYSATRIAFTKIGCLQVLFDDMADIFATLDELKSFTGV
 KRWDTSLLHEIPECMQTCFKVWFKLMEEVNNDVVKVQGRDMLAHIRKPWELYFNC
 YVQEREWLEAGYIPTFEEYLKTYAISVGLG PCTLQPILLMGELVKDDVVEKVHYP
 SNMFELVLSWRLTNDTKTYQAEKARGQQA SGIACYMKDNPGATEEDA IKHICRVVDR
 ALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIF YKFDGYGIANEEIKDYIRKVYIDPI
 QV (SEQ ID NO:44)

ATGTCTAGCTCTACGGGTACGTCTAAAGTCGTGAGTGAAACCTCATCGACGATCG
 TGGACGATATCCACGCTTGTCGGCGAACTATCATGGAGATCTGTGGCATCATAA

CGTCATTGACATTGGAAACCCCGTTTCGCGAAAAGTAGCACCTACCAGGAACG
 GGCAGATGAATTAGTCGTGAAAATCAAAGATATGTTTAAATGCATTAGGAGATGG
 AGACATCTCGCCAGCGCATATGATACGGCGTGGGTGGCTCGGTTGGCCACGATT
 AGCTCCGATGGCAGTGAAAAGCCGCGTTTCCCGCAGGCGCTGAACTGGGTGTTT
 AATAATCAATTGCAGGATGGCAGCTGGGGCATTGAATCTCACTTTAGCCTCTGTG
 ACCGGTTACTCAACACGACAAACTCCGTAATTGCGTTGTCAGTTTGGAAAACGGG
 CCATAGCCAGGTTC AACAGGGCGCGGAATTTATCGCTGAAAATCTGCGCCTGCTG
 AACGAGGAGGACGAACTGTACCCCGATTTTCAGATTATTTTTCCGGCTTACTCC
 AGAAAGCCAAAGCCTTAGGCATCAACCTGCCATATGATCTGCCGTTTCAATCAAGTA
 TCTGTCTACTACCCGCGAAGCCCGTCTCACTGACGTTCTGCGGCGGGCGGACAAT
 ATTCCAGCGAACATGCTGAACGCACTGGAAGGGCTGGAAGAGGTTATCGACTGG
 AATAAAATCATGCGCTTCCAAAAGCAAGGACGGTAGCTTCTTAAGCAGCCAGCA
 TCTACTGCTTGTGTTCTGATGAATACCGGAGACGAAAAGTGCTTTACGTTTCTGA
 ACAATCTGCTGGACAAAATTTGGGGGTTGTGTTCCCTTGTA'IGTA'ITCCA'ITGATCTG
 TTGGAACGTCTGTCGCTGG'ICGATAACATTGAACACTTAGGTA'PCGGCCGCCACT
 TCAAACAAGAAATCAAGGGGGCGTTGGATTATGTATACCGTCA'ITGGAGCGAGC
 GTGGTATTGGTTGGGGGCGCGATAGCTTGGTACCTGATCTGAACACCACTGCTTT
 GGGACTGCGCACTCTTCGTATGCACGGATACAACGTTAGTTCCGATGTCCTCAAT
 AATTTCAAGGACGAGAACGGCCGTTTTTTTCAGCTCGGCCGGTCAAGACGATGTTG
 AACTGCGGTCCGTAGTCAATCTCTTTTCGCGCTAGTGAATCTGGCCTTCCCGACGA
 GCGCGCTATGGACGATGCACGGAAGTTTGCCGAGCCGTATCTCCGCGAAGCCCT
 GGCCACCAAAA'ITTC AACCAACACCAAGCTTTTCAAAGAAATGAGTATGTAGT
 AGAGTATCCGTGGCATA'GTCTA'ITCCGCGCCTGGAAGCCCGC'ICGTATATCGAT
 TCTTACGATGACAATTA'GTGTGGCAACGCAAAACACTGTACCGTATGCCCAGCC
 TGTCAAATAGTAAGTGTCTGGAGCTGGCGAAACTGGATTTCAACATTGTGCAATC
 CCTGCACCAAGAAGAGCTGAAA'ITACTGACTCGCTGGTGGAAAGGAATCCGGCAT
 GGCAGACATCAATTTIACGCGTACCCGTGTTGACAGAGGTGTACTTCTCCTCGGCG
 ACCTTTGAGCCGGAGTATTCGGCCACACGTATTGCATTTACCAAGATTGGCTGCC
 TTCAGGTGCTTTTTGACGATATGGCGGATA'ITTTT'GCGACACTTGATGAGCTTAA
 ATCATTTACCGAAGGCGTGAAGCGTIGGGATACCTCTCTG'ITGCA'TGAAATCCCC
 GAATGTATGCAGACCTGC'ITCAAAGTTTGGTTCAAACCTGATGGAAGAAGTGAAC
 AACGACGTCGTGAAAAGTTCAAGGTCGTGATA'GTTAGCACACATCCGCAAGCCG
 TGGGAACTCTATTTCAAATTGCTATGTGCAGGAGCGTGAATGGTTAGAAGCGGGCT
 ACATTCCTACCTTCGAAGAGTACTTAAAAACCTATGCCATTTCCGTCGGTTTAGG
 CCCGTGCACTCTGCAGCCTATCTTGCTGATGGGTGAGCTGGTAAAGGATGATGTG
 GTGGAAAAAGTTCACTACCCGTGCAATA'G'ITTTGAACTGGTAAGTCTGAGTTGGC
 GTCTGACAAACGACACCAAAACGTACCAGGCAGAAAAGGCACGTGGGCAACAG
 GCAAGCGGTATCGCGTGTATATGAAGGATAATCCGGGCGCTACTGAGGAAGAT
 GCCATTAAGCATATCTGCCGTGTGTGGATCGCGCTCTTAAAGAAGCGTCATTTCG
 AATATTTTAAACCTAGTAATGATATTTCCGATGGGTTGTAAGTCA'ITTCATTTICAAT
 CTTCGCC'ITGTGCGTGCAAAATTTTTTACAAATTTATTGACGGCTACGGAATCGCCA
 ACGAAGAAATCAAAGACTATATTCGTAAAGTTTACATCGATCCAATCCAGGTC
 (SEQ ID NO:45)

Citocromo P450 Taxadieno 5 α -hidroxilasa (*T5 α OH*)

MDALYKSTVAKFNEVTQLDCSTESFSIALSAIAGILLLLLLFRSKRHSSLKLPKGKLGIP
 FIGESFIFLRALRSNSLEQFFDERVKKFGLVFKTSLIGHPTVVLCGPAGNRLIISNEEKL
 VQMSWPAQFMKLMGENSVATRRGEDHIVMRSALAGFFGPGALQS YIGKMNTEIQS

HINEKWKGKDEVNVLPLVRELVFNISAILFFNIYDKQEQRDLHKLLETILVGSFALPID
 LPGFGRALQGRAKLNKIMLSLIKRRKEDLQSGSATATQDLLSVLUTFRDDKGTPL
 TNDEILDNFSSLLHASDYDTTSPMALIFKLLSSNPECYQKVYQEQLEILSNKEEGEET
 WKDLKAMKYTWQVAQETLRMFPPVFGTFRKAITDIQYDGYTIPKGWKLWTTYST
 HPKDLYFNEPEKFMPSRFDQEGKHVAPYTFLPFGGGQRSCVGVWFEFSKMEILLFVHHF
 VKTFSSYTPVDPDEKISGDPLPPLPSKGFSLKLPFRP (SEQ ID NO:46)

ATGGATGCCCTCTATAAGTCTACCGTGGCGAAATTTAACGAAGTAACCCAGCTGG
 ATTGCAGCACTGAGTCATTTAGCATCGCTTTGAGTGCAATTGCCGGGATCTTGCT
 GTTGCTCCTGCTGTTTCGCTCGAAACGTCATAGTAGCCTGAAATTACCTCCGGGC
 AAAC'TGGGCATCCGTTTATCGGTGAGTCCTTTATTTTTTTCGCGCGCTGCGCAG
 CAAITTCCTGGAACAGTCTTTGATGAACGTGTGAAGAAGTTCGGCCTGGTATTT
 AAAACGTCCCTTATCGGTACCCGACGGTTGTCCTGTGCGGGCCCCGACGGTAATC
 GCCTCATCCTGAGCAACGAAGAAAAGCTGGTACAGATGTCCTGGCCGGCGCAGT
 TTATGAAGCTGATGGGAGAGAAGTCAAGTTCGACCCCGCGTGGTGAAGATCACA
 TTGTTATGCGCTCCGCGTGGCAGGCTTTTTCGGCCCCGGGAGCTCTGCAATCCTAT
 ATCGGCAAGATGAACACGGAAATCCAAAGCCATATTAATGAAAAGTGGAAAGG
 GAAGGACGAGGTTAATGTCTTACCCCTGGTGCAGGAACTGGTTTTTAACATCAGC
 GCTATTCTGTTCTTTAACATTTACGATAAGCAGGAACAAGACCGTCTGCACAAGT
 TGTTAGAAACCATTCTGGTAGGCTCGTTTGCCTTACCAATTGATTTACCGGGTTTC
 GGGTTTACCCGCGCTTTACAAGGTCGTGCAAAACTCAATAAAATCATGTTGTCCG
 TTATTAATAAACGTAAGAGGACTTACAGTCGGGATCGGCCACCCGCGACGCAGG
 ACCTGTTGTCTGTGCTTCTGACTTTCGCTGATGATAAGGGCACCCCGTTAACCAA
 TGACGAAATCCTGGACAACTTTAGCTACTGCTTACGCCTCTTACGACACCACG
 ACTAGTCCAATGGCTCTGATTTTCAAATTAAGTCAAGTAACCCTGAATGCTATC
 AGAAAGTCGTGCAAGAGCAACTCGAGATTCTGAGCAATAAGGAAGAAGGTGAA
 GAAATTACCTGGAAAGATCTTAAGGCCATGAAATACACGTGGCAGGTTGCGCAG
 GAGACACTTCGCATGTTTCCACCGGTGTTTCGGGACCTTCCGCAAAGCGATCACGG
 ATATTCAGTATGACGGATACACAATCCCGAAAGGTTGGAAACTGTTGTGGACTA
 CCTATAGCACTCATCCTAAGGACCTTTACTTCAACGAACCGGAGAAATTTATGCC
 TAGTCGTTTCGATCAGGAAGGCAAACATGTTGCGCCCTATACCTTCCCTGCCCTTT
 GGAGGCGGTCAGCGGAGTTGTGTGGGTTGGGAGTTCTCTAAGATGGAGATTCTC
 CTCCTTCGTGCATCATTTTCGTGAAAACATTTTCGAGCTATACCCCGGTCGATCCCG
 ATGAAAAAATTTCCGGCGATCCACTGCCGCCGTTACCGAGCAAAGGGTTTTCAAT
 CAAACIGTTCCTTCGTCCG (SEQ ID NO:47)

NADPH:citocromo P450 reductasa de Taxus (TCPR)

MQANSNTVEGASQKSLDISRLDHIFALLNGKGGDLGAMTGSALILTENSQNLMI
 LITLALAVLVACVFFVWRRGGSDTQKPAVRPTPLVKEEEDDEEDSAKKKVTIFFGT
 QTGTAEGFAKALAEAKARYEKA VFKVVDLDNYAADDEQYEEKLKEKLAFFMI.A
 TYGDGEPTDNAARFYKWFLE GKEREPWLSDLTYGVFGLGNRQYEHFNKVAKAVDE
 VLIEQGAKRLVPVGLGDDDDQIEDDFAWREQVWPELDQLLRDEDDEPTSAPTYTA
 APEYRVEIYDSVVSVEETHALKQNGQAVYDIHHPCRSNVAVRRELHTPLSDRSCIH
 LEFDISDTGLIYETGDHVGVTENSIEVVEEAKLLGYQLDTIFSVHGDKEDGTPGG
 SSLPPFPGPCTLRALARYADLLNPPRKA AFLALAAHASDPAEAERLKFSSPAGKD
 EYSQWVTASQRSLEIMAEFSAKPPLGVFFAAIAPRLQPRYYSSISSPRFAPSRIHVTC
 ALVYGPSPTGRIHKGVC SNWMKNSLPSEETHDCSWAPVFRQSNFKLPADSTTPIVM
 VGPGTGFAFRGFLQERAKLQEAGEKLGPAVLFFGCRNRQMDYIYEDELKGYVEKG

ILTNLIVAFSREGATKEYVQHKMLEKASDTWSLIAQGGYLYVCGDAKGMARDVHR
 TLHTIVQEQUESVDSSKAFLVKKLQMDGRYLRDIW (SEQ ID NO:48)

ATGCAGGCCGAATTCTAATACGGTTGAAGGCGCGAGCCAAGGCAAGTCTCTTCTG
 GACATTAGTCGCCTCGACCATATCTTCGCCCTGCTGTTGAACGGGAAAGGCGGAG
 ACCTTGGTGCGATGACCGGGTCGGCCTTAATTCTGACGGAAAAATAGCCAGAACTT
 GATGATTCTGACCACTGCGCTGGCCGTTCTGGTCGCTTGCCTTTTTTTTTTCGTTT
 GCGCGCGTGGTGAAGTGATACACAGAAGCCCCCGCTACGTCCACACCTCTTG
 TAAAGAAGAGGACGAAGAAGAAGAAGATGATAGCGCCAAGAAAAAGGTCACA
 ATATTTTTTGGCACCCAGACCGGCACCGCCGAAGGTTTTCGCAAAGGCCTTAGCTG
 AGGAAGCAAAGGCACGTTATGAAAAGGCGGTATTTAAAGTCGTGGATTTGGATA
 ACTATGCAGCGGATGACGAACAGTACGAAGAGAAGTTGAAAAAGGAAAAGCTA
 GCGTTCTTCATGCTCGCCACCTACGGTGACGGCGAACCGACTGATAATGCCGCTC
 GCTTTTATAAATGGTTTTCTCGAGGGTAAAGAGCGCGAGCCATGGTTGTCAGATCT
 GACTTATGGCGTGTTTGGCTTAGGTAACCGTCAGTATGAACACTTTAACAAGGTC
 GCGAAAGCGGTGGACGAAGTGCTCATTGAACAAGGCGCCAAACGTCTGGTACCG
 GTAGGGCTTGGTGATGATGATCAGTGCATTGAGGACGACTTCACTGCCTGGAGA
 GAACAAGTGTGGCCTGAGCTGGATCAGCTTTACGTGATGAAGATGACGAGCCG
 ACGTCTGCGACCCCGTACACGGCGGCTATTCCAGAATACCGGGTGGAAATCTAC
 GACTCAGTAGTGTGCGTCTATGAGGAAACCCATGCGCTGAAACA AAAATGGACAA
 GCCGTATACGATATCCACCACCCGTCGCGCAGCAACGTGGCAGTACGTCTGAG
 CTGCATACCCCGCTGTCGGATCGTAGTTGTATTCATCTGGAATTCGATAITAGTG
 AACTGGGTTAATCTATGAGACGGGCGACCACGTTGGAGTTCATACCGAGAATTC
 AATTGAAACCGTGAAGAAGCAGCTAAACTGTTAGGTTACCAACTGGATACAAT
 CTTACAGCGTGCATGGGGACAAGGAAGATGGAACACCATTGGGCGGGAGTAGCCT
 GCCACCGCCGTTTCCGGGGCCCTGCACGCTGCGGACGGCGCTGGCACGTTACGC
 GGACCTGCTGAACCCTCCGCGCAAAGCCGCTTCTGGCACTGGCCGCACACGC
 GTCAGATCCGGCTGAAGCTGAACGCCTTAAATTTCTCAGTTCTCCAGCCGAAAA
 CCGAATTCGCCAGCGCGAAACCGCCGCTGGGAGTGTTTTTCGCCGCAATAGCGCC
 GCGCTTGCAACCTAGGTATTATAGCATCTCCTCCTCCCCGCTTTCGCGCCGTCCTC
 GTATCCATGTAACGTGCGCGCTGGTCTATGGTCCTAGCCCTACGGGGCGTATTCA
 TAAAGGTGTGTGCAGCAACTGGATGAAGAA TTCTTTGCCCTCCGAAGAAACCCA
 CGATTGCAGCTGGGCACCGGTC TTTGTGCGCCAGTCAAAC TTTAAACTGCCCGCC
 GATTCGACGACGCCAATCGTGATGGTTGGACCTGGAACCGGCTTCGCTCCATTT
 GCGGCTTCTTCAGGAACGCGCAAAACTGCAGGAAGCGGGCGAAAAATTGGGCC
 CGGCAGTGCTGTTTTTTGGGTGCCGCAACCGCCAGATGGATTACATCTATGAAGA
 TGAGCTTAAGGGTTACGTTGAAAAAGGTATTCTGACGAATCTGATCGTTGCATTT
 TCACGAGAAGGCGCCACCAAAGAGTATGTTTCAGCACAAGATGTTAGAGAAAGCC
 TCCGACACGTGGTCTTTAATCGCCCAGGGTGGTTATCTGTATGTTTGGCGTGATG
 CGAAGGGTATGGCCAGAGACGTACATCGCACCCATGCATACAATCGTTTCAGGAAC
 AAGAATCCGTAGACTCGTCAAAGCGGAGTTTTTAGTCAAAAAGCTGCAAAATGG
 ATGGACGCTACTTACGGGATATTTGG (SEQ ID NO:49)

Referencias

1. Kingston, D. G. The shape of things to come: structural and synthetic studies of taxol and related compounds. *Phytochemistry* 68, 1844-54 (2007).
- 5 2. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. & McPhail, A. T. Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc* 93, 2325-7 (1971).
3. Suffness M, W. M. Discovery and development of taxol. (ed. (ed), S. M.) (CRC Press, Boca Raton, 1995).
4. Nicolaou, K. C. et al. Total synthesis of taxol. *Nature* 367, 630-4 (1994).
- 10 5. Holton, R. A. et al. First total synthesis of taxol. 2. Completion of the C and D rings. *Journal of the American Chemical Society* 116, 1599-1600 (1994).
6. Walji, A. M. & MacMillan, D. W. C. Strategies to Bypass the Taxol Problem. *Enantioselective Cascade Catalysis, a New Approach for the Efficient Construction of Molecular Complexity*. *Synlett* 18, 1477-1489 (2007).
7. Holton RA, B. R., Boatman PD. Semisynthesis of taxol and taxotere (ed. M, S.) (CRC Press, Boca Raton, 1995).
- 15 8. Frense, D. Taxanes: perspectives for biotechnological production. *Applied microbiology and biotechnology* 73, 1233-1240 (2007).
9. Roberts, S. C. Production and engineering of terpenoids in plant cell culture. *Nature Chemical Biology* 3, 387-395 (2007).
- 20 10. Goodman, J. & Walsh, V. The story of taxol : nature and politics in the pursuit of an anticancer drug (Cambridge University Press, Cambridge ; New York, 2001).
11. Tyo, K. E., Alper, H. S. & Stephanopoulos, G. N. Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends Biotechnol* 25, 132-7 (2007).
12. Ajikumar, P. K. et al. Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol Pharm* 5, 167-90 (2008).
- 25 13. Jennewein, S. & Croteau, R. Taxol: biosynthesis, molecular genetics, and biotechnological applications. *Appl Microbiol Biotechnol* 57, 13-9 (2001).
14. Jennewein, S., Wildung, M. R., Chau, M., Walker, K. & Croteau, R. Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci U S A* 101, 9149-54 (2004).
- 30 15. Croteau, R., Ketchum, R. E. B., Long, R. M., Kaspera, R. & Wildung, M. R. Taxol biosynthesis and molecular genetics. *Phytochemistry Reviews* 5, 75-97 (2006).
16. Walker, K. & Croteau, R. Taxol biosynthetic genes. *Phytochemistry* 58, 1-7 (2001).
17. Dejong, J. M. et al. Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 93, 212-24 (2006).
- 35 18. Engels, B., Dahm, P. & Jennewein, S. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic engineering* 10, 201-206 (2008).

19. Chang, M. C. Y. & Keasling, J. D. Production of isoprenoid pharmaceuticals by engineered microbes. *Nature chemical biology* 2, 674-681 (2006).
20. Khosla, C. & Keasling, J. D. Metabolic engineering for drug discovery and development. *Nat Rev Drug Discov* 2, 1019-25 (2003).
- 5 21. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* 23, 612-6 (2005).
22. Chang, M. C., Eachus, R. A., Trieu, W., Ro, D. K. & Keasling, J. D. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat Chem Biol* 3, 274-7 (2007).
- 10 23. Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol* 21, 796-802 (2003).
24. Huang, Q., Roessner, C. A., Croteau, R. & Scott, A. I. Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. *Bioorg Med Chem* 9, 2237-42 (2001).
25. Kim, S. W. & Keasling, J. D. Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnol Bioeng* 72, 408-15 (2001).
- 15 26. Farmer, W. R. & Liao, J. C. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nature Biotechnology* 18, 533-537 (2000).
27. Farmer, W. R. & Liao, J. C. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. *Biotechnology progress* 17 (2001).
- 20 28. Yuan, L. Z., Rouviere, P. E., Larossa, R. A. & Suh, W. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metab Eng* 8, 79-90 (2006).
29. Jin, Y. S. & Stephanopoulos, G. Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metabolic Engineering* 9, 337-347 (2007).
30. Wang, H. H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* (2009).
- 25 31. Klein-Marcuschamer, D., Ajikumar, P. K. & Stephanopoulos, G. Engineering microbial cell factories for biosynthesis of isoprenoid molecules: beyond lycopene. *Trends in Biotechnology* 25, 417-424 (2007).
32. Sandmann, G. Combinatorial biosynthesis of carotenoids in a heterologous host: a powerful approach for the biosynthesis of novel structures. *ChemBioChem* 3 (2002).
33. Brosius, J., Erfle, M. & Storella, J. Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity. *J Biol Chem* 260, 3539-41 (1985).
- 30 34. Brunner, M. & Bujard, H. Promoter recognition and promoter strength in the *Escherichia coli* system. *Embo J* 6, 3139-44 (1987).
35. Nishizaki, T., Tsuge, K., Itaya, M., Doi, N. & Yanagawa, H. Metabolic engineering of carotenoid biosynthesis in *Escherichia coli* by ordered gene assembly in *Bacillus subtilis*. *Appl Environ Microbiol* 73, 1355-61 (2007).
- 35 36. Sorensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol* 115, 113-28 (2005).
37. Jones, K. L., Kim, S. W. & Keasling, J. D. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab Eng* 2, 328-38 (2000).

38. Hoffmann, F. & Rinas, U. Stress induced by recombinant protein production in *Escherichia coli*. *Advances in Biochemical Engineering Biotechnology* 89, 73-92 (2005).
39. Hoffmann, F., Weber, J. & Rinas, U. Metabolic adaptation of *Escherichia coli* during temperature-induced recombinant protein production: 1. Readjustment of metabolic enzyme synthesis. *Biotechnology and bioengineering* 80 (2002).
40. Chang, D. E., Smalley, D. J. & Conway, T. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. *Molecular microbiology* 45, 289-306 (2002).
41. Kaspera, R. & Croteau, R. Cytochrome P450 oxygenases of Taxol biosynthesis. *Phytochemistry Reviews* 5, 433-444 (2006).
42. Jennewein, S., Long, R. M., Williams, R. M. & Croteau, R. Cytochrome p450 taxadiene 5 α -hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. *Chem Biol* 11, 379-87 (2004).
43. Schuler, M. A. & Werck-Reichhart, D. FUNCTIONAL GENOMICS OF P450 S. *Annual Review of Plant Biology* 54, 629-667 (2003).
44. Leonard, E. & Koffas, M. A. G. Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. *Applied and Environmental Microbiology* 73, 7246 (2007).
45. Nelson, D. R. Cytochrome P450 and the individuality of species. *Archives of Biochemistry and Biophysics* 369, 1-10 (1999).
46. Jennewein, S. et al. Coexpression in yeast of *Taxus* cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis. *Biotechnol Bioeng* 89, 588-98 (2005).
47. Rontein, D. et al. CYP725A4 from yew catalyzes complex structural rearrangement of taxa-4(5),11(12)-diene into the cyclic ether 5(12)-oxa-3(11)-cyclotaxane. *J Biol Chem* 283, 6067-75 (2008).
48. Shigemori, H. & Kobayashi, J. Biological activity and chemistry of taxoids from the Japanese yew, *Taxus cuspidata*. *J Nat Prod* 67, 245-56 (2004).
49. Xu, F., Tao, W., Cheng, L. & Guo, L. Strain improvement and optimization of the media of taxol-producing fungus *Fusarium maire*. *Biochemical Engineering Journal* 31, 67-73 (2006).
50. Hefner, J., Ketchum, R. E. & Croteau, R. Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cells induced for taxol production. *Arch Biochem Biophys* 360, 62-74 (1998).
51. Wildung, M. R. & Croteau, R. A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. *J Biol Chem* 271, 9201-4 (1996).
52. Kodumal, S. J. et al. Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proceedings of the National Academy of Sciences* 101, 15573-15578 (2004).
53. Tyo, K. E. J., Ajikumar, P. K. & Stephanopoulos, G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nature Biotechnology* (2009).
54. Datsenko, K. A. & Wanner, B. L. 6640-6645 (*National Acad Sciences*, 2000).
55. Rost, B., Yachdav, G. & Liu, J. The predictprotein server. *Nucleic acids research* 32, W321 (2004).

56. Shalel-Levanon, S., San, K. Y. & Bennett, G. N. Effect of ArcA and FNR on the expression of genes related to the oxygen regulation and the glycolysis pathway in Escherichia coli under microaerobic growth conditions. *Biotechnology and bioengineering* 92 (2005).
57. Heinig, U. & Jennewein, S. Taxol: A complex diterpenoid natural product with an evolutionarily obscure origin. *African Journal of Biotechnology* 8,1370-1385 (2009).
58. Walji, A. M. & MacMillan, D. W. C. Strategies to Bypass the Taxol Problem. *Enantioselective Cascade Catalysis, a New Approach for the Efficient Construction of Molecular Complexity. Synlett* 18, 1477-1489 (2007).
59. Chau, M., Jennewein, S., Walker, K. & Croteau, R. Taxol biosynthesis: Molecular cloning and characterization of a cytochrome P450 taxoid 7 beta-hydroxylase. *Chem Biol* 11, 663-72 (2004).
60. Williams, D. C. et al. Heterologous expression and characterization of a "Pseudomature" form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Arch Biochem Biophys* 379, 137-46 (2000).
61. Morrone D. et al. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl Microbiol Biotechnol.* 85(6):1893-906 (2010)
- 15 Habiendo descrito, por lo tanto, varios aspectos de al menos una realización de esta invención, se ha de apreciar que diversas alteraciones, modificaciones y mejoras se les ocurrirán fácilmente a los expertos en la técnica. Por consiguiente, la descripción anterior y los dibujos son sólo a modo de ejemplo. Los expertos en la técnica reconocerán, o serán capaces de determinar usando no más que una experimentación rutinaria, muchos equivalentes a las realizaciones específicas de la invención descritas en esta memoria. Tales equivalentes están
- 20 destinados a quedar abarcados por las siguientes reivindicaciones.

LISTADO DE SECUENCIAS

<110> Massachusetts Institute of Technology

<110> National University of Singapore

25 <120> INGENIERÍA MICROBIANA PARA LA PREPARACIÓN DE PRODUCTOS QUÍMICOS Y FARMACÉUTICOS A PARTIR DE LA VÍA ISOPRENOIDE

<130> P203436EP1

<140>

< 141> 10-11-2011

30 <150> US 61/280,877

< 151> 10-11-2009

<150> US 61/388,543

< 151> 30-09-2010

<160> 52

<170> PatentIn version 3.5

35 <210> 1

< 211> 33

< 212> ADN

< 213> Secuencia Artificial

	<220> < 223> Oligonucleótido Sintético	
	<400> 1 cggcatatga gtttgat tgccaaatac ccg	33
5	<210> 2 < 211> 33 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 2 cggctagctt atgccagcca ggccttgatt ttg	33
15	<210> 3 < 211> 52 < 212> ADN < 213> Secuencia Artificial	
20	<220> < 223> Oligonucleótido Sintético	
	<400> 3 cgcggctagc gaaggagata tacatatgca aacggaacac gtcattttat tg	52
25	<210> 4 < 211> 32 < 212> ADN < 213> Secuencia Artificial	
25	<220> < 223> Oligonucleótido Sintético	
	<400> 4 gcggaattcgc tcacaacccc ggcaaatgtc gg	32
30	<210> 5 < 211> 49 < 212> ADN < 213> Secuencia Artificial	
35	<220> < 223> Oligonucleótido Sintético	
	<400> 5 gcgaattcga aggagatata catatggcaa ccaactcattt ggatgtttg	49
40	<210> 6 < 211> 37 < 212> ADN < 213> Secuencia Artificial	

	<220> < 223> Oligonucleótido Sintético	
	<400> 6 gcgctcgagt catttggtg ccttaatgag tagcgcc	37
5	<210> 7 < 211> 33 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 7 taaaccatgg gtttgatg tgccaaatac ccg	33
15	<210> 8 < 211> 36 < 212> ADN < 213> Secuencia Artificial	
20	<220> < 223> Oligonucleótido Sintético	
	<400> 8 cggggtacct catttggtg ccttaatgag tagcgc	36
25	<210> 9 < 211> 36 < 212> ADN < 213> Secuencia Artificial	
25	<220> < 223> Oligonucleótido Sintético	
	<400> 9 cggctcgagt catttggtg ccttaatgag tagcgc	36
30	<210> 10 < 211> 37 < 212> ADN < 213> Secuencia Artificial	
35	<220> < 223> Oligonucleótido Sintético	
	<400> 10 cgtaaccggg gcctctgcta accatgtca tgcctc	37
40	<210> 11 < 211> 24 < 212> ADN < 213> Secuencia Artificial	

	<220> < 223> Oligonucleótido Sintético	
	<400> 11 ctccttcgct agcttatgcc agcc	24
5	<210> 12 < 211> 39 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 12 cgtagaattc actcacaact gacgaaacgc aatgtaac	39
15	<210> 13 < 211> 47 < 212> ADN < 213> Secuencia Artificial	
20	<220> < 223> Oligonucleótido Sintético	
	<400> 3 cgtagaattc agaaggagat atacatatgg ctagtcttac gggtagc	47
25	<210> 14 < 211> 37 < 212> ADN < 213> Secuencia Artificial	
	<220> < 223> Oligonucleótido Sintético	
	<400> 14 gatggctgac ttagacctgg attggatcga tgtaaac	37
30	<210> 15 < 211> 27 < 212> ADN < 213> Secuencia Artificial	
35	<220> < 223> Oligonucleótido Sintético	
	<400> 15 cgtaccatgg ctagtcttac gggtagc	27
40	<210> 16 < 211> 37 < 212> ADN < 213> Secuencia Artificial	

ES 2 575 790 T3

	<220> < 223> Oligonucleótido Sintético	
	<400> 16 cgtagaattc ttagacctgg attggatcga tgtaaac	37
5	<210> 17 < 211> 61 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 17 cgtagaattc agaaggagat atacatatgt ttgatttcaa tgaatatatg aaaagtaagg	60
	c	61
15	<210> 18 < 211> 37 < 212> ADN < 213> Secuencia Artificial	
20	<220> < 223> Oligonucleótido Sintético	
	<400> 18 gatggtcgac tcacaactga cgaaacgcaa tgtaac	37
	<210> 19 < 211> 37 < 212> ADN < 213> Secuencia Artificial	
25	<220> < 223> Oligonucleótido Sintético	
	<400> 19 gatgctcgag ttagacctgg attggatcga tgtaaac	37
30	<210> 20 < 211> 25 < 212> ADN < 213> Secuencia Artificial	
	<220> < 223> Oligonucleótido Sintético	
35	<400> 20 gccgtcgacc atcatcatca tcatc	25
	<210> 21 < 211> 35 < 212> ADN	

	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 21	
5	gcagtctaga gccagaaccg ttatgatgtc ggcgc	35
	<210> 22	
	< 211> 35	
	< 212> ADN	
	< 213> Secuencia Artificial	
10	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 22	
	cggtgccgga gcatctaacg ctgaggtta ggcgc	35
	<210> 23	
15	< 211> 32	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
20	<400> 23	
	gcagtctaga ggaaacctgt cgtgccagct gc	32
	<210> 24	
	< 211> 91	
	< 212> ADN	
25	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 24	
	gacgctcgag gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg	60
	ttttttgctt gtgtaggctg gagctgcttc g	91
30	<210> 25	
	< 211> 34	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
35	< 223> Oligonucleótido Sintético	
	<400> 25	
	gacgagtact gaacgtcgga attgatccgt cgac	34
	<210> 26	

ES 2 575 790 T3

< 211> 91
 < 212> ADN
 < 213> Secuencia Artificial

5 <220>
 < 223> Oligonucleótido Sintético

<400> 26

gacggagctc gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg 60

ttttttgctt gtgtaggctg gagctgcttc g 91

10 <210> 27
 < 211> 63
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 223> Oligonucleótido Sintético

<400> 27

atgacgattt ttgataatta tgaagtgtgg tttgtcattg cattaattgc gttgcgctca 60

15 **ctg 63**

<210> 28
 < 211> 64
 < 212> ADN
 < 213> Secuencia Artificial

20 <220>
 < 223> Oligonucleótido Sintético

<400> 28

atgacgattt ttgataatta tgaagtgtgg tttgtcattg gcatccgctt acagacaagc 60

tgtg 64

25 <210> 29
 < 211> 64
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 223> Oligonucleótido Sintético

30 <400> 29

ttagcgacga aacccgtaat acacttcggt ccagcgcagc cgacgtcgga attgatccgt 60

cgac 64

<210> 30
 < 211> 59

	< 212> ADN < 213> Secuencia Artificial	
	<220> < 223> Oligonucleótido Sintético	
5	<400> 30 cgtacatatg gctctggtat tagcagtttt tgggcaaaa ttaacgaag taaccagc	59
	<210> 31 < 211> 56 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 31 cgtacatatg gctctggtat tagcagtttt ttttagcatc gcttgagtg caattg	56
15	<210> 32 < 211> 58 < 212> ADN < 213> Secuencia Artificial	
	<220> < 223> Oligonucleótido Sintético	
20	<400> 32 cgtacatatg gctctggtat tagcagtttt tttcgctcg aaacgcata gtagcctg	58
	<210> 33 < 211> 49 < 212> ADN < 213> Secuencia Artificial	
25	<220> < 223> Oligonucleótido Sintético	
	<400> 33 cgcgggatcc ggtgctgccc ggacgagga acagttgat tgaaaacc	49
	<210> 34 < 211> 34 < 212> ADN < 213> Secuencia Artificial	
35	<220> < 223> Oligonucleótido Sintético	
	<400> 34 cgcgggatcc cgccgtggtg gaagtgatac acag	34
40	<210> 35 < 211> 40 < 212> ADN	

	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 35	
5	cgcggtcgac ttaccaaata tccgtaagt agcgtccatc	40
	<210> 36	
	< 211> 20	
	< 212> ADN	
	< 213> Secuencia Artificial	
10	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 36	
	attcaaaagc ttccggtcct	20
	<210> 37	
15	< 211> 20	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
20	<400> 37	
	atctggcgac attcgtttc	20
	<210> 38	
	< 211> 20	
	< 212> ADN	
25	< 213> Secuencia Artificial	
	<220>	
	< 223> Oligonucleótido Sintético	
	<400> 38	
	gacgaactgt caccgattt	20
30	<210> 39	
	< 211> 20	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
35	< 223> Oligonucleótido Sintético	
	<400> 39	
	gcttcgctgg tagtagacag	20
	<210> 40	
	< 211> 20	
40	< 212> ADN	
	< 213> Secuencia Artificial	

	<220> < 223> Oligonucleótido Sintético	
	<400> 40 aggccttcgg gttgtaaagt	20
5	<210> 41 < 211> 20 < 212> ADN < 213> Secuencia Artificial	
10	<220> < 223> Oligonucleótido Sintético	
	<400> 41 attccgatta acgcttgac	20
15	<210> 42 < 211> 296 < 212> PRT < 213> Secuencia Artificial	
	<220> < 223> Polipéptido Sintético	
	<400> 42	

ES 2 575 790 T3

Met Phe Asp Phe Asn Glu Tyr Met Lys Ser Lys Ala Val Ala Val Asp
 1 5 10 15

Ala Ala Leu Asp Lys Ala Ile Pro Leu Glu Tyr Pro Glu Lys Ile His
 20 25 30

Glu Ser Met Arg Tyr Ser Leu Leu Ala Gly Gly Lys Arg Val Arg Pro
 35 40 45

Ala Leu Cys Ile Ala Ala Cys Glu Leu Val Gly Gly Ser Gln Asp Leu
 50 55 60

Ala Met Pro Thr Ala Cys Ala Met Glu Met Ile His Thr Met Ser Leu
 65 70 75 80

Ile His Asp Asp Leu Pro Cys Met Asp Asn Asp Asp Phe Arg Arg Gly
 85 90 95

Lys Pro Thr Asn His Lys Val Phe Gly Glu Asp Thr Ala Val Leu Ala
 100 105 110

Gly Asp Ala Leu Leu Ser Phe Ala Phe Glu His Ile Ala Val Ala Thr
 115 120 125

Ser Lys Thr Val Pro Ser Asp Arg Thr Leu Arg Val Ile Ser Glu Leu
 130 135 140

Gly Lys Thr Ile Gly Ser Gln Gly Leu Val Gly Gly Gln Val Val Asp
 145 150 155 160

Ile Thr Ser Glu Gly Asp Ala Asn Val Asp Leu Lys Thr Leu Glu Trp
 165 170 175

Ile His Ile His Lys Thr Ala Val Leu Leu Glu Cys Ser Val Val Ser
 180 185 190

ES 2 575 790 T3

Gly Gly Ile Leu Gly Gly Ala Thr Glu Asp Glu Ile Ala Arg Ile Arg
195 200 205

Arg Tyr Ala Arg Cys Val Gly Leu Leu Phe Gln Val Val Asp Asp Ile
210 215 220

Leu Asp Val Thr Lys Ser Ser Glu Glu Leu Gly Lys Thr Ala Gly Lys
225 230 235 240

Asp Leu Leu Thr Asp Lys Ala Thr Tyr Pro Lys Leu Met Gly Leu Glu
245 250 255

Lys Ala Lys Glu Phe Ala Ala Glu Leu Ala Thr Arg Ala Lys Glu Glu
260 265 270

Leu Ser Ser Phe Asp Gln Ile Lys Ala Ala Pro Leu Leu Gly Leu Ala
275 280 285

Asp Tyr Ile Ala Phe Arg Gln Asn
290 295

<210> 43

< 211> 888

< 212> PRT

5 < 213> Secuencia Artificial

<220>

< 223> Polinucleótido Sintético

<400> 43

ES 2 575 790 T3

```

atgtttgatt tcaatgaata tatgaaaagt aaggctgttg cggtagacgc ggctctggat      60
aaagcgattc cgctggaata tcccgagaag attcacgaat cgatgcgcta ctccctgtta      120
gcaggagggg aacgcgttcg tccggcatta tgcacgcggg cctgtgaact cgtcggcggg      180
tcacaggact tagcaatgcc aactgcttgc gcaatggaaa tgattcacac aatgagcctg      240
attcatgatg atttgcttg catggacaac gatgactttc ggcgcggtaa acctactaat      300
cataaggttt ttggcgaaga tactgcagtg ctggcggggc atgcgctgct gtcgtttgcc      360
ttcgaacata tcgccgtcgc gacctcgaag accgtcccgt cggaccgtac gcttcgctg      420
atcccgagc tgggaaagac catcggtctc caaggactcg tgggtgggca ggtagttgat      480
atcacgtctg aggtgacgc gaacgtggac ctgaaaaccc tggagtggat ccatattcac      540
aaaacggccg tgctgctgga atgtagcgtg gtgtcagggg ggatcttggg ggcgcccacg      600
gaggatgaaa tcgcgcgtat tcgtcgttat gcccgctgtg ttggactggt atttcaggtg      660
gtggatgaca tcctggatgt cacaaaatcc agcgaagagc ttggcaagac cgcgggcaaa      720
gaccttctga cggataaggc tacatacccg aaattgatgg gcttggagaa agccaaggag      780
ttcgcagctg aacttgccac gcgggcgaag gaagaactct cttctttcga tcaaatcaaa      840
gccgcgccac tgctgggctc cgccgattac attgcgtttc gtcagaac      888

```

5 <210> 44
 <211> 802
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Polipéptido Sintético

10 <400> 44

ES 2 575 790 T3

Met Ser Ser Ser Thr Gly Thr Ser Lys Val Val Ser Glu Thr Ser Ser
 1 5 10 15

Thr Ile Val Asp Asp Ile Pro Arg Leu Ser Ala Asn Tyr His Gly Asp
 20 25 30

Leu Trp His His Asn Val Ile Gln Thr Leu Glu Thr Pro Phe Arg Glu
 35 40 45

Ser Ser Thr Tyr Gln Glu Arg Ala Asp Glu Leu Val Val Lys Ile Lys
 50 55 60

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

Asp Thr Ala Trp Val Ala Arg Leu Ala Thr Ile Ser Ser Asp Gly Ser
 85 90 95

Glu Lys Pro Arg Phe Pro Gln Ala Leu Asn Trp Val Phe Asn Asn Gln
 100 105 110

Leu Gln Asp Gly Ser Trp Gly Ile Glu Ser His Phe Ser Leu Cys Asp
 115 120 125

Arg Leu Leu Asn Thr Thr Asn Ser Val Ile Ala Leu Ser Val Trp Lys
 130 135 140

Thr Gly His Ser Gln Val Gln Gln Gly Ala Glu Phe Ile Ala Glu Asn
 145 150 155 160

Leu Arg Leu Leu Asn Glu Glu Asp Glu Leu Ser Pro Asp Phe Gln Ile
 165 170 175

Ile Phe Pro Ala Leu Leu Gln Lys Ala Lys Ala Leu Gly Ile Asn Leu
 180 185 190

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

ES 2 575 790 T3

Arg Leu Thr Asp Val Ser Ala Ala Ala Asp Asn Ile Pro Ala Asn Met
 210 215 220

Leu Asn Ala Leu Glu Gly Leu Glu Glu Val Ile Asp Trp Asn Lys Ile
 225 230 235 240

Met Arg Phe Gln Ser Lys Asp Gly Ser Phe Leu Ser Ser Pro Ala Ser
 245 250 255

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

Leu Asn Asn Leu Leu Asp Lys Phe Gly Gly Cys Val Pro Cys Met Tyr
 275 280 285

Ser Ile Asp Leu Leu Glu Arg Leu Ser Leu Val Asp Asn Ile Glu His
 290 295 300

Leu Gly Ile Gly Arg His Phe Lys Gln Glu Ile Lys Gly Ala Leu Asp
 305 310 315 320

Tyr Val Tyr Arg His Trp Ser Glu Arg Gly Ile Gly Trp Gly Arg Asp
 325 330 335

Ser Leu Val Pro Asp Leu Asn Thr Thr Ala Leu Gly Leu Arg Thr Leu
 340 345 350

Arg Met His Gly Tyr Asn Val Ser Ser Asp Val Leu Asn Asn Phe Lys
 355 360 365

Asp Glu Asn Gly Arg Phe Phe Ser Ser Ala Gly Gln Thr His Val Glu
 370 375 380

Leu Arg Ser Val Val Asn Leu Phe Arg Ala Ser Asp Leu Ala Phe Pro
 385 390 395 400

Asp Glu Arg Ala Met Asp Asp Ala Arg Lys Phe Ala Glu Pro Tyr Leu
 405 410 415

Arg Glu Ala Leu Ala Thr Lys Ile Ser Thr Asn Thr Lys Leu Phe Lys
 420 425 430

Glu Ile Glu Tyr Val Val Glu Tyr Pro Trp His Met Ser Ile Pro Arg
 435 440 445

Leu Glu Ala Arg Ser Tyr Ile Asp Ser Tyr Asp Asp Asn Tyr Val Trp
 450 455 460

ES 2 575 790 T3

Gln Arg Lys Thr Leu Tyr Arg Met Pro Ser Leu Ser Asn Ser Lys Cys
465 470 475 480

Leu Glu Leu Ala Lys Leu Asp Phe Asn Ile Val Gln Ser Leu His Gln
485 490 495

Glu Glu Leu Lys Leu Leu Thr Arg Trp Trp Lys Glu Ser Gly Met Ala
500 505 510

Asp Ile Asn Phe Thr Arg His Arg Val Ala Glu Val Tyr Phe Ser Ser
515 520 525

Ala Thr Phe Glu Pro Glu Tyr Ser Ala Thr Arg Ile Ala Phe Thr Lys
530 535 540

Ile Gly Cys Leu Gln Val Leu Phe Asp Asp Met Ala Asp Ile Phe Ala
545 550 555 560

Thr Leu Asp Glu Leu Lys Ser Phe Thr Glu Gly Val Lys Arg Trp Asp
565 570 575

Thr Ser Leu Leu His Glu Ile Pro Glu Cys Met Gln Thr Cys Phe Lys
580 585 590

Val Trp Phe Lys Leu Met Glu Glu Val Asn Asn Asp Val Val Lys Val
595 600 605

Gln Gly Arg Asp Met Leu Ala His Ile Arg Lys Pro Trp Glu Leu Tyr
610 615 620

Phe Asn Cys Tyr Val Gln Glu Arg Glu Trp Leu Glu Ala Gly Tyr Ile
625 630 635 640

Pro Thr Phe Glu Glu Tyr Leu Lys Thr Tyr Ala Ile Ser Val Gly Leu
645 650 655

Gly Pro Cys Thr Leu Gln Pro Ile Leu Leu Met Gly Glu Leu Val Lys
660 665 670

Asp Asp Val Val Glu Lys Val His Tyr Pro Ser Asn Met Phe Glu Leu
675 680 685

Val Ser Leu Ser Trp Arg Leu Thr Asn Asp Thr Lys Thr Tyr Gln Ala
690 695 700

Glu Lys Ala Arg Gly Gln Gln Ala Ser Gly Ile Ala Cys Tyr Met Lys
705 710 715 720

Asp Asn Pro Gly Ala Thr Glu Glu Asp Ala Ile Lys His Ile Cys Arg

ES 2 575 790 T3

	725			730			735
Val Val Asp Arg Ala Leu Lys Glu Ala Ser Phe Glu Tyr Phe Lys Pro							
	740			745			750
Ser Asn Asp Ile Pro Met Gly Cys Lys Ser Phe Ile Phe Asn Leu Arg							
	755			760			765
Leu Cys Val Gln Ile Phe Tyr Lys Phe Ile Asp Gly Tyr Gly Ile Ala							
	770			775			780
Asn Glu Glu Ile Lys Asp Tyr Ile Arg Lys Val Tyr Ile Asp Pro Ile							
	785			790			800

Gln Val

<210> 45
 < 211> 2406
 < 212> ADN
 5 < 213> Secuencia Artificial

<220>
 < 223> Polinucleótido Sintético

<400> 45

```

atgtctagct ctacgggtac gtctaaagtc gtgagtgaaa cctcatcgac gatcgtggac      60
gatattccac gcttgtcggc gaactatcat ggagatctgt ggcatcataa cgtcattcag      120
acattggaaa ccccgtttcg cgaaagtagc acctaccagg aacgggcaga tgaattagtc      180
gtgaaaatca aagatatgtt taatgcatta ggagatggag acatctcgcc cagcgcatat      240
gatacggcgt gggtagctcg gttggccacg attagctccg atggcagtga aaagccgcgt      300
ttcccgcagg cgctgaactg ggtgtttaat aatcaattgc aggatggcag ctggggcatt      360
gaatctcact ttagcctctg tgaccgggta ctcaacacga caaactccgt aatgvcgttg      420
tcagtttgga aaacgggcca tagccaggtt caacagggcg cggaatttat cgctgaaaat      480
ctgcgcctgc tgaacgagga ggaogaactg tcacccgatt ttcagattat ttttccggct      540
ttactccaga aagccaaagc cttaggcatc aacctgccat atgatctgcc gttcatcaag      600
tatctgtcta ctaccgcga agcccgtctc actgacgtct ctgvcggcggc ggacaatatt      660
ccagcgaaca tgctgaacgc actggaaggg ctggaagagg ttatcgactg gaataaaatc      720
atgvccttcc aaagcaagga cggtagcttc ttaagcagcc cagcatctac tgcttgtggt      780
ctgatgaata ccggagacga aaagtgcttt acgtttctga acaatctgct ggacaaattt      840
gggggttggt ttccttgat gtattccatt gatctgttg aacgtctgtc gctggtcgat      900
aacattgaac acttaggtat cggccgccac ttcaacaag aatcaaggg gcggttgat      960
tatgtatacc gtcattggag cgagcgtggt attggttggg ggcgcgatag cttggtacct     1020

```

ES 2 575 790 T3

gatctgaaca ccactgcttt gggactgcmc actcttcgta tgcacggata caacgttagt 1080
tccgatgtcc tcaataatth caaggacgag aacggccgtt ttttcagctc ggccggtcag 1140
acgcatgttg aactgcmcgc cgtagtcaat ctctttcgcg ctagtgatct ggccttcccc 1200
gacgagcmcgc ctatggacga tgcacggaag tttgcmcgcg cgtatctccg cgaagccctg 1260
gccaccaaaa tttcaaccaa caccaagctt ttcaaagaaa ttgagtatgt agtagagtat 1320
ccgtggcata tgtctattcc gcgcctggaa gcccgctcgt atatcgattc ttacgatgac 1380
aattatgtgt ggcaacgcaa aacctgtac cgtatgccca gcctgtcaa tagtaagtgt 1440
ctggagctgg cgaactgga tttcaacatt gtgcaatccc tgcaccaaga agagctgaaa 1500
ttactgactc gctggtggaa ggaatccggc atggcagaca tcaattttac gcgtcacctg 1560
gttgacgagg tgtacttctc ctccgcmgacc tttgagccgg agtattcggc cacacgtatt 1620
gcatttacca agattggctg ccttcaggtg ctttttgacg atatggcmgga tatttttgcmg 1680
acacttgatg agcttaaatc atttaccgaa ggcgtgaagc gttgggatac ctctctgttg 1740
catgaaatcc ccgaatgtat gcagacctgc ttcaaagttt ggttcaaact gatggaagaa 1800
gtgaacaacg acgtcgtgaa agttcagggc cgtgatatgt tagcacacat ccgcaagccg 1860
tgggaactct atttcaattg ctatgtcmgag gagcgtgaaat ggttagaagc gggctacatt 1920
cctacctcg aagagtactt aaaaacctat gccatttccg tcmggttagg cccgtgcact 1980
ctgcagccta tcttgctgat gggtagctg gtaaaggatg atgtggtgga aaaagttcac 2040
taccctcga atatgtttga actggtaagt ctgagttgcmg gtctgacaaa cgacaccaaaa 2100
acgtaccagg cagaaaagcmg acgtgggcaa caggcaagcmg gtatcgcctg ttatatgaag 2160
gataatccgg cmgctactga ggaagatgcc attaagcata tctgccctgt tgtggatcmg 2220
gctcttaaag aagcmgtcatt cgaatatttt aaacctagta atgatattcc gatgggttgt 2280
aagtcattca ttttcaatct tcmgctgtgc gtgcaaatth tttacaaatt tattgaccmg 2340
tacggaatcmg ccaacgaaga aatcaaagac tatattccta aagtttcat cmgtccaatc 2400
caggtc 2406

5 <210> 46
< 211> 499
< 212> PRT
< 213> Secuencia Artificial

<220>
< 223> Polipéptido Sintético

10 <400> 46

ES 2 575 790 T3

Met Asp Ala Leu Tyr Lys Ser Thr Val Ala Lys Phe Asn Glu Val Thr
 1 5 10 15

Gln Leu Asp Cys Ser Thr Glu Ser Phe Ser Ile Ala Leu Ser Ala Ile
 20 25 30

Ala Gly Ile Leu Leu Leu Leu Leu Leu Phe Arg Ser Lys Arg His Ser
 35 40 45

Ser Leu Lys Leu Pro Pro Gly Lys Leu Gly Ile Pro Phe Ile Gly Glu
 50 55 60

Ser Phe Ile Phe Leu Arg Ala Leu Arg Ser Asn Ser Leu Glu Gln Phe
 65 70 75 80

Phe Asp Glu Arg Val Lys Lys Phe Gly Leu Val Phe Lys Thr Ser Leu
 85 90 95

Ile Gly His Pro Thr Val Val Leu Cys Gly Pro Ala Gly Asn Arg Leu
 100 105 110

Ile Leu Ser Asn Glu Glu Lys Leu Val Gln Met Ser Trp Pro Ala Gln
 115 120 125

Phe Met Lys Leu Met Gly Glu Asn Ser Val Ala Thr Arg Arg Gly Glu
 130 135 140

Asp His Ile Val Met Arg Ser Ala Leu Ala Gly Phe Phe Gly Pro Gly
 145 150 155 160

Ala Leu Gln Ser Tyr Ile Gly Lys Met Asn Thr Glu Ile Gln Ser His
 165 170 175

Ile Asn Glu Lys Trp Lys Gly Lys Asp Glu Val Asn Val Leu Pro Leu
 180 185 190

Val Arg Glu Leu Val Phe Asn Ile Ser Ala Ile Leu Phe Phe Asn Ile
 195 200 205

Tyr Asp Lys Gln Glu Gln Asp Arg Leu His Lys Leu Leu Glu Thr Ile
 210 215 220

Leu Val Gly Ser Phe Ala Leu Pro Ile Asp Leu Pro Gly Phe Gly Phe
 225 230 235 240

His Arg Ala Leu Gln Gly Arg Ala Lys Leu Asn Lys Ile Met Leu Ser
 245 250 255

Leu Ile Lys Lys Arg Lys Glu Asp Leu Gln Ser Gly Ser Ala Thr Ala
 260 265 270

Thr Gln Asp Leu Leu Ser Val Leu Leu Thr Phe Arg Asp Asp Lys Gly
 275 280 285

ES 2 575 790 T3

Thr Pro Leu Thr Asn Asp Glu Ile Leu Asp Asn Phe Ser Ser Leu Leu
290 295 300

His Ala Ser Tyr Asp Thr Thr Thr Ser Pro Met Ala Leu Ile Phe Lys
305 310 315 320

Leu Leu Ser Ser Asn Pro Glu Cys Tyr Gln Lys Val Val Gln Glu Gln
325 330 335

Leu Glu Ile Leu Ser Asn Lys Glu Glu Gly Glu Glu Ile Thr Trp Lys
340 345 350

Asp Leu Lys Ala Met Lys Tyr Thr Trp Gln Val Ala Gln Glu Thr Leu
355 360 365

Arg Met Phe Pro Pro Val Phe Gly Thr Phe Arg Lys Ala Ile Thr Asp
370 375 380

Ile Gln Tyr Asp Gly Tyr Thr Ile Pro Lys Gly Trp Lys Leu Leu Trp
385 390 395 400

Thr Thr Tyr Ser Thr His Pro Lys Asp Leu Tyr Phe Asn Glu Pro Glu
405 410 415

Lys Phe Met Pro Ser Arg Phe Asp Gln Glu Gly Lys His Val Ala Pro
420 425 430

Tyr Thr Phe Leu Pro Phe Gly Gly Gly Gln Arg Ser Cys Val Gly Trp
435 440 445

Glu Phe Ser Lys Met Glu Ile Leu Leu Phe Val His His Phe Val Lys
450 455 460

Thr Phe Ser Ser Tyr Thr Pro Val Asp Pro Asp Glu Lys Ile Ser Gly
465 470 475 480

Asp Pro Leu Pro Pro Leu Pro Ser Lys Gly Phe Ser Ile Lys Leu Phe
485 490 495

Pro Arg Pro

<210> 47

< 211> 1497

< 212> ADN

< 213> Secuencia Artificial

5

<220>

< 223> Polinucleótido Sintético

<400> 47

ES 2 575 790 T3

```

atggatgcc tctataagtc taccgtggcg aaatttaacg aagtaaccga gctggattgc      60
agcactgagt catttagcat cgctttgagt gcaattgccg ggatcttgct gttgctcctg      120
ctgtttcgct cgaaacgtca tagtagcctg aaattacctc cgggcaaaact gggcattccg      180
tttatcggtg agtcctttat ttttttgccg gcgctgcgca gcaattctct ggaacagttc      240
tttgatgaac gtgtgaagaa gttcggcctg gtatttaaaa cgtcccttat cggtcacccg      300
acggttgctc tgtgcccggc cgcaggtaat cgcctcatcc tgagcaacga agaaaagctg      360
gtacagatgt cctggccggc gcagtttatg aagctgatgg gagagaactc agttgcgacc      420
cgccgtggtg aagatcacat tgttatgcgc tccgcgttgg caggcttttt cggcccggga      480
gctctgcaat cctatatcgg caagatgaac acggaaatcc aaagccatat taatgaaaag      540
tggaaagggg aggacgaggt taatgtctta cccctggtgc gggaaactgt ttttaacatc      600
agcgtatttc tgttctttta catttacgat aagcaggaac aagaccgtct gcacaagttg      660
ttagaaacca ttctggtagg ctogtttgc ttaccaattg atttaccggg tttcgggttt      720
caccgcgctt tacaaggtcg tgcaaaaactc aataaaatca tgttgcgct tattaaaaaa      780
cgtaaagagg acttacagtc gggatcggcc accgcgacgc aggacctgtt gtctgtgctt      840
ctgactttcc gtgatgataa gggcaccocg ttaaccaatg acgaaatcct ggacaacttt      900
agctcactgc ttcacgctc ttacgacacc acgactagtc caatggctct gattttcaaa      960
ttactgtcaa gtaaccctga atgctatcag aaagtcgtgc aagagcaact cgagattctg     1020
agcaataagg aagaaggtga agaaattacc tggaaagatc ttaaggccat gaaatacacg     1080
tggcaggttg cgcaggagac acttcgcatg ttccaccgg tgttcgggac cttccgcaaa     1140
gcgatcacgg atattcagta tgacggatac acaatcccga aaggttgaa actggttggtg     1200
actacctata gcactcatcc taaggacctt tacttcaacg aaccggagaa atttatgcct     1260
agtcgtttcg atcaggaagg caaacatggt gcgccctata ccttcctgcc ctttgagggc     1320
ggtcagcggg gttgtgtggg ttgggagttc tctaagatgg agattctcct cttcgtgcat     1380
catttcgtga aaacattttc gagctatacc ccggtcgatc ccgatgaaaa aatttcgggc     1440
gatccactgc cgccgttacc gagcaaaggg ttttcaatca aactgttccc tegtccg      1497

```

<210> 48
 <211> 717
 5 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Polipéptido Sintético

<400> 48

10 Met Gln Ala Asn Ser Asn Thr Val Glu Gly Ala Ser Gln Gly Lys Ser
 1 5 10 15

ES 2 575 790 T3

Leu Leu Asp Ile Ser Arg Leu Asp His Ile Phe Ala Leu Leu Leu Asn
 20 25 30
 Gly Lys Gly Gly Asp Leu Gly Ala Met Thr Gly Ser Ala Leu Ile Leu
 35 40 45
 Thr Glu Asn Ser Gln Asn Leu Met Ile Leu Thr Thr Ala Leu Ala Val
 50 55 60
 Leu Val Ala Cys Val Phe Phe Phe Val Trp Arg Arg Gly Gly Ser Asp
 65 70 75 80
 Thr Gln Lys Pro Ala Val Arg Pro Thr Pro Leu Val Lys Glu Glu Asp
 85 90 95
 Glu Glu Glu Glu Asp Asp Ser Ala Lys Lys Lys Val Thr Ile Phe Phe
 100 105 110
 Gly Thr Gln Thr Gly Thr Ala Glu Gly Phe Ala Lys Ala Leu Ala Glu
 115 120 125
 Glu Ala Lys Ala Arg Tyr Glu Lys Ala Val Phe Lys Val Val Asp Leu
 130 135 140
 Asp Asn Tyr Ala Ala Asp Asp Glu Gln Tyr Glu Glu Lys Leu Lys Lys
 145 150 155 160
 Glu Lys Leu Ala Phe Phe Met Leu Ala Thr Tyr Gly Asp Gly Glu Pro
 165 170 175
 Thr Asp Asn Ala Ala Arg Phe Tyr Lys Trp Phe Leu Glu Gly Lys Glu
 180 185 190
 Arg Glu Pro Trp Leu Ser Asp Leu Thr Tyr Gly Val Phe Gly Leu Gly
 195 200 205
 Asn Arg Gln Tyr Glu His Phe Asn Lys Val Ala Lys Ala Val Asp Glu
 210 215 220
 Val Leu Ile Glu Gln Gly Ala Lys Arg Leu Val Pro Val Gly Leu Gly
 225 230 235 240
 Asp Asp Asp Gln Cys Ile Glu Asp Asp Phe Thr Ala Trp Arg Glu Gln
 245 250 255
 Val Trp Pro Glu Leu Asp Gln Leu Leu Arg Asp Glu Asp Asp Glu Pro
 260 265 270

ES 2 575 790 T3

Thr Ser Ala Thr Pro Tyr Thr Ala Ala Ile Pro Glu Tyr Arg Val Glu
 275 280 285
 Ile Tyr Asp Ser Val Val Ser Val Tyr Glu Glu Thr His Ala Leu Lys
 290 295 300
 Gln Asn Gly Gln Ala Val Tyr Asp Ile His His Pro Cys Arg Ser Asn
 305 310 315 320
 Val Ala Val Arg Arg Glu Leu His Thr Pro Leu Ser Asp Arg Ser Cys
 325 330 335
 Ile His Leu Glu Phe Asp Ile Ser Asp Thr Gly Leu Ile Tyr Glu Thr
 340 345 350
 Gly Asp His Val Gly Val His Thr Glu Asn Ser Ile Glu Thr Val Glu
 355 360 365
 Glu Ala Ala Lys Leu Leu Gly Tyr Gln Leu Asp Thr Ile Phe Ser Val
 370 375 380
 His Gly Asp Lys Glu Asp Gly Thr Pro Leu Gly Gly Ser Ser Leu Pro
 385 390 395 400
 Pro Pro Phe Pro Gly Pro Cys Thr Leu Arg Thr Ala Leu Ala Arg Tyr
 405 410 415
 Ala Asp Leu Leu Asn Pro Pro Arg Lys Ala Ala Phe Leu Ala Leu Ala
 420 425 430
 Ala His Ala Ser Asp Pro Ala Glu Ala Glu Arg Leu Lys Phe Leu Ser
 435 440 445
 Ser Pro Ala Gly Lys Asp Glu Tyr Ser Gln Trp Val Thr Ala Ser Gln
 450 455 460
 Arg Ser Leu Leu Glu Ile Met Ala Glu Phe Pro Ser Ala Lys Pro Pro
 465 470 475 480
 Leu Gly Val Phe Phe Ala Ala Ile Ala Pro Arg Leu Gln Pro Arg Tyr
 485 490 495
 Tyr Ser Ile Ser Ser Ser Pro Arg Phe Ala Pro Ser Arg Ile His Val
 500 505 510
 Thr Cys Ala Leu Val Tyr Gly Pro Ser Pro Thr Gly Arg Ile His Lys
 515 520 525

ES 2 575 790 T3

Gly Val Cys Ser Asn Trp Met Lys Asn Ser Leu Pro Ser Glu Glu Thr
 530 535 540

His Asp Cys Ser Trp Ala Pro Val Phe Val Arg Gln Ser Asn Phe Lys
 545 550 555 560

Leu Pro Ala Asp Ser Thr Thr Pro Ile Val Met Val Gly Pro Gly Thr
 565 570 575

Gly Phe Ala Pro Phe Arg Gly Phe Leu Gln Glu Arg Ala Lys Leu Gln
 580 585 590

Glu Ala Gly Glu Lys Leu Gly Pro Ala Val Leu Phe Phe Gly Cys Arg
 595 600 605

Asn Arg Gln Met Asp Tyr Ile Tyr Glu Asp Glu Leu Lys Gly Tyr Val
 610 615 620

Glu Lys Gly Ile Leu Thr Asn Leu Ile Val Ala Phe Ser Arg Glu Gly
 625 630 635 640

Ala Thr Lys Glu Tyr Val Gln His Lys Met Leu Glu Lys Ala Ser Asp
 645 650 655

Thr Trp Ser Leu Ile Ala Gln Gly Gly Tyr Leu Tyr Val Cys Gly Asp
 660 665 670

Ala Lys Gly Met Ala Arg Asp Val His Arg Thr Leu His Thr Ile Val
 675 680 685

Gln Glu Gln Glu Ser Val Asp Ser Ser Lys Ala Glu Phe Leu Val Lys
 690 695 700

Lys Leu Gln Met Asp Gly Arg Tyr Leu Arg Asp Ile Trp
 705 710 715

<210> 49

< 211> 2151

< 212> ADN

5 < 213> Secuencia Artificial

<220>

< 223> Polinucleótido Sintético

<400> 49

ES 2 575 790 T3

```
atgcaggcga attctaatac ggttgaaggc gcgagccaag gcaagtctct tctggacatt      60
agtcgcctcg accatatctt cgccctgctg ttgaacggga aagcgggaga ccttgggtgcg    120
atgaccgggt cggccttaat tctgacggaa aatagccaga acttgatgat tctgaccact    180
gcgctggccg ttctggtcgc ttgcgttttt tttttcgttt ggcgcctggg tggaaagtgat    240
```

ES 2 575 790 T3

acacagaagc cgcgctacg tcccacacct cttgttaaag aagaggacga agaagaagaa 300
gatgatagcg ccaagaaaaa ggtcacaata ttttttggca ccagaccgg caccgcccga 360
ggtttcgcga aggccttagc tgaggaagca aaggcacggt atgaaaaggc ggtattttaa 420
gtcgtggatt tggataacta tgcagcggat gacgaacagt acgaagagaa gttgaaaaag 480
gaaaagctag cgttcttcat gctcgccacc tacggtgacg gcgaaccgac tgataatgcc 540
gctcgctttt ataaatgggt tctcgagggt aaagagcgcg agccatgggt gtcagatctg 600
acttatggcg tgtttggcct aggtaaccgt cagtatgaac actttaaca ggtcgcgaaa 660
gcggtggacg aagtgctcat tgaacaaggc gccaaacgtc tggtagcgggt agggcttgg 720
gatgatgac agtgattga ggacgactc actgcctgga gagaacaagt gtggcctgag 780
ctggatcagc tcttacgtga tgaagatgac gagccgacgt ctgagacccc gtacacggcg 840
gctattccag aataccgggt ggaatctac gactcagtag tgcggtcta tgaggaaacc 900
catgcgctga acaaaaatgg acaagccgta tacgatatcc accaccggtg tcgcagcaac 960
gtggcagtac gtcgtgagct gcataccccg ctgctggatc gtagttgtat tcatctggaa 1020
ttcgatatta gtgatactgg gtaaatctat gagacgggcg accacgttgg agttcatacc 1080
gagaattcaa ttgaaaccgt ggaagaagca gctaaactgt taggttacca actggatata 1140
atcttcagcg tgcattggga caaggaagat ggaacacat tgggcgggag tagcctgcca 1200
ccgcccgttc cggggccctg cacgctgcgg acggcgtggt cacgttacgc ggacctgctg 1260
aaccctccgc gcaaagccgc ctctctggca ctggccgac acgcgtcaga tccggctgaa 1320
gctgaagccc ttaatttct cagttctcca gccgaaaag acgaatactc acagtgggtc 1380
actgcgtccc aacgcagcct cctcgagatt atggccgaat tccccagcg gaaaccgccc 1440
ctgggagtg ttttcgccc aatagcggcg cgcttgcaac ctaggatta tagcatctcc 1500
tcctccccgc gtttcgccc gtctcgtatc catgtaacgt gcgctggt ctatggtcct 1560
agccctacg ggcgtattca taaaggtgtg tgcagcaact ggatgaagaa ttctttgccc 1620
tccgaagaaa cccacgattg cagctgggca ccggtctttg tgcgccagtc aaactttaa 1680
ctgcccggcg attcgacgac gccaatcgtg atggttggac ctggaaccgg ctctgctcca 1740
tttcgaggct tccttcagga acgcgcaaaa ctgcaggaag cgggcgaaaa attgggccc 1800
gcagtgtgt tttttgggtg ccgcaaccgc cagatggatt acatctatga agatgagctt 1860
aagggttacg ttgaaaaagg tattctgacg aatctgatcg ttgcattttc acgagaaggc 1920
gccaccaaag agtatgttca gcacaagatg ttagagaaa cctccgacac gtggtcttta 1980
atcggccagg gtggttatct gtatgtttgc ggtgatgca agggatggc cagagacgta 2040
catcgacccc tgcatacaat cgttcaggaa caagaatccg tagactcgtc aaaagcggag 2100
tttttagtca aaaagctgca aatggatgga cgctacttac gggatatttg g 2151

<210> 50

<211> 5

<212> PRT

<213> Secuencia Artificial

5

<220>
< 223> Polipéptido Sintético

<400> 50

Gly Ser Thr Gly Ser
1 5

5 <210> 51
< 211> 8
< 212> PRT
< 213> Secuencia Artificial

10 <220>
< 223> Polipéptido Sintético

Met Ala Leu Leu Leu Ala Val Phe
1 5

15 <210> 52
< 211> 41
< 212> ADN
< 213> Secuencia Artificial

<220>
< 223> Oligonucleótido Sintético
cgtaccatgg ttgattcaa tgaatatatg aaaagtaagg c

41

REIVINDICACIONES

- 5 1. Un método para aumentar la producción de terpenoides en una célula de *E. coli* que produce uno o más terpenoides, que comprende controlar la acumulación de indol en la célula o en un cultivo de las células a menos de 100 mg/L y sobre-exresar en la célula uno o más componentes de la vía no mevalonato (MEP), aumentando con ello la producción de terpenoides en la célula, en el que la etapa de controlar la acumulación de indol en la célula o en un cultivo de las células comprende (i) equilibrar una vía MEP aguas arriba con una vía de síntesis de terpenoides aguas abajo y/o modificar o regular una vía indol, o (ii) separar del cultivo celular el indol acumulado a través de métodos químicos.
- 10 2. El método de la reivindicación 1, en el que se sobre-expresan el uno o más componentes de la vía MEP para amplificar la producción de isopentenil difosfato (IPP) y dimetilalil difosfato (DMAPP).
3. El método de la reivindicación 1 ó 2, en el que la célula expresa de forma recombinante una enzima terpenoide sintasa y una enzima geranylgeranyl difosfato sintasa (GGPPS).
4. El método de una cualquiera de las reivindicaciones 1-3, en el que el uno o más componentes de la vía MEP se seleccionan de *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* e *ispB*.
- 15 5. El método de la reivindicación 4, en el que la célula comprende un operón *dxs-idi-ispDF* heterólogo.
6. El método de una cualquiera de las reivindicaciones 1-5, en el que la vía MEP de aguas arriba está equilibrada con respecto a las vías de síntesis de terpenoides de aguas abajo en uno o más de:
- manipular el número de copias de genes u operones para una o más enzimas de las vías de aguas arriba y/o aguas abajo,
- 20 regular la expresión de genes u operones de la vía de aguas arriba y/o aguas abajo utilizando promotores con diferentes fuerzas.
7. El método de una cualquiera de las reivindicaciones 1-5, en el que el indol acumulado se separa del cultivo celular utilizando absorbentes o agentes depuradores.
- 25 8. El método de una cualquiera de las reivindicaciones 1-7, en el que la producción se realiza en un recipiente de reacción ventilado, y en el que el método comprende recuperar el terpenoide del cultivo celular.
9. El método de una cualquiera de las reivindicaciones 1-7, en el que el método comprende recuperar el terpenoide de la fase gaseosa del cultivo celular.
10. El método de la reivindicación 9, en el que el método comprende añadir al cultivo celular una capa orgánica tal como dodecano y recuperar el terpenoide de la capa orgánica.
- 30 11. El método de una cualquiera de las reivindicaciones 1-10, en el que el uno o más terpenoides es un monoterpeneoide, un sesquiterpeneoide, un diterpeneoide, un triterpeneoide o un tetraterpeneoide.
12. El método de la reivindicación 11, en el que el uno o más terpenoides es un taxadieno o un precursor de taxol.
13. El método de una cualquiera de las reivindicaciones 1-12, que comprende, además, medir la cantidad o concentración de indol en la célula o en un cultivo de las células.
- 35 14. El método de la reivindicación 13, en el que el método comprende medir dos o más veces la cantidad o concentración de indol.
15. El método de la reivindicación 13 o la reivindicación 14, en el que la cantidad o concentración medida de indol se utiliza para guiar un procedimiento de producir uno o más terpenoides y/o para guiar la construcción de la cepa.

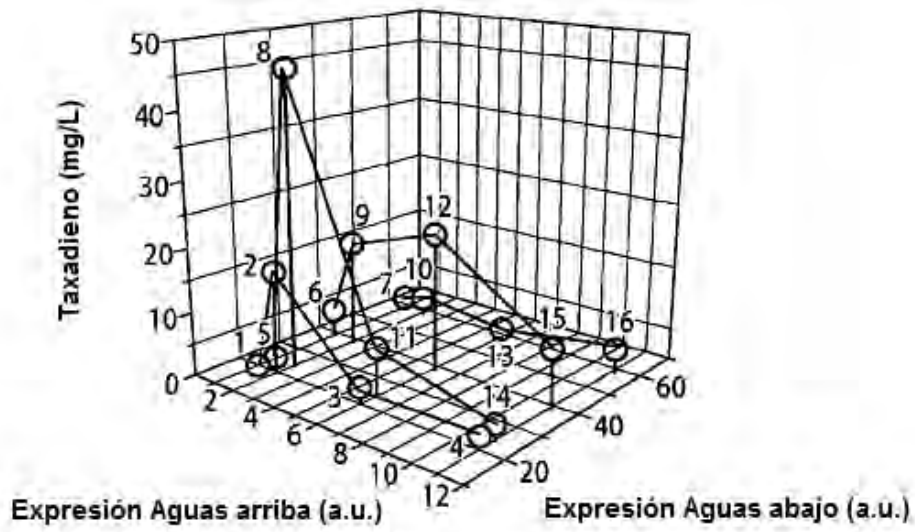


Fig. 2a

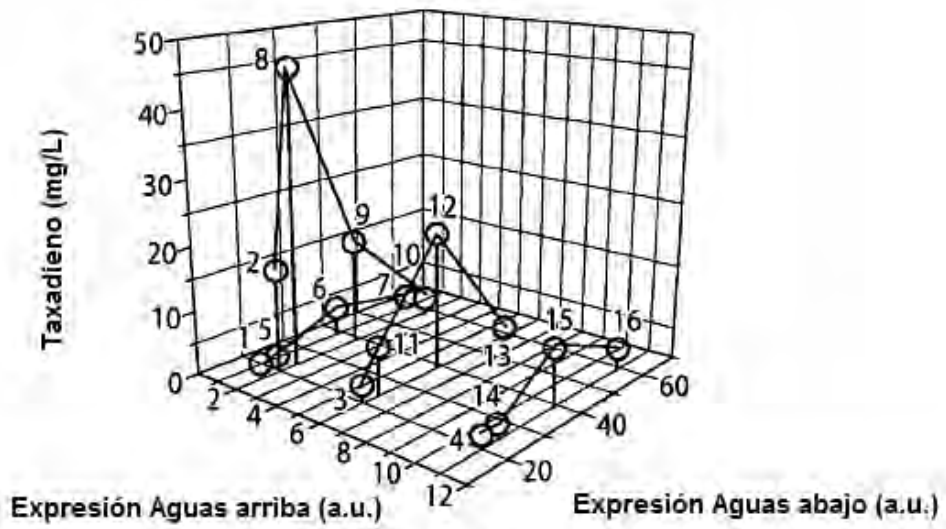


Fig. 2b

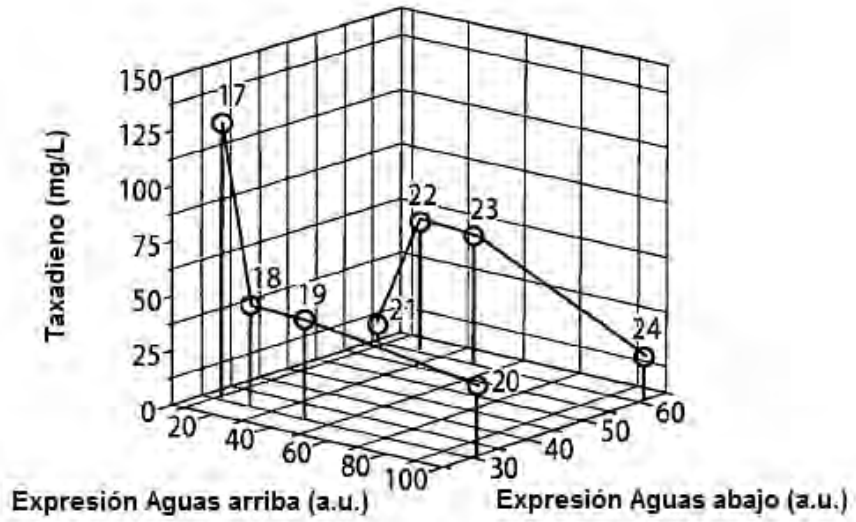


Fig. 2c

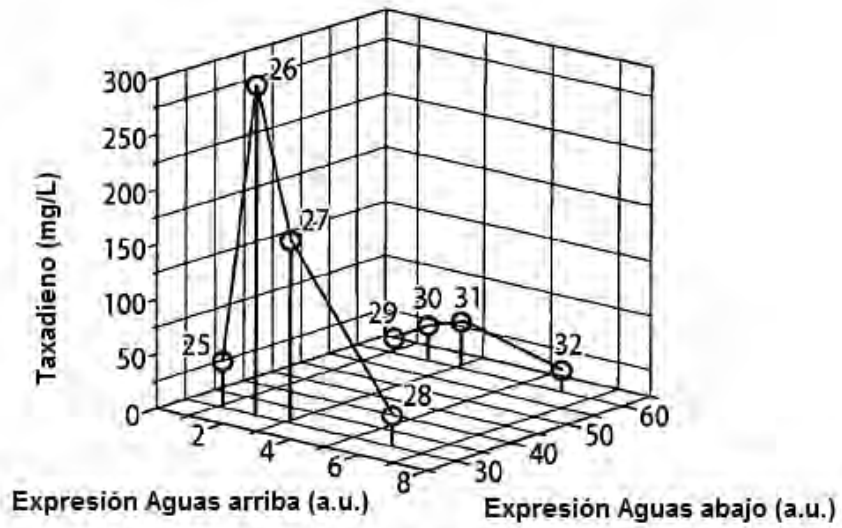


Fig. 2d

1. Ep20TrcGT
2. ECh1TrcMEPp20GT
3. Ep5TrcMEPp20TrcGT
4. Ep10TrcMEPp20TrcGT
5. Ep20TrcTG
6. Ep20T5GT
7. Ep20T5GTTrcT
8. ECh1TrcMEPp20TrcTG
9. ECh1TrcMEPp20T5GT
10. ECh1TrcMEPp20T5GTTrcT
11. Ep5TrcMEPp20TrcTG
12. Ep5TrcMEPp20T5GT
13. Ep5TrcMEPp20T5GTTrct
14. Ep10TrcMEPp20TrcTG
15. Ep10TrcMEPp20T5GT
16. Ep10TrcMEPp20T5GTTrcT
17. EDE3p10TrcMEPp5T7TG
18. EDE3p20TrcMEPp5T7TG
19. EDE3p20T5MEPp5T7TG
20. EDE3p20T7MEPp5T7TG
21. EDE3p5TrcMEPp10T7TG
22. EDE3p20TrcMEPp10T7TG
23. EDE3p20T5MEPp10T7TG
24. EDE3p20T7MEPp10T7TG
25. EDE3p5T7TG
26. EDE3Ch1TrcMEPp5T7TG
27. EDE3Ch1T5MEPp5T7TG
28. EDE3Ch1T7MEPp5T7TG
29. EDE3p10T7TG
30. EDE3Ch1TrcMEPp10T7TG
31. EDE3Ch1T5MEPp10T7TG
32. EDE3Ch1T7MEPp10T7TG

Fig. 2e

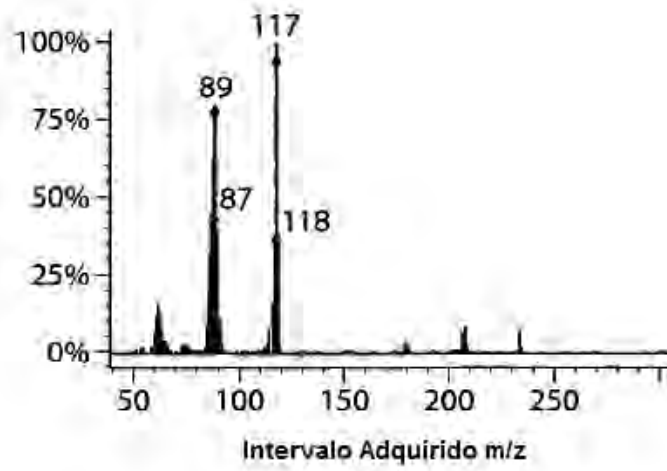


Fig. 3a

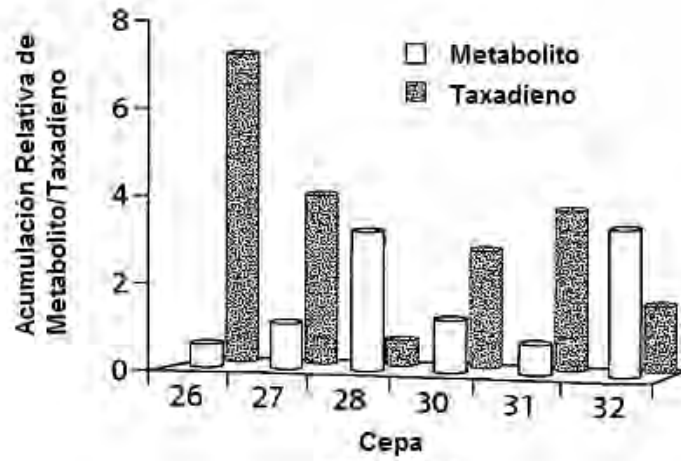


Fig. 3b

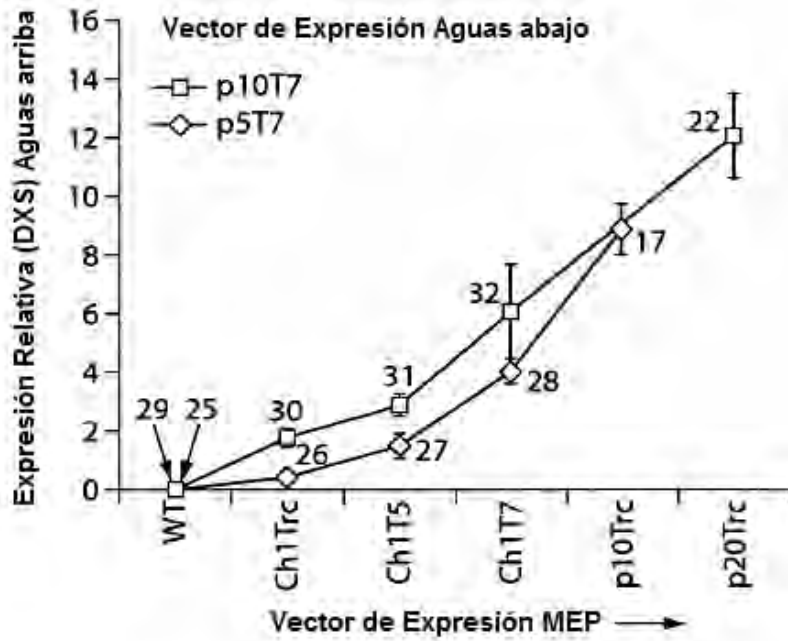


Fig. 4a

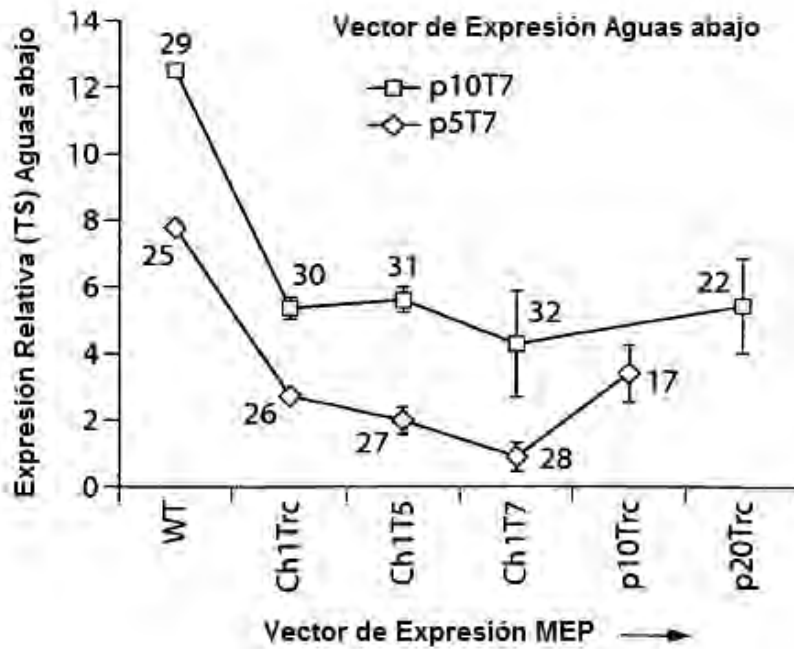


Fig. 4b

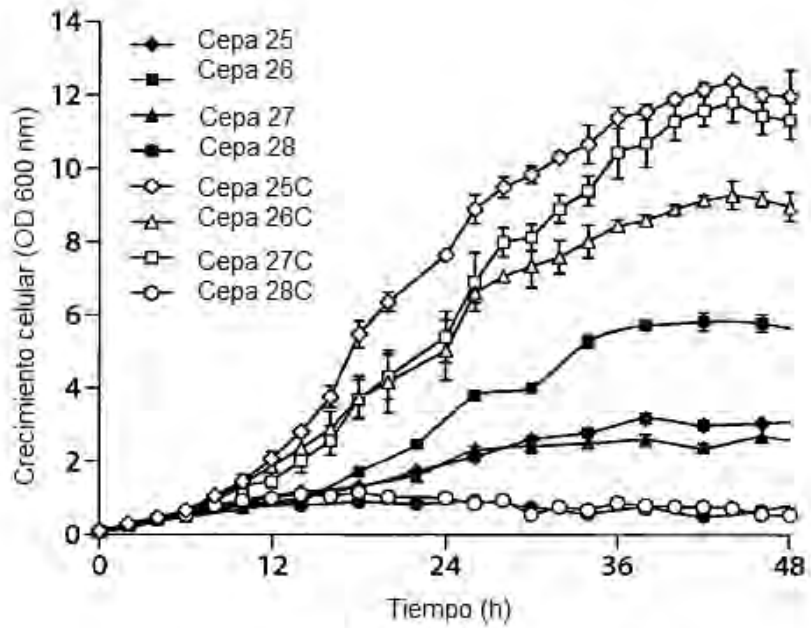


Fig. 4c

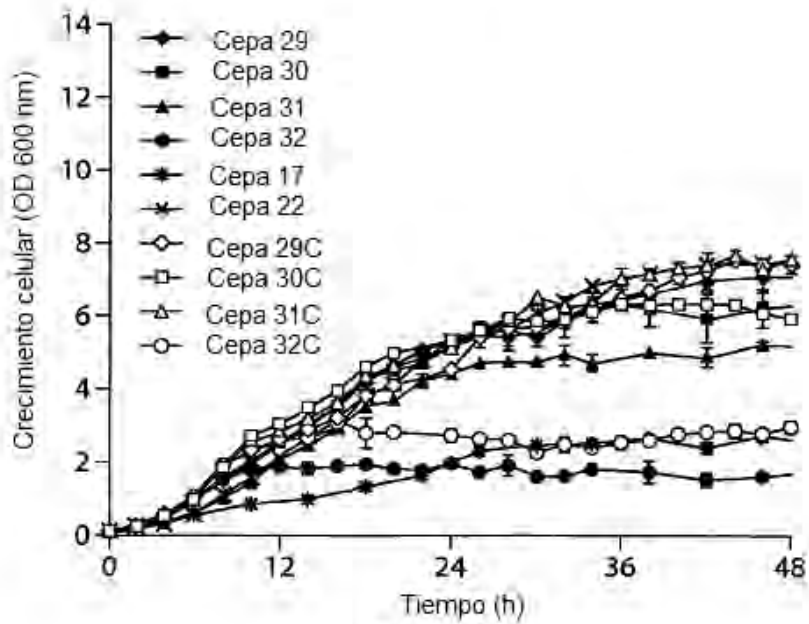


Fig. 4d

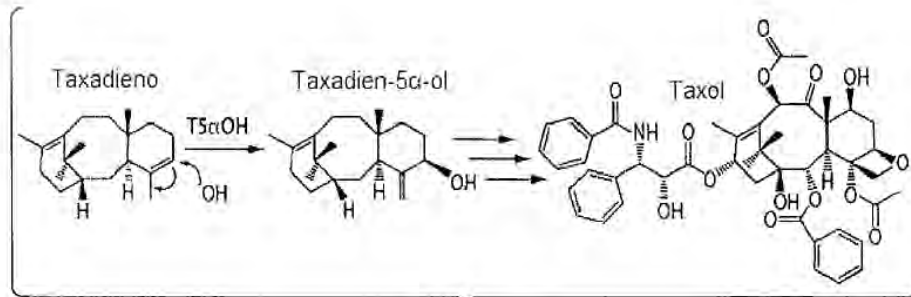


Fig. 5a

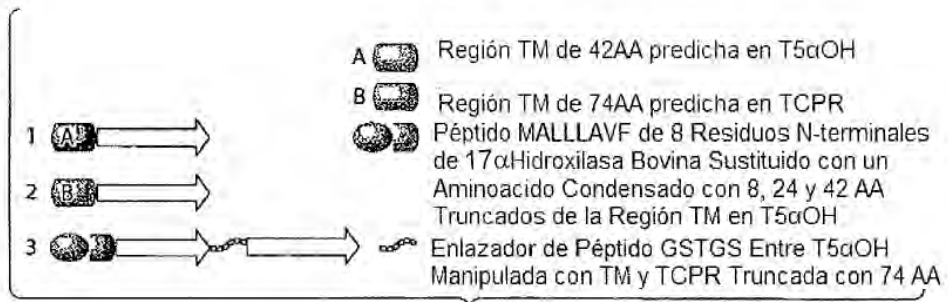


Fig. 5b

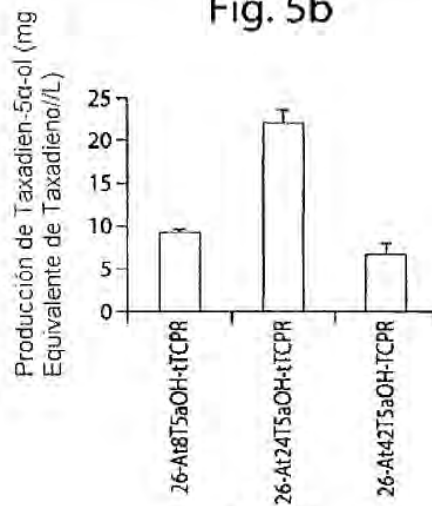


Fig. 5c

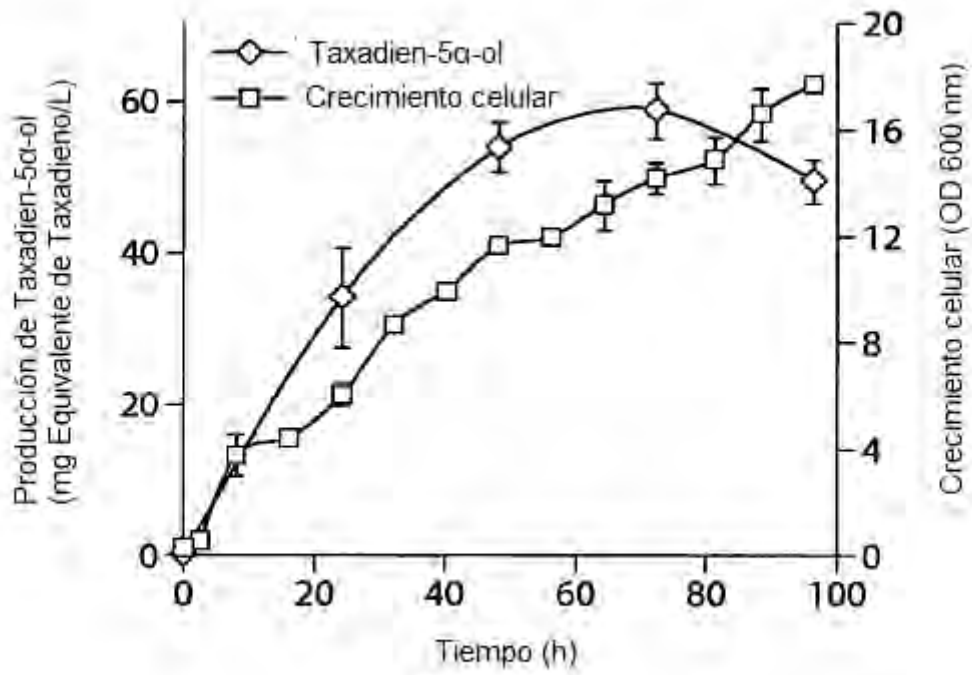


Fig. 5d

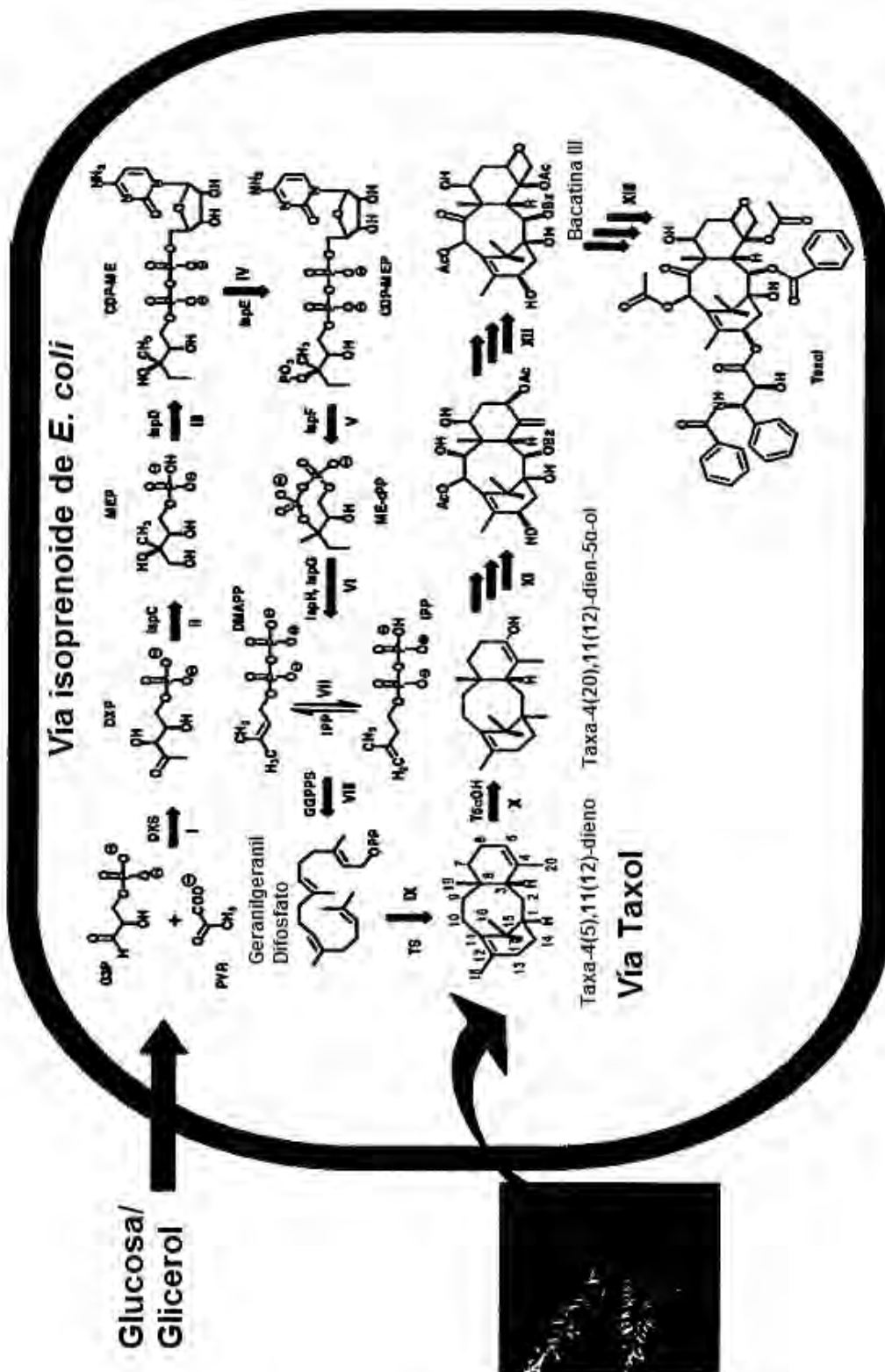


Fig.6

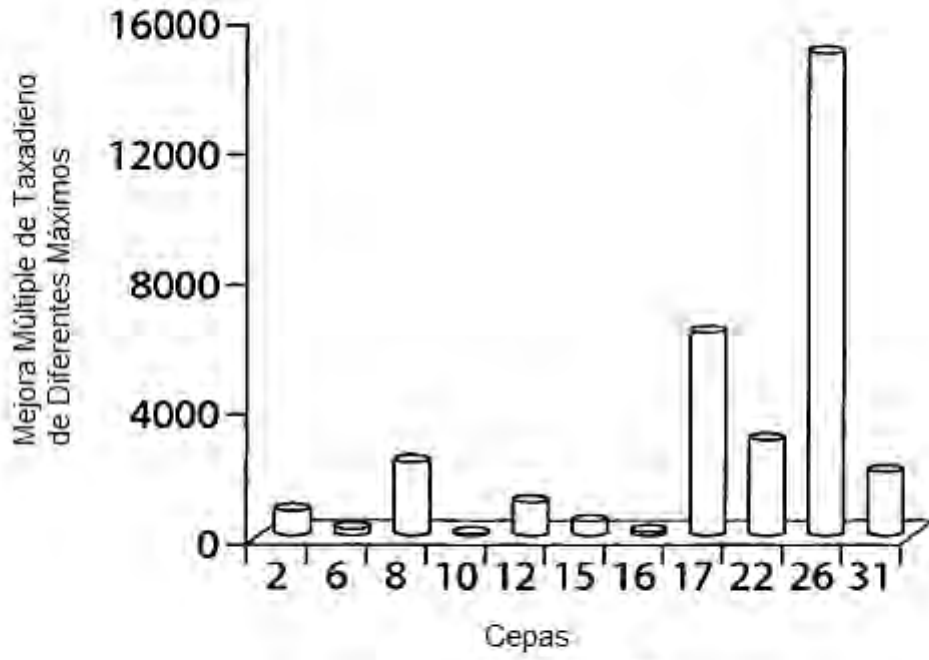


Fig. 7

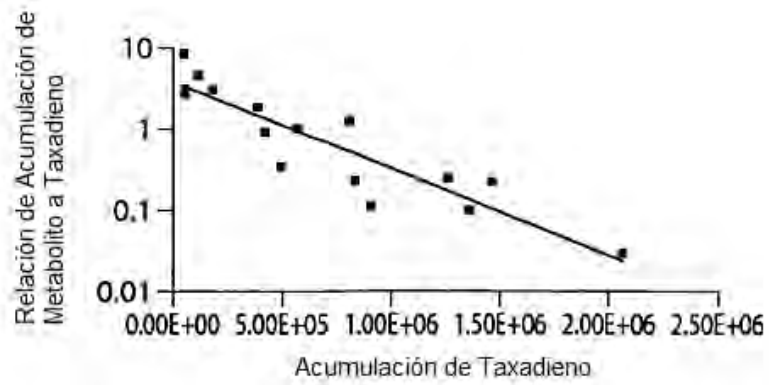


Fig. 8a

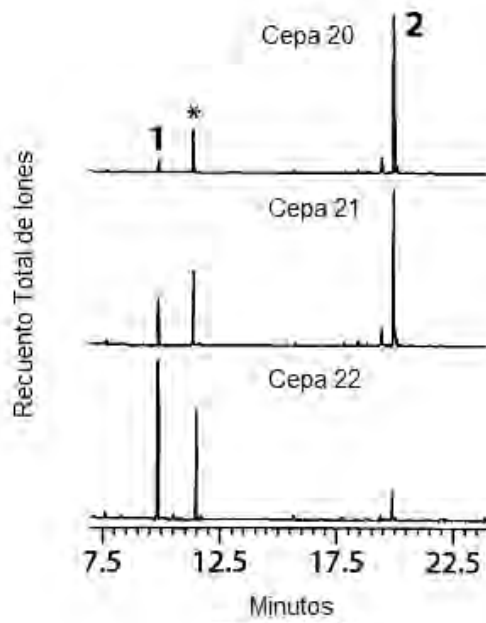


Fig. 8b

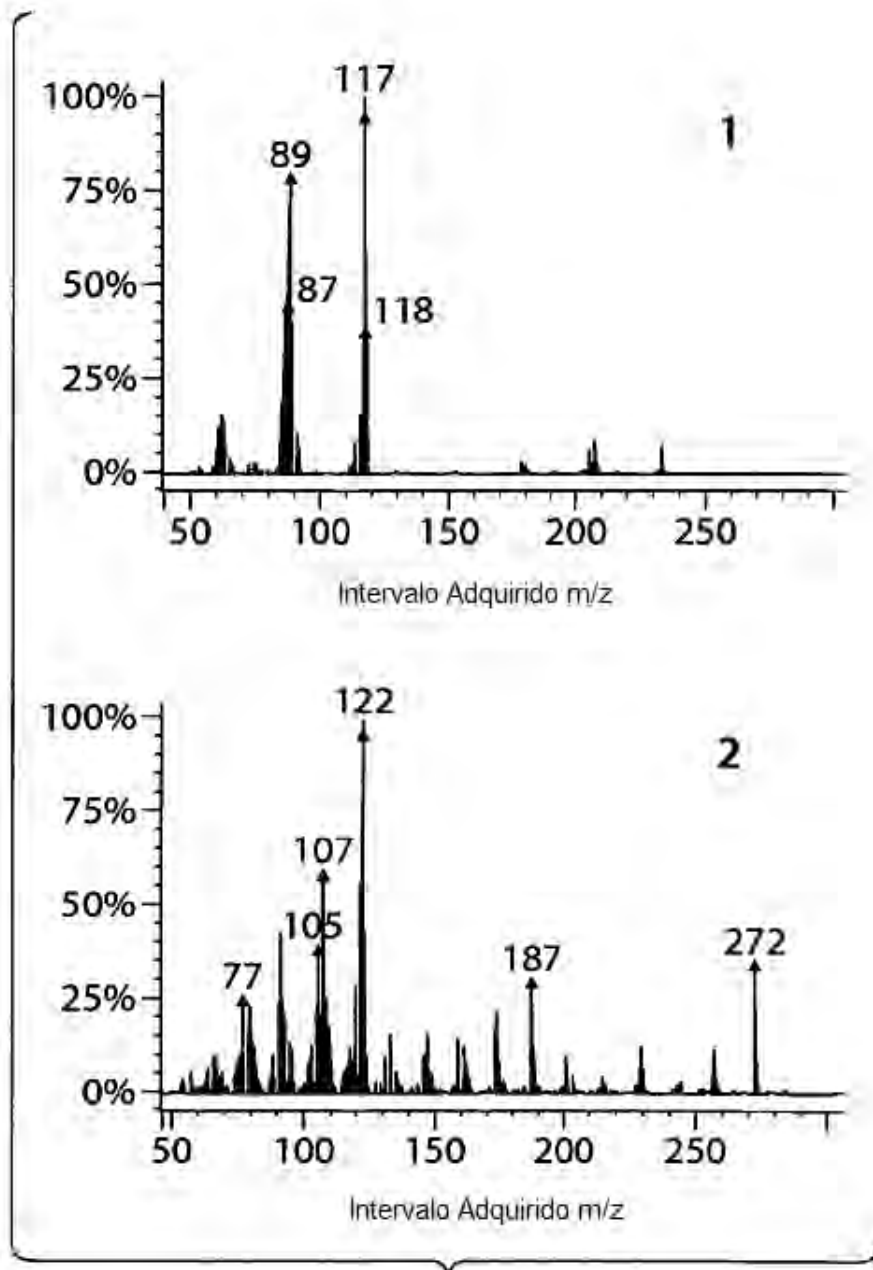


Fig. 8c

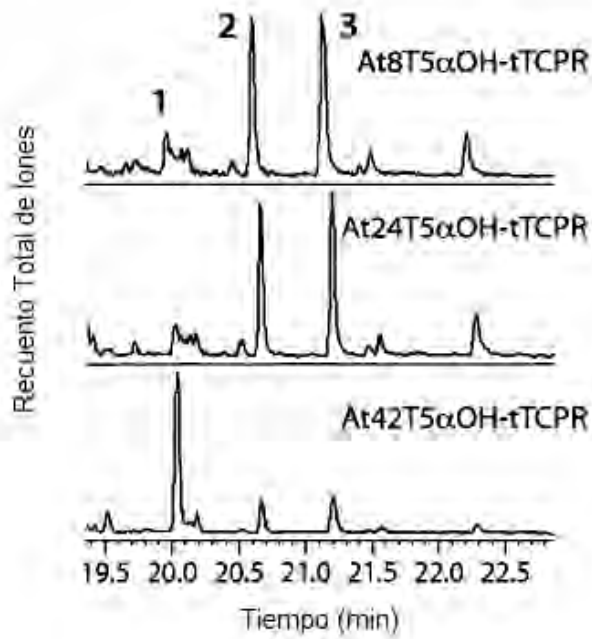


Fig. 9a

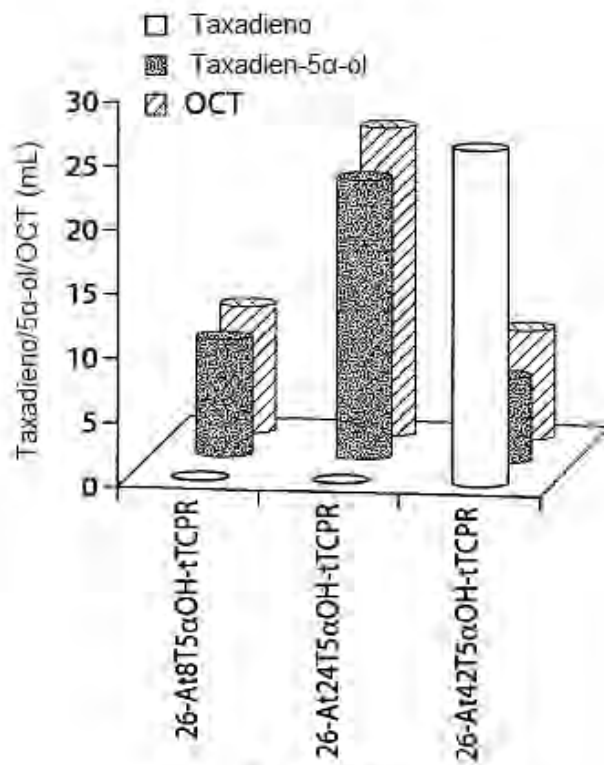


Fig. 9b

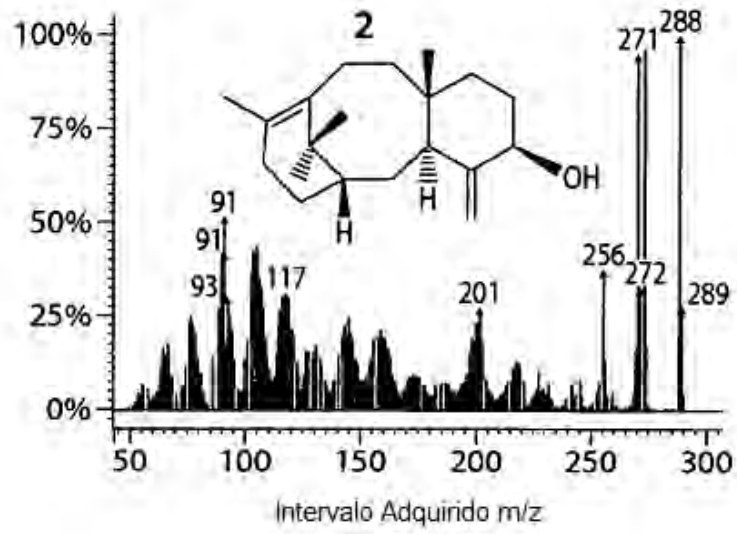


Fig. 9c

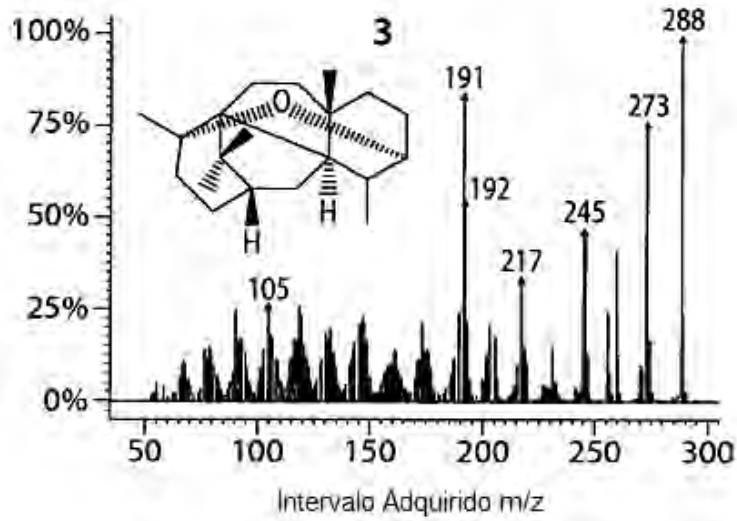


Fig. 9d

Vía Biosintética Terpenoide: Diversidad en Química de Productos Naturales

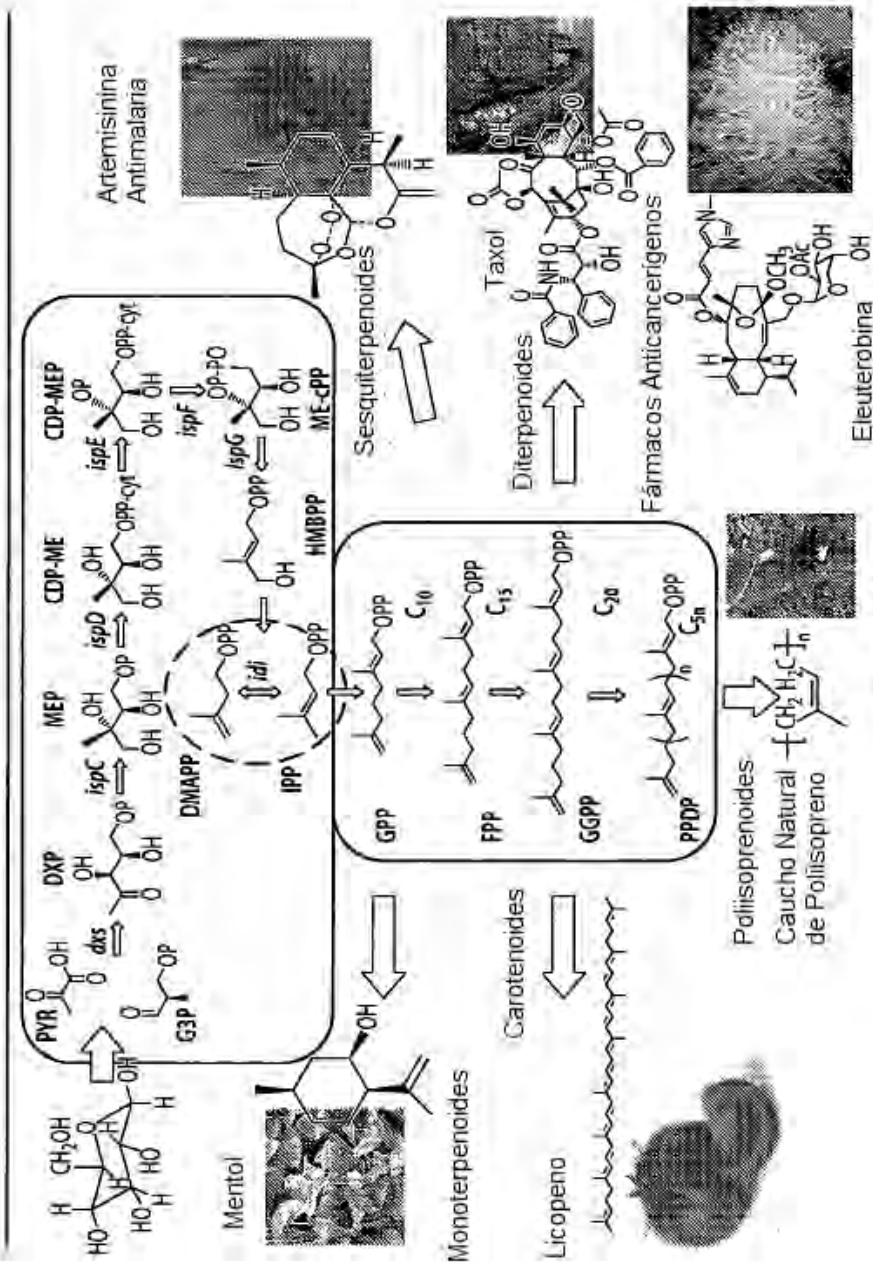


Fig. 10

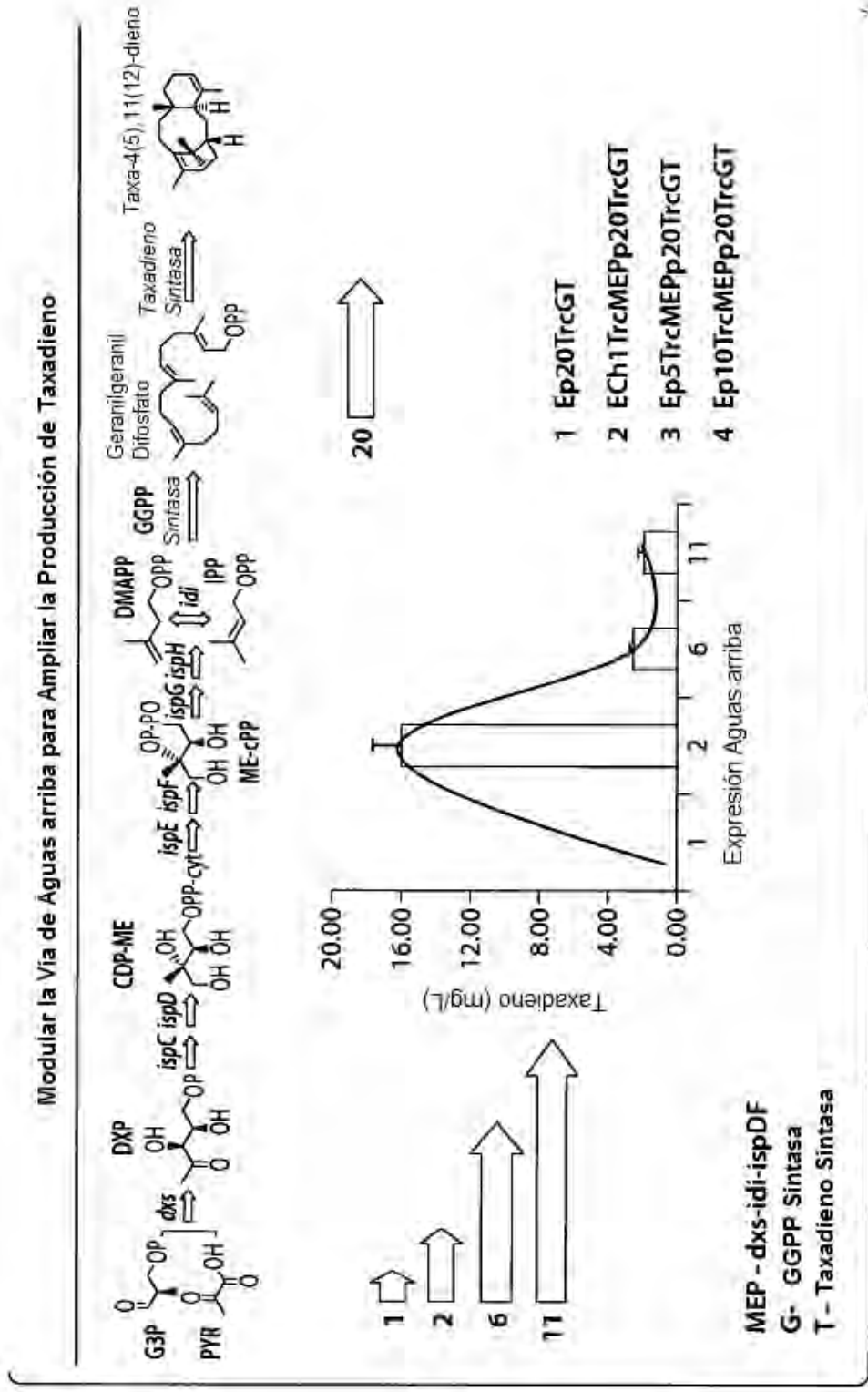
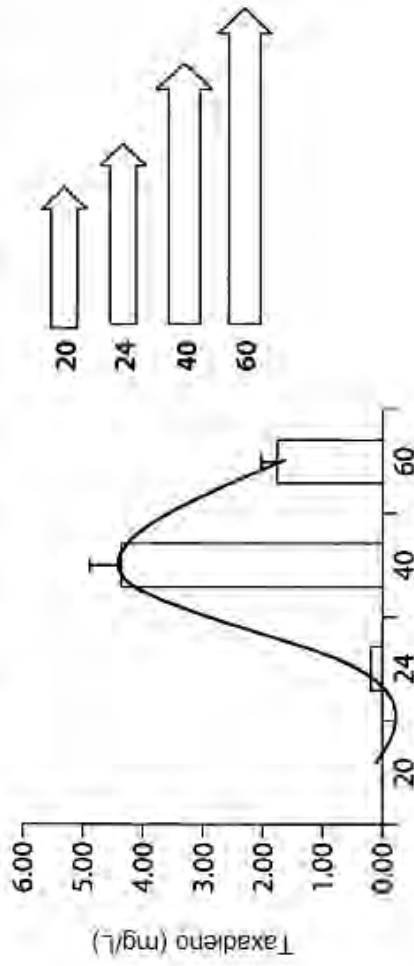
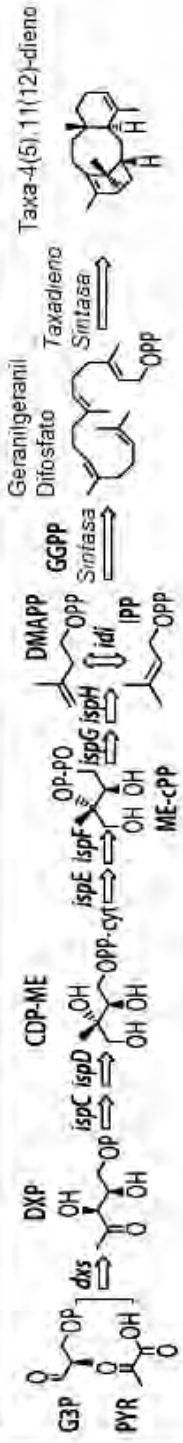


Fig. 11

Modular la Via de Aguas abajo para Ampliar la Producción de Taxadieno



- 1 Ep20TrcGT
- 5 Ep20TrcTG
- 6 Ep20T5GT
- 7 Ep20T5GTTrcT

Fig. 12

La Diversificación de la Vía Recientemente Identificada no es la Característica de la Vía Sintética Aguas abajo

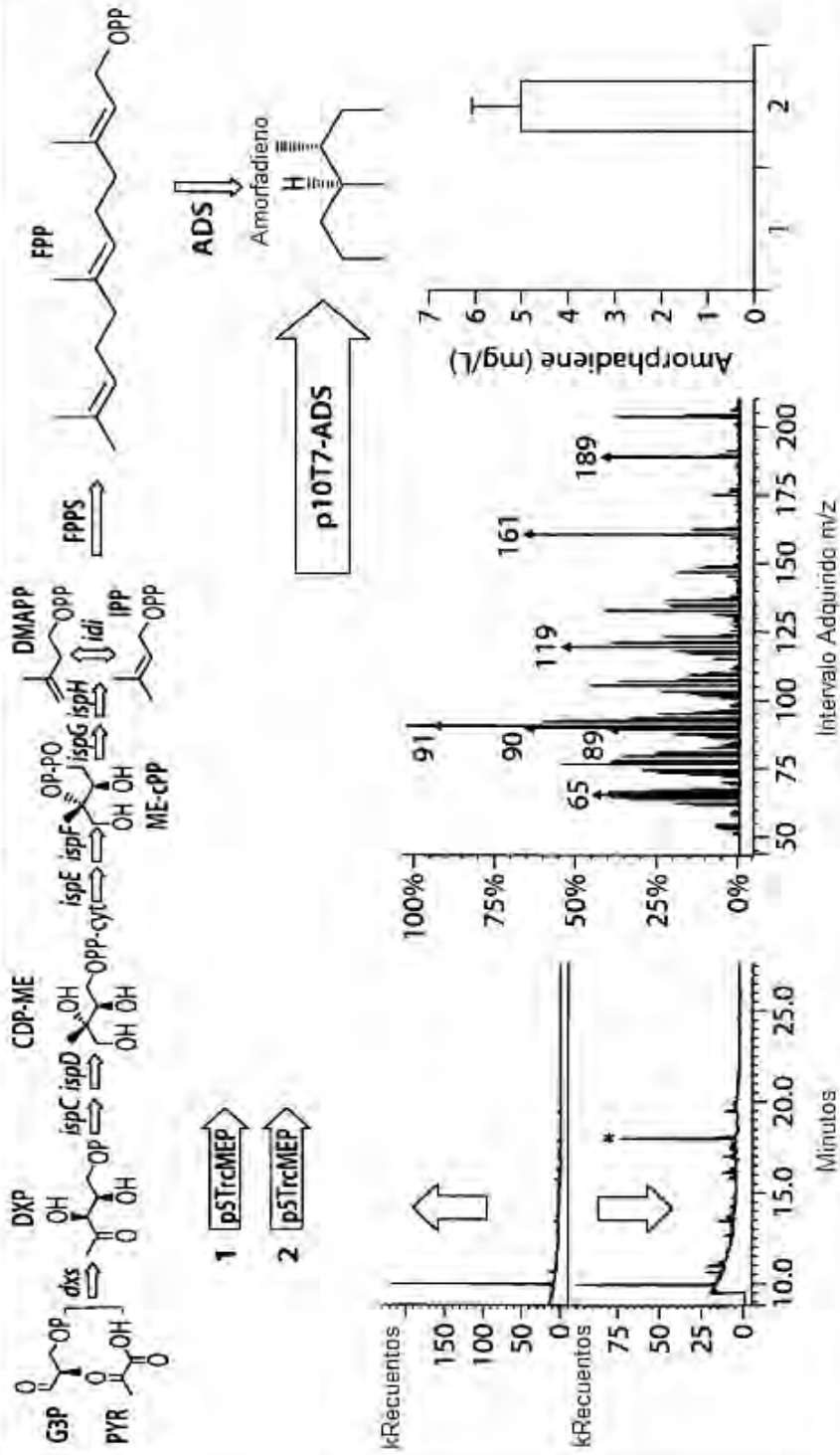


Fig. 13

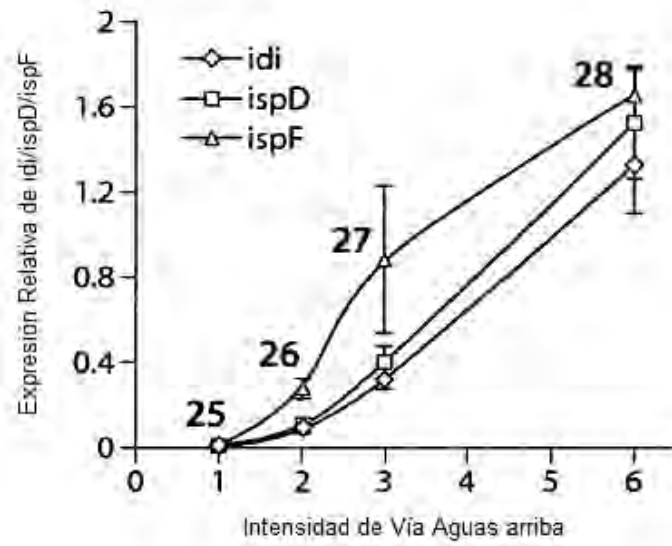


Fig. 14a

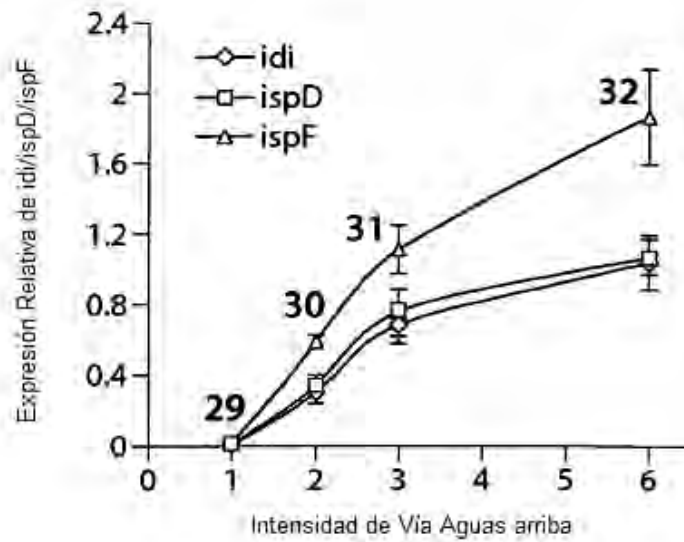


Fig. 14b

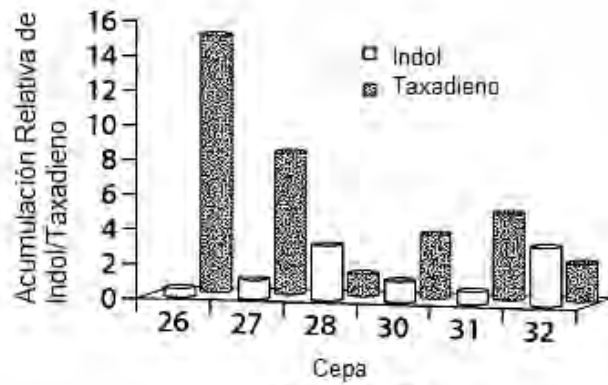


Fig. 15a

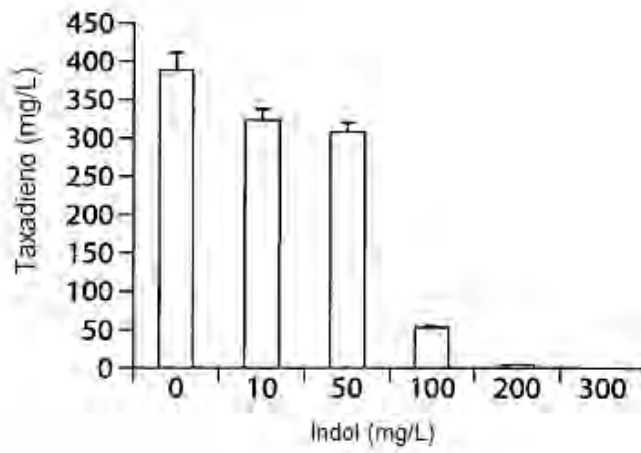


Fig. 15b

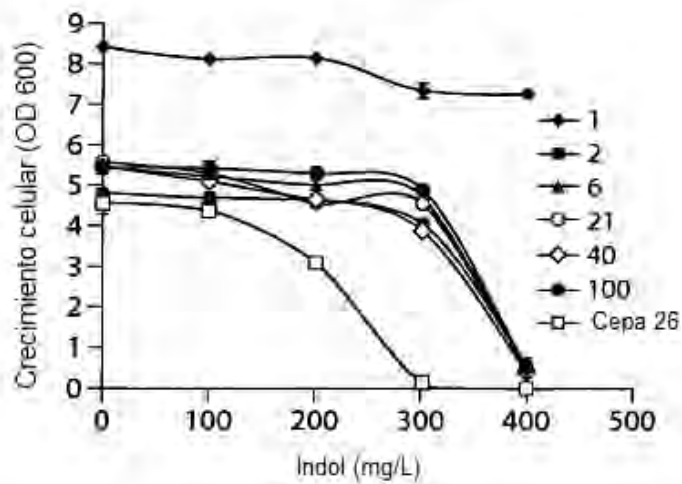


Fig. 15c

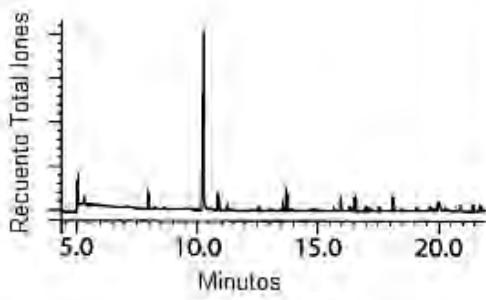


Fig. 16A

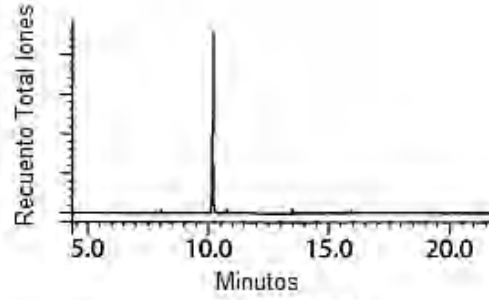


Fig. 16B

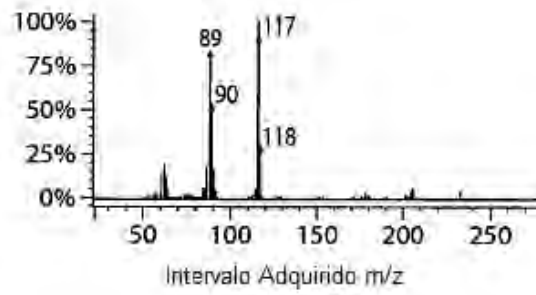


Fig. 16a

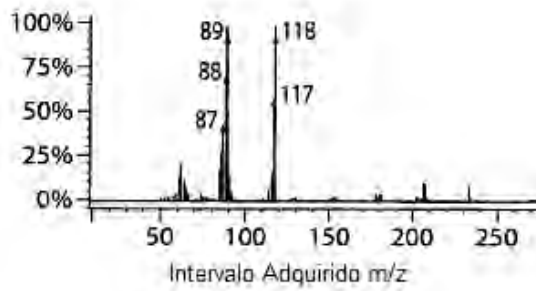


Fig. 16b

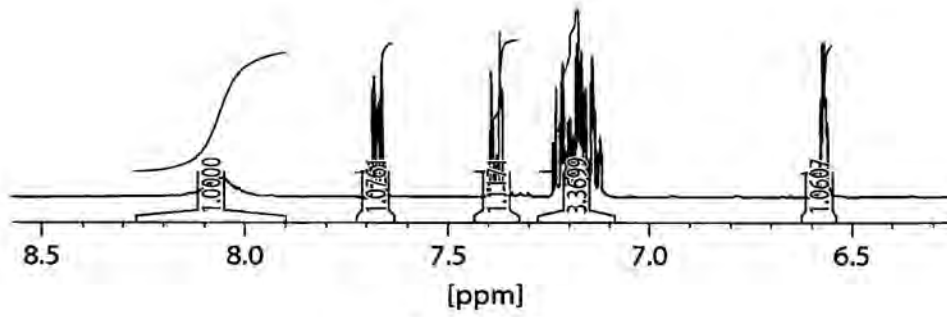


Fig. 16c

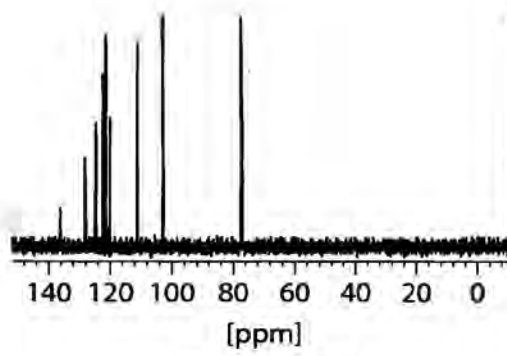


Fig. 16d

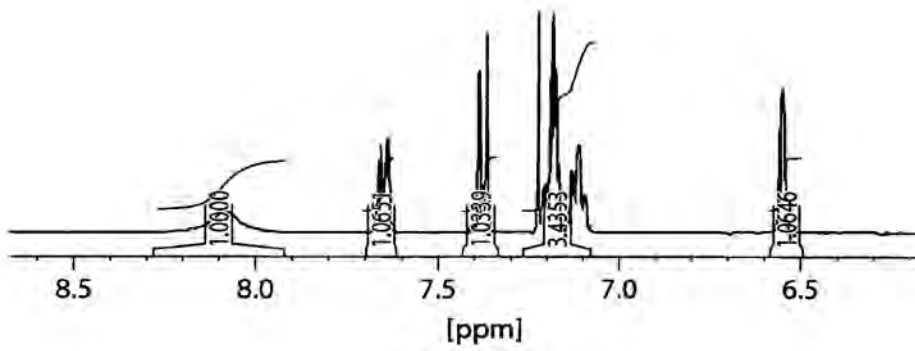


Fig. 16e

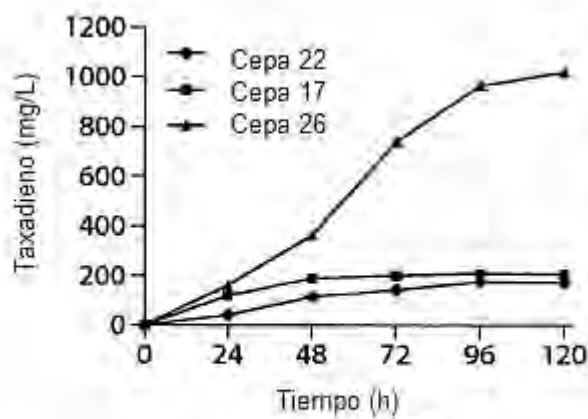


Fig. 17a

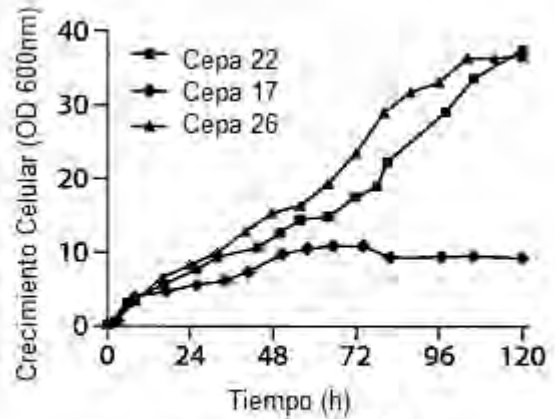


Fig. 17b

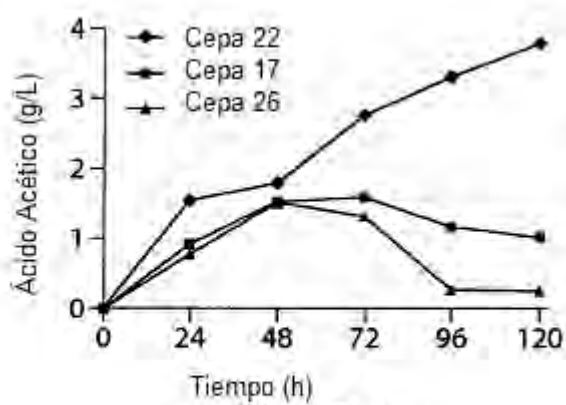


Fig. 17c

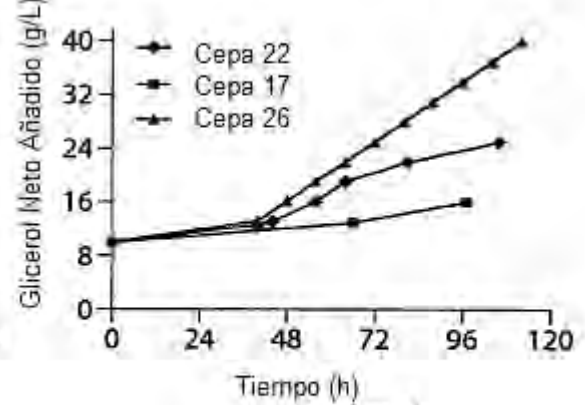


Fig. 17d