

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 699 479**

51 Int. Cl.:

**G01N 33/53** (2006.01)

**C12Q 1/68** (2008.01)

**C12N 15/31** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.09.2014 PCT/US2014/055318**

87 Fecha y número de publicación internacional: **19.03.2015 WO15038853**

96 Fecha de presentación y número de la solicitud europea: **12.09.2014 E 14844036 (5)**

97 Fecha y número de publicación de la concesión europea: **29.08.2018 EP 3044591**

54 Título: **Microorganismos recombinantes y sus métodos de uso**

30 Prioridad:

**12.09.2013 US 201361877272 P**

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

**11.02.2019**

73 Titular/es:

**LANZATECH NEW ZEALAND LIMITED (100.0%)  
239 Onehunga Mall Onehunga  
Auckland 1061, NZ**

72 Inventor/es:

**WALKER, DAVID JEFFREY FRASER;  
NAGARAJU, SHILPA;  
KOEPE, MICHAEL y  
MUELLER, ALEXANDER PAUL**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 699 479 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**

Microorganismos recombinantes y sus métodos de uso

Referencia cruzada a una solicitud relacionada

**Campo de la invención**

- 5 La presente invención se refiere a marcadores de selección de uso en la preparación de Clostridium spp. recombinante.

**Antecedentes de la invención**

- 10 Se conocen procedimientos para producir organismos recombinantes. Por lo general, implican la transformación de un organismo con un vector de ácido nucleico exógeno, que puede integrarse con el genoma del huésped o permanecer en un estado estable independiente (por ejemplo, extracromosómico).

- 15 La integración de un ácido nucleico exógeno en el genoma del huésped implica un evento de doble entrecruzamiento entre el vector y un ácido nucleico endógeno. La recombinación de doble entrecruzamiento se produce a frecuencias que son demasiado bajas para identificar, de manera confiable, los integrantes solo por casualidad. Por lo tanto, cualquier medio para seleccionar uno o ambos entrecruzamientos tiene un gran beneficio en la frecuencia de identificación, tanto en tiempo como en mano de obra.

- 20 Se conocen marcadores de selección de uso en la detección de eventos de recombinación. Dichos marcadores son típicamente secuencias codificantes de proteínas que confieren una ventaja selectiva (selección positiva) o desventaja (contra-selección) frente a un organismo huésped. Se conocen una serie de marcadores de selección positiva y de contra-selección que pueden ser útiles en la detección de organismos en los que se ha producido un evento de recombinación deseado. Un marcador de selección positiva generalmente comprende un gen que, cuando se expresa, permite que un organismo sobreviva en un entorno de crecimiento particular. Un marcador de contra-selección típicamente comprende un gen que, cuando se expresa, produce una toxina que es letal para un organismo.

- 25 En bacterias distintas a Clostridia, hay una gran cantidad de marcadores de contra-selección disponibles, pero desafortunadamente, ya sea por razones fisiológicas o genéticas, la gran mayoría no funciona en Clostridia.

Un objeto de la invención es superar una o más de las desventajas de la técnica anterior o, al menos, proporcionar al público una opción útil.

**Sumario de la invención**

- 30 En un primer aspecto, la invención proporciona el uso de ThiK y/o PheS como marcador de contra-selección en un método para producir un microorganismo recombinante a partir de un microorganismo progenitor, en donde el microorganismo progenitor es una especie de Clostridium, y en donde PheS incluye al menos una alteración en comparación con una PheS natural, de modo que, con el uso de fenilalanina ARNt sintetasa, sea capaz de aminoacilar ARNt utilizando un análogo de fenilalanina.

- 35 En un segundo aspecto, la invención proporciona el uso de un ácido nucleico que codifica ThiK y/o PheS en un plásmido de uso en la producción de un microorganismo recombinante a partir de un microorganismo progenitor, en donde el microorganismo progenitor es una especie de Clostridium, y en donde PheS incluye al menos una alteración en comparación con una PheS natural, de modo que, con el uso de fenilalanina ARNt sintetasa, sea capaz de aminoacilar ARNt utilizando un análogo de fenilalanina.

- 40 En un tercer aspecto, la invención proporciona el uso de un plásmido que comprende un ácido nucleico que codifica ThiK y/o PheS para producir un microorganismo recombinante a partir de un microorganismo progenitor, en donde el microorganismo progenitor es una especie de Clostridium, y en donde PheS incluye al menos una alteración en comparación con una PheS natural, de manera tal que, con su uso, la fenilalanina ARNt sintetasa sea capaz de aminoacilar ARNt utilizando un análogo de fenilalanina.

- 45 En un cuarto aspecto, la invención proporciona un método para la producción de un microorganismo recombinante a partir de un microorganismo progenitor, comprendiendo el método al menos las etapas de:

a) transformar un microorganismo progenitor con un plásmido que comprende:

- 50 1. al menos una secuencia de ácido nucleico que codifica al menos un marcador de contra-selección elegido del grupo que consiste en PheS y ThiK, en donde PheS incluye, al menos, una alteración en comparación con una PheS natural tal que en el uso de fenilalanina ARNt sintetasa sea capaz de aminoacilar ARNt utilizando un análogo de fenilalanina;

2. al menos una secuencia de ácido nucleico que codifica, al menos, un marcador de selección positiva; y,

3. dos secuencias de ácido nucleico homólogas a las regiones seleccionadas alrededor de una ubicación objetivo dentro del genoma del microorganismo progenitor, que permiten la recombinación del plásmido con el genoma del microorganismo progenitor;

b) seleccionar uno o más microorganismos que expresen al menos un marcador de selección positiva; y,

5 c) seleccionar uno o más microorganismos que no expresen al menos un marcador de contra-selección.

En una realización, los pasos de selección b) y c) se llevan a cabo simultáneamente. En otra realización, los pasos de selección b) y c) se realizan secuencialmente.

En una realización, el plásmido comprende, además, al menos, una secuencia de ácido nucleico de interés para insertarse en el genoma progenitor.

10 En un cuarto aspecto, la invención proporciona un ácido nucleico que codifica un PheS, en el que el PheS se altera en comparación con un PheS natural, de modo que en el uso de fenilalanina tRNA sintetasa sea capaz de aminoacilar el ARNt usando un análogo de fenilalanina, y en el que la PheS natural se derive de un *Clostridium* spp o sea una variante funcionalmente equivalente de ésta.

15 En un quinto aspecto, la invención proporciona un vector de ácido nucleico que comprende un ácido nucleico de acuerdo con el cuarto aspecto de la invención. En una realización, el vector es un plásmido.

En una realización, el vector es un plásmido y también comprende uno o más de:

a. al menos una secuencia de ácido nucleico que codifica al menos un marcador de selección positiva; y,

20 b. dos secuencias de ácido nucleico homólogas a regiones seleccionadas alrededor de una ubicación objetivo dentro del genoma de un microorganismo progenitor, que permiten la recombinación del plásmido con el genoma del microorganismo progenitor.

En una realización, el plásmido comprende además al menos una secuencia de ácido nucleico de interés que se desea insertar en el genoma de un microorganismo progenitor.

25 En un sexto aspecto, la invención proporciona un PheS, en el que el PheS comprende una o más alteraciones en comparación con un PheS natural, de modo que, con el uso de fenilalanina tRNA sintetasa, sea capaz de aminoacilar el ARNt usando un análogo de fenilalanina, y en el que el PheS natural se derive de un *Clostridium* spp o sea una variación funcionalmente equivalente de ésta.

En un séptimo aspecto, la invención proporciona una célula que comprende un ácido nucleico de acuerdo con el cuarto aspecto de la invención, un vector de acuerdo con el quinto aspecto de la invención y/o un PheS de acuerdo con el sexto aspecto de la invención.

30 En una realización de los aspectos anteriores de la invención, la PheS natural y/o el ácido nucleico natural que codifica PheS se deriva de un *Clostridium* spp o es una variante funcionalmente equivalente de ésta.

En una realización de los aspectos anteriores y las realizaciones de la invención, la al menos una alteración en PheS, en comparación con una natural, incluye una o más sustituciones, deleciones y/o adiciones de aminoácidos.

35 En una realización particular de los aspectos y realizaciones anteriores de la invención, la al menos una alteración en PheS está localizada dentro del sitio de especificidad del sustrato. En una realización, el sitio de especificidad de sustrato se localiza entre los aminoácidos 306 y 313 leídos con relación a la posición de aminoácidos de PheS natural de *C. autoethanogenum* (SEQ ID 21). En una realización, la al menos una alteración es una sustitución de aminoácido en la posición 311. En una realización, la al menos una alteración es la sustitución de Cys por Gly en el aminoácido 311.

40 En una realización de los aspectos y realizaciones anteriores de la invención, la PheS se deriva de *Clostridium autoethanogenum* o es una variante funcionalmente equivalente de la misma.

En una realización de los aspectos y realizaciones anteriores de la invención, la PheS alternada comprende la secuencia de aminoácidos de la SEQ ID N° 21.

45 En una realización de los aspectos y realizaciones anteriores de la invención, el ácido nucleico que codifica PheS que incluye al menos una alteración en comparación con una PheS natural, comprende al menos una alteración en comparación con un ácido nucleico que codifica una PheS natural. En una realización, la al menos una alteración en un ácido nucleico que codifica PheS incluye una o más sustituciones, deleciones y/o adiciones de nucleótidos. En una realización, la una o más alteraciones en la secuencia de ácido nucleico está localizada dentro de una región del ácido nucleico que codifica el sitio de especificidad de sustrato de PheS. En una realización, la región de un ácido nucleico que codifica el sitio de especificidad de sustrato se localiza entre las bases 918 y 939, leída con relación a la posición de nucleótido del gen que codifica PheS natural de *C. autoethanogenum* (SEQ ID 12). En una

50

realización, la al menos una alteración es una sustitución de nucleótido en la base 932. En una realización, la al menos una alteración es la sustitución de C por G en la base 932.

En una realización de los aspectos y realizaciones anteriores de la invención, el ácido nucleico que codifica PheS se deriva de *Clostridium autoethanogenum* o es una variante funcionalmente equivalente de ésta.

- 5 En una realización de los aspectos y realizaciones anteriores de la invención, el ácido nucleico que codifica PheS alterado comprende la secuencia de la SEQ ID N° 14.

10 En una realización de los aspectos y realizaciones anteriores de la invención, el análogo de fenilalanina se elige entre clorofenilalanina, fluorofenilalanina y bromofenilalanina. En una realización particular, el análogo de la fenilalanina se elige entre la DL-4-clorofenilalanina y la p-clorofenilalanina, p-fluoro-L-fenilalanina, p-fluoro-DL-fenilalanina, p-bromo-L-fenilalanina.

En una realización de los aspectos y realizaciones anteriores de la invención, el ThiK y/o el ácido nucleico que codifica ThiK proviene del Virus del Herpes Simple 1 o Virus del Herpes Simple 2 (VHS-TK), VZV, CMV, HHV7, HHV8, EBV o es una variante funcionalmente equivalente de uno o más de estos.

### Breve descripción de las figuras

- 15 Estos y otros aspectos de la presente invención, que deben considerarse en todos sus aspectos novedosos, se harán evidentes a partir de la siguiente descripción, que se ofrece a modo de ejemplo solamente, con referencia a las figuras adjuntas, en las que:

Figura 1: Muestra el mapa del plásmido pMTL83155-HSV-tk

20 Figura 2: Muestra la amplificación por PCR de un fragmento de ~1,5 kb que abarca el replicón gram positivo y el marcador catP en pMTL83155 y pMTL83155-Hsv-tk (h1 y h2) en transformantes de *C. autoethanogenum*.

Se utilizó *C. autoethanogenum* (C) no modificado como control. Se analizaron 2 colonias de LZ-pMTL83155 (P1 y P2) y LZ-pMTL83155-Hsv-tk (h1 y h2).

Figura 3: Muestra la alineación de la secuencia de nucleótidos traducida por pares de pheS de *E. coli* MG1655 (Seq ID 13) y *C. autoethanogenum* con la supuesta región de especificidad de sustrato en negrita y subrayada.

25 Figura 4: Muestra el mapa del plásmido pMTL85151-p/zeS\*

Figura 5: Muestra un mapa representativo de un plásmido que comprende HSV-tk.

Figura 6: Muestra un mapa representativo de un plásmido que comprende pheS.

30 Figura 7: Muestra los resultados de la electroforesis en gel de agarosa TAE de la selección por PCR con AM041 y AM042. El carril 1 contiene la escalera GeneRuler 1kb (Thermo). El carril 2 contiene PCR sin plantilla añadida. El carril 3 PCR con ADN genómico de *C. autoethanogenum* natural como plantilla que muestra el tamaño esperado del producto natural (3137 pb). Los carriles 4-7 contienen PCR con colonias resistentes a p-clorofenilalanina y tiamfenicol como plantilla. El carril 6 muestra un producto de PCR del tamaño esperado (3570 pb) para el reemplazo exitoso del doble entrecruzamiento del gen nativo de butanodiol deshidrogenasa con el gen de butanodiol deshidrogenasa de *K. pneumonia*.

35 Figura 8: Mapa genético que muestra los sitios dianas del intrón (211s, 287a, 388a, 400s, 433s y 552a). También se muestran los sitios de unión del cebador (flechas inferiores, horizontales).

Figura 9A -9C: Confirmación de las inserciones del intrón del grupo II. Fig. 9A: 433s y 388a (banda débil), Fig. 9B: 211s y 287a. Fig. 9C: 287a y 433

### Breve descripción de la información de la secuencia

40 Antes de las figuras mostradas en este documento en adelante, la memoria descriptiva incluye detalles de las secuencias de ácidos nucleicos y polipéptidos relevantes para la invención. Se proporcionan las siguientes secuencias:

Seq. ID.1: Secuencia de ácido nucleico de pMK-RQ-Hsv-tk

Seq. ID.2: Secuencia de ácido nucleico de pMTL83155-Hsv-tk

45 Seq. ID.3: Secuencia de ácido nucleico de pMTL83155

Seq. ID.4: Secuencia de ácido nucleico del cebador repHf

Seq. ID.5: Secuencia de ácido nucleico del cebador catr

- Seq. ID.6: Secuencia de ácido nucleico del cebador fD1
- Seq. ID.7: Secuencia de ácido nucleico del cebador rP2
- Seq. ID.8: Secuencia de ácido nucleico rRNA 16s de LZ-pMTL83155-1 obtenida usando el cebador rP2
- Seq. ID.9: Secuencia de ácido nucleico rRNA 16s de LZ-pMTL83155-2 obtenida usando el cebador rP2
- 5 Seq. ID.10: Secuencia de ácido nucleico rRNA 16s de LZ-pMTL83155-hsv-tk-1 obtenida usando el cebador rP2
- Seq. ID.11: Secuencia de ácido nucleico rRNA 16s de LZ-pMTL83155-hsv-tk-2 obtenida usando el cebador rP2
- Seq. ID.12: Secuencia de ácido nucleico que codifica pheS de *C. autoethanogenum*
- Seq. ID.13: Secuencia de ácido nucleico que codifica pheS de *E. coli* MG1655
- Seq. ID.14: Secuencia de ácido nucleico que codifica pheS\* alterado de *C. autoethanogenum*
- 10 Seq. ID.15: Secuencia de cebador directo utilizada para confirmar la presencia del plásmido PheS – M13F
- Seq. ID.16: Secuencia de cebador inverso utilizada para confirmar la presencia del plásmido PheS - M13R
- Seq. ID.17: Promotor sintético PpheS\*
- Seq. ID.18: Secuencia de nucleótidos de pMTL85151pheS\*
- 15 Seq. ID. 19: Secuencia de ácido nucleico que codifica HSV-TK del virus del herpes 1 humano (virus del herpes simple tipo 1)
- Seq. ID. 20: Secuencia de aminoácidos de PheS de *E. coli* MG1655
- Seq. ID. 21: Secuencia de aminoácidos de PheS de *C. autoethanogenum*
- Seq. ID. 22: Secuencia de ácido nucleico que codifica ThiK del virus del herpes 1 humano (virus del herpes simple tipo 1)
- 20 Seq. ID. 23: Secuencia de ácido nucleico que codifica CatP de *Clostridium perfringens*
- Seq. ID. 24: Secuencia de ácido nucleico que codifica ErmB de *Peptoclostridium difficile*
- Seq. ID. 25: Secuencia de ácido nucleico que codifica TetA de *Escherichia coli*
- Seq ID 26: Secuencia de ácido nucleico del casete de PheS\* y ColE1 con el producto de la PCR traJ para el montaje de pPheS\*-ErmB
- 25 Seq ID 27: Secuencia de ácido nucleico del producto de la PCR del fragmento pPheS para el ensamblaje de pPheS\*-CaBDHXXKpBDH (Ejemplo 2)
- Seq ID 30: Secuencia de ácido nucleico del producto de la PCR del origen de replicación pCB102 utilizado para el ensamblaje de pPheS\*-ErmB
- 30 Seq ID 33: Secuencia de ácido nucleico del producto de la PCR del casete de la resistencia a la eritromicina utilizado para el ensamblaje de pPheS\*-ErmB
- Seq ID 36: Secuencia de ácido nucleico de la plantilla pPheS-ErmB para la amplificación del plásmido principal para el ensamblaje de pPheS-CaBDHXXKpBDH
- Seq ID 39: Secuencia de ácido nucleico del producto de PCR del brazo de homología anterior para el ensamblaje de pPheS-CaBDHXXKpBDH
- 35 Seq ID 42: Secuencia de ácido nucleico del producto de la PCR del gen de *K. pneumoniae* butanodiol deshidrogenasa para el ensamblaje de pPheS-CaBDHXXKpBDH
- Seq ID 45: Secuencia de ácido nucleico del producto de la PCR del casete de cloranfenicol acetiltransferasa para el ensamblaje de pPheS-CaBDHXXKpBDH
- 40 Seq ID 48: Secuencia de ácido nucleico del producto de la PCR del brazo de homología posterior para el ensamblaje de pPheS\*-CaBDHXXKpBDH
- Seq ID 53: Secuencia de aminoácidos de pheS\* alterado de *C. autoethanogenum*

Seq ID 54: Secuencia de aminoácidos del virus del herpes humano 1 (virus del herpes simple tipo 1)

Seq ID 55: Secuencia de ácido nucleico de alcohol deshidrogenasa primaria:secundaria de *C. autoethanogenum*

Seq ID 56: Región de reconocimiento del intrón

Seq ID 57: Secuencia del cebador para 156F

5 Seq ID 58: Secuencia del cebador para 939R

Otras secuencias de relevancia para la invención se describen en otra parte en este documento. Por ejemplo, véase la Tabla 3 en el Ejemplo 2.

### Descripción detallada de la invención

10 Lo siguiente es una descripción de la presente invención, que incluye realizaciones preferidas de la misma, dadas en términos generales. La invención se explica adicionalmente a partir de la descripción dada bajo el encabezado "Ejemplos" a continuación en el presente documento, que proporciona datos experimentales que respaldan la invención, ejemplos específicos de diversos aspectos de la invención y medios para realizar la invención.

15 La producción de microorganismos recombinantes puede implicar la introducción de un ácido nucleico exógeno en un microorganismo progenitor, con un evento de recombinación de doble entrecruzamiento que ocurra entre el ácido nucleico exógeno y el genoma del microorganismo, de modo que se pueda introducir al menos una alteración genética deseada en el genoma. La recombinación de doble entrecruzamiento se produce en frecuencias que normalmente son demasiado bajas para identificar de manera confiable los integrantes solo por casualidad. Por lo tanto, los inventores creen que un medio para seleccionar uno o ambos entrecruzamientos tiene un gran beneficio en la frecuencia de identificación tanto en tiempo como en mano de obra. Los inventores han observado en su laboratorio que la frecuencia de recombinación de un solo entrecruzamiento, aunque baja, se puede encontrar al seleccionar un número apropiado de colonias, sin embargo, este no ha sido el caso con el segundo evento de entrecruzamiento. La presente invención proporciona un medio para seleccionar el segundo evento mediante una contra-selección frente a un producto génico letal de la condición presente en el ácido nucleico exógeno introducido en el microorganismo progenitor. Esto significa que en cualquier microorganismo en el que solo haya ocurrido un evento de entrecruzamiento único en presencia de un agente de contra-selección, la expresión del producto genético letal de la condición, matará cualquier célula que no haya sufrido el evento de entrecruzamiento secundario y haya liberado el gen que contiene el ácido nucleico que codifica el marcador de contra-selección.

20 Se conocen marcadores de contra-selección. Sin embargo, no son necesariamente transferibles para su uso en diferentes géneros de bacterias. Desafortunadamente, ya sea por razones fisiológicas o genéticas, la gran mayoría no trabaja en *Clostridia*. Los inventores han identificado sorprendentemente que ThiK y una versión alterada de PheS pueden usarse como marcadores de contra-selección en *Clostridium* spp.

### Definiciones

25 Los "ácidos nucleicos exógenos" son ácidos nucleicos que se originan fuera del microorganismo en el que se introducen. Los ácidos nucleicos exógenos pueden derivarse de cualquier fuente apropiada, incluidos, entre otros, el microorganismo en el que se van a introducir, las cepas o especies de organismos que difieren del organismo en el que se van a introducir, o pueden ser creados artificial o recombinantemente.

30 Una "modificación genética" debe tomarse ampliamente y se pretende que incluya, por ejemplo, la introducción de una mutación en un sitio genético, agregando o eliminando del genoma uno o más nucleótidos, la sustitución de uno o más nucleótidos con diferentes nucleótidos, la sustitución de un gen, la eliminación de un gen, la adición de un gen y similares.

35 En el presente documento se puede hacer referencia a una "PheS alterada", una "PheS que está alterada" o una PheS que incluye una o más o al menos una "alteración" en comparación con una PheS natural tal que, en su uso, la fenilalanina ARNt sintetasa puede aminoacilar el ARNt utilizando un análogo de fenilalanina. Una "alteración" debe considerarse ampliamente e incluye, por ejemplo, uno o una combinación de sustitución de uno o más aminoácidos, la eliminación de uno o más aminoácidos y/o la adición de uno o más aminoácidos en comparación con un PheS original. Un PheS "alterado" también puede incluir una o más alteraciones además de las que permiten a la fenilalanina ARNt sintetasa aminoacilar el ARNt utilizando un análogo de fenilalanina, siempre que aún sea capaz de realizar sustancialmente su función deseada.

40 Se puede hacer referencia en este documento a un ácido nucleico que codifique una PheS que comprenda una o más alteraciones en comparación con un ácido nucleico que codifique una PheS natural. Una "alteración" debe considerarse ampliamente e incluye, por ejemplo, uno o una combinación de sustitución de uno o más nucleótidos, eliminación de uno o más nucleótidos, y/o adición de uno o más nucleótidos en comparación con un ácido nucleico que codifique un PheS original. El ácido nucleico también puede incluir una o más alteraciones además de las que permiten a la fenilalanina ARNt sintetasa aminoacilar el ARNt utilizando un análogo de fenilalanina, siempre que la

PheS aún pueda realizar sustancialmente su función deseada.

Una o más alteraciones de PheS o un ácido nucleico que codifica PheS se pueden describir en este documento con referencia a una región específica o posición de aminoácidos o nucleótidos en un PheS natural (o ácido nucleico que codifica el mismo) de un organismo específico. Se apreciará que la ubicación precisa de una región particular, aminoácido o nucleótido puede variar ligeramente de un PheS o ácido nucleico que codifica PheS a otro, por ejemplo, en diferentes cepas o especies de organismos. Para tener en cuenta esta variación, donde la ubicación de una región específica, nucleótido o aminoácido se menciona en este documento, se describe como "leída en relación con" o "leída con relación a" la posición de aminoácidos de PheS natural de *C. autoethanogenum* (SEQ ID No. 21) (o al Phe-S natural de *C. ljungdahlii* DSM13528 (GenBank: ADK16487.1)) o la posición del nucleótido del ácido nucleico que codifica PheS de *C. autoethanogenum* (SEQ ID 12 en este documento) (o al Phe-S natural de *C. ljungdahlii* DSM13528 (GenBank: ADK16487.1), en donde el aminoácido de la primera posición en la SEQ ID 21 (o ADK16487.1) y el primer nucleótido en la SEQ ID 12 (o ADK16487.1) son la posición 1. Dichas frases deben tomarse de manera amplia y están destinadas a abarcar regiones equivalentes, aminoácidos o nucleótidos en otras proteínas PheS (o ácidos nucleicos que codifican las mismas), aunque puedan estar en una ubicación diferente. Los expertos en la materia a los que se refiere la invención podrán identificar fácilmente la ubicación o posición de una región particular, aminoácido o nucleótido en un PheS particular o ácido nucleico que codifica a través de la alineación de secuencia rutinaria y con la información contenida en este documento.

La referencia a una región particular de un PheS o ácido nucleico "entre" dos aminoácidos o nucleótidos particulares debe tomarse como una región que comprende dichos nucleótidos o aminoácidos. En otras palabras, la región incluye los nucleótidos terminales o aminoácidos referidos. Por ejemplo, un sitio específico de sustrato entre los aminoácidos en la posición 306 y 313 incluye los aminoácidos presentes en las posiciones 306 y 313.

La expresión "análogo de fenilalanina" debe tomarse de manera amplia e incluye un análogo o derivado de fenilalanina que puede incorporarse a péptidos y proteínas en lugar de fenilalanina, lo que resulta en cierta toxicidad para un microorganismo. En una realización, el análogo de fenilalanina se elige entre clorofenilalanina, fluorofenilalanina y bromofenilalanina. En una realización particular, el análogo de fenilalanina se elige entre DL-4-clorofenilalanina y p-clorofenilalanina, p-fluoro-L-fenilalanina, p-fluoro-DL-fenilalanina, p-bromo-L-fenilalanina. Las personas expertas pueden apreciar fácilmente otros análogos de fenilalanina de uso en la invención.

Se puede hacer referencia en este documento a un vector de ácido nucleico que incluya una "secuencia de ácido nucleico de interés" o frases similares. Dichas frases deben tomarse de forma amplia e incluir uno o más nucleótidos, genes, promotores, secuencias reguladoras, otros elementos genéticos y pueden ser codificantes o no codificantes. Puede incluir una secuencia de nucleótidos que esté diseñada para introducir una o más modificaciones genéticas en una o más ubicaciones diana en el genoma del huésped, incluida una o una combinación de una eliminación, adición o sustitución de uno o más nucleótidos. En algunas realizaciones, la secuencia de ácido nucleico o ácido nucleico de interés puede diseñarse para eliminar un gen presente en el genoma del microorganismo progenitor.

Las expresiones "ubicación diana" y "secuencia de ácido nucleico diana", como se usan en este documento, deben tomarse de manera amplia para incluir cualquier sitio, región o secuencia de nucleótidos en un genoma progenitor o huésped donde se desee introducir una o más modificaciones genéticas (incluida la inserción, supresión y/o sustitución de uno o más nucleótidos), e incluye un gen, una región intergénica, un promotor y/o una secuencia reguladora de interés, por ejemplo.

Se puede hacer referencia en el presente documento a un vector de la invención que incluya "dos secuencias de ácido nucleico homólogas a las regiones seleccionadas alrededor de una ubicación diana "dentro del genoma de un microorganismo progenitor". Dichas secuencias de ácido nucleico también pueden denominarse en el presente documento como "brazos de homología".

Se puede hacer referencia en este documento a proteínas (PheS y/o ThiK) o ácidos nucleicos que codifican tales proteínas que son "de" o "derivadas de" un organismo particular. Esto debe tomarse en términos generales para significar que la proteína o el ácido nucleico tiene la secuencia de la proteína relevante o el ácido nucleico que codifica la proteína relevante en ese organismo. No debe entenderse que significa que la proteína o el ácido nucleico se han extraído físicamente de ese organismo. Dichas proteínas y ácidos nucleicos pueden prepararse usando síntesis química y similares, por ejemplo.

Un "microorganismo progenitor" es un microorganismo utilizado para generar un microorganismo recombinante de acuerdo con la invención. En una realización, el microorganismo progenitor puede ser uno que se produzca en la naturaleza (es decir, un microorganismo natural) o uno que se haya modificado previamente (un microorganismo genéticamente modificado o recombinante). De acuerdo con la presente invención, un "microorganismo progenitor" es un *Clostridium* spp.

Las personas capacitadas podrán identificar fácilmente a los microorganismos de uso en la invención de *Clostridium* ssp. Sin embargo, a modo de ejemplo, el grupo puede incluir: *Clostridium autoethanogenum*, *Clostridium ljungdahlii*, *Clostridium ragsdalei*, *Clostridium carboxidivorans*, *Clostridium drakei*, *Clostridium scatologenes*, *Clostridium*

5 aceticum, Clostridium formicoaceticum, Clostridium magnum, Clostridium coskatii, Clostridium acetobutylicum, Clostridium beijerinckii, Clostridium sacharoperbutylaceticum, Clostridium saccharobutylicum, Clostridium thermocellum, Clostridium cellulolyticum, Clostridium phytofermentans, Clostridium pasterianum, Clostridium kluveri, Clostridium difficile, Clostridium botulinum, Clostridium sporogenes, Clostridium perfringens, Clostridium acetobutylicum, Clostridium acidisoli, Clostridium aciditolerans, Clostridium acidurici, Clostridium aerotolerans, Clostridium akagii, Clostridium aldenense, Clostridium algidicarnis, Clostridium algidixylanolyticum, Clostridium alkalicellulosi, Clostridium aminovalericum, Clostridium amygdalinum, Clostridium arcticum, Clostridium argentinense, Clostridium aurantibutyricum, Clostridium baratii, Clostridium botulinum, Clostridium bowmanii, Clostridium butyricum, Clostridium beijerinckii, Clostridium cadaveris, Clostridium caminithermale, Clostridium carboxidivorans, Clostridium carnis, Clostridium celatum, Clostridium celerecrescens, Clostridium cellulolyticum, Clostridium cellulosi, Clostridium chartatabidum, Clostridium clostridioforme, Clostridium coccoides, Clostridium cochlearium, Clostridium cocleatum, Clostridium colinum, Peptoclostridium difficile, Clostridium diolis, Clostridium disporicum, Clostridium drakei, Clostridium durum, Clostridium esterteticum, Clostridium fallax, Clostridium felsineum, Clostridium ervidum, Clostridium fimetarium, Clostridium formicaceticum, Clostridium ghonii, Clostridium glycolicum, Clostridium glycyrrhizinilyticum, Clostridium haemolyticum, Clostridium halophilum, Clostridium histolyticum, Clostridium perfringens, Clostridium phytofermentans, Clostridium piliforme, Clostridium polysaccharolyticum, Clostridium populeti, Clostridium propionicum, Clostridium proteoclasticum, Clostridium proteolyticum, Clostridium psychrophilum, Clostridium puniceum, Clostridium puri, Clostridium putrefaciens, Clostridium putrificum, Clostridium quercicolum, Clostridium quinii, Clostridium ramosum, Clostridium roseum, Clostridium saccharobutylicum, Clostridium saccharolyticum, Clostridium sacharoperbutylaceticum, Clostridium sardiniense, Clostridium stercorarium, Clostridium sticklandii, Clostridium paradoxum, Clostridium paraperfringens, Clostridium paraputrificum, Clostridium pasculi, Clostridium pasteurianum, Clostridium novyi, Clostridium septicum, Clostridium histolyticum, Clostridium hydroxybenzoicum, Clostridium hylemonae, Clostridium innocuum, Clostridium kluveri, Clostridium lactatifermentans, Clostridium lacusfiyxellense, Clostridium laramiense, Clostridium lentocellum, Clostridium lentoputrescens, Clostridium methoxybenzovorans, Clostridium methylpentosum, Clostridium nitrophenolicum, Clostridium novyi, Clostridium oceanicum, Clostridium oroticum, Clostridium oxalicum, Clostridium tertium, Clostridium tetani, Clostridium tetanomorphum, Clostridium thermaceticum, Clostridium thermautotrophicum, Clostridium thermoalcaliphilum, Clostridium thermobutyricum, Clostridium thermocellum, Clostridium thermocopriae, Clostridium thermohydrosulfuricum, Clostridium thermolacticum, Clostridium thermopalmarium, Clostridium thermopapyrolyticum, Clostridium thermosaccharolyticum, Clostridium thermosulfirigenes, Clostridium tyrobutyricum, Clostridium uliginosum, Clostridium ultunense, Clostridium villosum, Clostridium viride, Clostridium xylanolyticum, Clostridium xylanovorans, Clostridium bifermentans, y Clostridium sporogenes.

35 En una realización particular, el organismo progenitor se selecciona de un grupo de Clostridium spp acetogénico. En una realización particular, el microorganismo progenitor se selecciona del grupo de organismos carboxidotróficos acetogénicos que comprenden las especies Clostridium autoethanogenum, Clostridium ljungdahlii, Clostridium ragsdalei, Clostridium carboxidivorans, Clostridium drakei, Clostridium scatologenes, Clostridium aceticum, Clostridium formicoaceticum, Clostridium magnum, y Clostridium coskatii.

40 En una realización, el microorganismo progenitor se selecciona de un grupo de clostridios carboxidotróficos que comprenden la especie C. autoethanogenum, C. ljungdahlii y "C. ragsdalei" y aislados relacionados. Estos incluyen, entre otros, las cepas C. autoethanogenum JAI-1<sup>T</sup> (DSM10061) (Abrini, Naveau y Nyns, 1994), C. autoethanogenum LBS1560 (DSM19630) (documento WO/2009/064200), C. autoethanogenum LBS1561 (DSM23693), C. ljungdahlii PETC<sup>T</sup> (DSM13528 = ATCC 55383) (Tanner, Miller y Yang, 1993), C. ljungdahlii ERI-2 (ATCC 55380) (patente de EE. UU. 5.593.886), C. ljungdahlii C-01 (ATCC 55988) (patente de EE. UU. 6.368.819), C. ljungdahlii O-52 (ATCC 55989) (patente de EE. UU. 6.368.819), o "C. ragsdalei P11<sup>T</sup>" (ATCC BAA-622) (documento WO 2008/028055), y aislados relacionados como "C. coskatii" (patente de EE.UU. 2011/0229947), y sus cepas mutantes como C. ljungdahlii OTA-1 (Tirado-Acevedo O. "Production of Bioethanol from Synthesis Gas Using Clostridium ljungdahlii". Tesis doctoral, North Carolina State University, 2010).

50 Estas cepas forman un subgrupo dentro del clúster I de ARNr clostridial (Collins et al., 1994), con al menos un 99% de identidad en el nivel del gen 16S ARNr, aunque son especies distintas según lo determinado por la reasociación de ADN-ADN y los experimentos de huellas dactilares de ADN (documento WO 2008/028055, patente de EE.UU. 2011/0229947).

Las cepas de este grupo se definen por características comunes, que tienen un genotipo y un fenotipo similares, y todas comparten el mismo modo de conservación de energía y metabolismo fermentativo. Las cepas de este grupo carecen de citocromos y conservan energía a través de un complejo Rnf.

55 Todas las cepas de este grupo tienen un tamaño de genoma de aproximadamente 4,2 MBp (Köpke et al., 2010) y una composición de GC de alrededor del 32% en moles (Abrini et al., 1994; Köpke et al., 2010; Tanner et al., 1993) (documento WO 2008/028055; patente de EE.UU. 2011/0229947), y se conservaron los operones de genes clave esenciales que codificaban para las enzimas de la ruta Wood-Ljungdahl (monóxido de carbono deshidrogenasa, formil-tetrahidrofolato sintetasa, metilen-tetrahidrofolato deshidrogenasa, formil-tetrahidrofolato-ciclohidrolasa, metilen-tetrahidrofolato reductasa, formil-tetrahidrofolato ciclohidrolasa, metilentetrahidrofolato reductasa, y monóxido de carbono deshidrogenasa/acetil-CoA sintasa), hidrogenasa, formato deshidrogenasa, complejo Rnf (rnfCDGEAB), piruvato:ferredoxina oxidoreductasa, aldehído:ferredoxina oxidoreductasa (Köpke et al., 2010, 2011).

Se ha encontrado que la organización y el número de genes de la ruta Wood-Ljungdahl, responsables de la captación de gas, son iguales en todas las especies, a pesar de las diferencias en las secuencias de aminoácidos y ácidos nucleicos (Köpke et al., 2011).

5 Las cepas del grupo tienen una morfología y tamaño similares (las células de crecimiento logarítmico están entre 0,5-0,7 x 3-5 µm), son mesofílicas (temperatura de crecimiento óptima entre 30-37°C) y estrictamente anaerobias (Abrini et al, 1994; Tanner et al, 1993) (documento WO 2008/028055). Además, todas comparten las mismas características filogenéticas principales, como el mismo rango de pH (pH 4-7,5, con un pH inicial óptimo de 5,5-6), un fuerte crecimiento autótrofo en gases que contienen CO con tasas de crecimiento similares y un perfil metabólico con etanol y ácido acético como principal producto final de la fermentación, con pequeñas cantidades de 2,3-  
10 butanodiol y ácido láctico formado bajo ciertas condiciones (Abrini et al., 1994; Köpke et al., 2011; Tanner et al., 1993) (documento WO 2008/028055). Se ha observado la producción de indol con todas las especies. Sin embargo, las especies se diferencian en la utilización del sustrato de diversos azúcares (por ejemplo, ramnosa, arabinosa), ácidos (por ejemplo, gluconato, citrato), aminoácidos (por ejemplo, arginina, histidina) u otros sustratos (por ejemplo, betaína, butanol). Se encontró que algunas de las especies eran auxotróficas de ciertas vitaminas (por ejemplo, tiamina, biotina), mientras que otras no lo eran. La reducción de los ácidos carboxílicos en sus correspondientes  
15 alcoholes se ha demostrado en una variedad de estos organismos (Pérez, Richter, Loftus y Angenent, 2012).

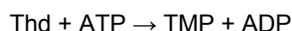
Por lo tanto, los rasgos descritos no son específicos de un organismo como *C. autoethanogenum* o *C. ljungdahlii*, sino que son rasgos generales para la Clostridia carboxidotrófica que sintetiza etanol. Se puede anticipar que la invención funcionará no solo en estas cepas, sino en todas las especies de Clostridia, aunque puede haber  
20 diferencias en el comportamiento.

En realizaciones particulares, el microorganismo progenitor se selecciona del grupo que comprende *Clostridium autoethanogenum*, *Clostridium ljungdahlii* y *Clostridium ragsdalei*. En una realización, el grupo también comprende *Clostridium coskatii*. En una realización particular, el microorganismo progenitor es *Clostridium autoethanogenum* DSM23693.

25 Un microorganismo progenitor puede o no contener ácidos nucleicos que codifican fenilalanina tRNA sintetasa o expresar fenilalanina tRNA sintetasa.

A lo largo de esta memoria descriptiva, se proporciona información de secuencia ejemplar para PheS, proteínas/péptidos PheS y ThiK alterados y ácidos nucleicos que codifican los mismos.

A lo largo de esta memoria descriptiva, se proporciona una información ilustrativa de las secuencias para identificar  
30 proteínas/péptidos y ácidos nucleicos a modo de ejemplo aplicables a la invención y para permitir que cualquier experto en la materia ponga en práctica las realizaciones específicas de la invención sin una experimentación excesiva. Debe apreciarse que las secuencias de ácidos nucleicos y aminoácidos pueden diferir de un microorganismo a otro. Por consiguiente, la invención no debe considerarse limitada a estas realizaciones específicas, sino que se extiende a proteínas/péptidos y ácidos nucleicos que tienen secuencias diferentes pero que  
35 son sustancialmente capaces de realizar la misma función. Para PheS, la función deseada (como una subunidad de fenilalanina ARNt sintetasa) es la aminoacilación de ARNt<sup>Phe</sup> con fenilalanina. Para la PheS, la función deseada (como una subunidad de fenilalanina ARNt sintetasa) es la aminoacilación del ARNt<sup>Phe</sup> con fenilalanina. Para ThiK, la función deseada es catalizar la reacción:



40 donde Thd es desoxitimidina, ATP es adenosina 5'-trifosfato, TMP es desoxitimidina 5'-fosfato y ADP es adenosina 5'-difosfato.

Típicamente, tales proteínas/péptidos alternativos o distintos tendrán al menos aproximadamente un 75% de similitud de secuencia de aminoácidos con una PheS (que incluye una PheS alterada) o una proteína ThiK ejemplificada en este documento. En realizaciones particulares, tales proteínas alternativas tendrán al menos  
45 aproximadamente 80%, 85%, 90%, 95% o 99% de similitud de secuencia con un PheS (incluyendo un PheS alterado) o el ThiK ejemplificado en este documento. En realizaciones particulares, tales proteínas alternativas tendrán al menos aproximadamente el 75%, 80%, 85% o, 90%, 95% o 99% de identidad de secuencia con un PheS (incluido un PheS alterado) o el ThiK ejemplificado en el presente documento. A nivel de ácido nucleico, los genes que codifican tales proteínas alternativas o variantes tendrán típicamente al menos aproximadamente un 75% de homología de secuencia con un ácido nucleico que codifique un PheS (incluido un PheS alterado) o el ThiK  
50 ejemplificado en el presente documento. En realizaciones particulares, tales ácidos nucleicos distintos o alternativos tendrán al menos aproximadamente 80%, 85%, 90%, 95% o 99% de homología de secuencia con un ácido nucleico que codifica un PheS (incluido un PheS alterado) o el ThiK ejemplificado en este documento. En una realización particular, dichos ácidos nucleicos tendrán al menos aproximadamente 75%, 80%, 85%, 90%, 95% o 99% de  
55 identidad de secuencia con un ácido nucleico que codifique un PheS (incluido un PheS alterado) o el ThiK ejemplificado en este documento. Los ácidos nucleicos o proteínas/péptidos alternativos o distintos como se describen pueden denominarse en el presente documento "variantes funcionalmente equivalentes".

También debe tenerse en cuenta que la variante funcionalmente equivalente de PheS, PheS alterada o ThiK no

necesita tener el mismo nivel de actividad que una proteína/péptido del cual es una variante. Todo lo que se requiere es que se retenga algún nivel de la actividad deseada. Los expertos en la materia conocerán los ensayos de uso para evaluar la actividad de PheS, PheS o ThiK alterados. Sin embargo, a modo de ejemplo: la función o actividad de PheS puede probarse utilizando métodos que miden la aminoacilación. Los autores utilizaron velocidades de aminoacilación y parámetros cinéticos de pheS para probar las variaciones de actividad de pheS en la utilización de fenilalanina (Kast et al, 1991, J. Mol. Biol 222: 99-124). La función o actividad de una PheS alterada puede realizarse observando el crecimiento en presencia de un análogo tóxico de la fenilalanina utilizando métodos conocidos para cultivar o hacer crecer microorganismos (Kast et al, 1991, J. Mol. Biol 222: 99-124). La función o actividad de ThiK se puede probar utilizando un ensayo de actividad como lo describen Brockenbrough et al. (Nucl Med Biol. 2007, 34 (6): 619-23) y Jonsson & McIvor (Anal Biochem. 1991, 199 (2): 232-7) o con kits ELISA disponibles en el mercado, como por ejemplo de BioVendor (Cat. 901/902).

La referencia a "transformar" un microorganismo progenitor debe tomarse de manera amplia para incluir cualquier medio de transferencia o introducción de un ácido nucleico exógeno en un microorganismo que se conoce en la técnica. A modo de ejemplo, "transformar" incluye, entre otros, transfección, transducción, conjugación y electroporación.

Aspectos y realizaciones de la invención.

La invención proporciona el uso de ThiK y/o PheS como marcador de contra-selección en un método para producir un microorganismo recombinante a partir de un microorganismo progenitor. También proporciona ácido(s) nucleico(s) que codifican ThiK y/o PheS, vectores de ácido nucleico que comprenden dicho(s) ácido(s) nucleico(s) y el uso de dicho(s) ácido(s) nucleico(s) y/o plásmidos para producir un microorganismo recombinante a partir de un microorganismo progenitor. Además, la invención proporciona un método para producir un microorganismo recombinante a partir de un microorganismo progenitor. De acuerdo con la invención, el microorganismo progenitor es un *Clostridium* spp., como se describe en este documento anteriormente, y la PheS incluye al menos una alteración en comparación con una PheS natural tal que, en uso, la fenilalanina tRNA sintetasa es capaz de aminoacilar el ARNt usando un análogo de fenilalanina.

PheS

PheS es la subunidad alfa de la subunidad de la proteína fenilalanina tRNA sintetasa. La fenilalanina tRNA sintetasa es responsable de la aminoacilación del tRNA<sup>Phe</sup> con fenilalanina, que es crítica para la producción de proteínas en una célula. La enzima cataliza la acetilación de fenilalanina a su ARNt análogo. El ARNt<sup>Phe</sup> resultante es enviado al ribosoma por factores de alargamiento que luego se unen a su anti-codón presente en el ARNm. Una vez unido, el aminoácido se une covalentemente a su aminoácido precedente, aumentando así la cadena peptídica.

Un PheS de la invención es uno que incluye al menos una alteración en comparación con un PheS natural, de modo que en el uso de fenilalanina tRNA sintetasa sea capaz de aminoacilar ARNt utilizando análogos de fenilalanina. La incorporación de análogos de fenilalanina en proteínas celulares da como resultado proteínas inestables o no funcionales. Por lo tanto, cualquier célula que incluya la PheS alterada normalmente no podrá sobrevivir.

La PheS natural en la que se basa la PheS alterada puede ser de cualquier fuente apropiada, incluyendo cualquier número de diferentes organismos auto-replicantes, como plantas, animales, hongos y microorganismos. En una realización particular, la PheS es de un microorganismo. En una realización, la PheS es de un *Clostridium* spp. o es una variante funcionalmente equivalente de la misma. Sólo a modo de ejemplo, el *Clostridium* spp. puede incluir:

*Clostridium autoethanogenum*, *Clostridium ljungdahlii*, *Clostridium ragsdalei*, *Clostridium carboxidivorans*, *Clostridium drakei*, *Clostridium scatologenes*, *Clostridium aceticum*, *Clostridium formicoaceticum*, *Clostridium magnum*, *Clostridium coskatii*, *Clostridium acetobutylicum*, *Clostridium beijerinckii*, *Clostridium sacharoperbutylaceticum*, *Clostridium saccharobutylicum*, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium pasterianum*, *Clostridium kluyveri*, *Clostridium difficile*, *Clostridium botulinum*, *Clostridium sporogenes*, *Clostridium perfringens*, *Clostridium acetobutylicum*, *Clostridium acidisoli*, *Clostridium aciditolerans*, *Clostridium acidurici*, *Clostridium aerotolerans*, *Clostridium akagii*, *Clostridium aldenense*, *Clostridium algidicarnis*, *Clostridium algidixylanolyticum*, *Clostridium alkallicellulosi*, *Clostridium aminovalericum*, *Clostridium amygdalinum*, *Clostridium arcticum*, *Clostridium argentinense*, *Clostridium aurantibutyricum*, *Clostridium baratii*, *Clostridium botulinum*, *Clostridium bowmanii*, *Clostridium butyricum*, *Clostridium beijerinckii*, *Clostridium cadaveris*, *Clostridium caminithermale*, *Clostridium carboxidivorans*, *Clostridium carnis*, *Clostridium celatum*, *Clostridium celerecrescens*, *Clostridium cellulolyticum*, *Clostridium cellulosi*, *Clostridium chartatabidum*, *Clostridium clostridioforme*, *Clostridium coccoides*, *Clostridium cochlearium*, *Clostridium cocleatum*, *Clostridium colinum*, *Clostridium difficile*, *Clostridium diolis*, *Clostridium disporicum*, *Clostridium drakei*, *Clostridium durum*, *Clostridium esterteticum*, *Clostridium fallax*, *Clostridium felsineum*, *Clostridium ervidum*, *Clostridium fimetarium*, *Clostridium formicaceticum*, *Clostridium ghonii*, *Clostridium glycolicum*, *Clostridium glycyrrhizinilyticum*, *Clostridium haemolyticum*, *Clostridium halophilum*, *Clostridium tetani*, *Clostridium perfringens*, *Clostridium phytofermentans*, *Clostridium piliforme*, *Clostridium polysaccharolyticum*, *Clostridium populeti*, *Clostridium propionicum*, *Clostridium proteoclasticum*, *Clostridium proteolyticum*, *Clostridium psychrophilum*, *Clostridium puniceum*, *Clostridium puri*, *Clostridium putrefaciens*, *Clostridium putrificum*, *Clostridium quercicolum*, *Clostridium quinii*, *Clostridium ramosum*, *Clostridium roseum*,

5 Clostridium saccharobutylicum, Clostridium saccharolyticum, Clostridium saccharoperbutylacetonicum, Clostridium sardiniense, Clostridium stercorarium, Clostridium sticklandii, Clostridium paradoxum, Clostridium parapaperfringens, Clostridium paraputrificum, Clostridium pascui, Clostridium pasteurianum, Clostridium novyi, Clostridium septicum, Clostridium histolyticum, Clostridium hydroxybenzoicum, Clostridium hylemonae, Clostridium innocuum, Clostridium kluveri, Clostridium lactatifermentans, Clostridium lacusfixellense, Clostridium laramiense, Clostridium lentocellum, Clostridium lentoputrescens, Clostridium methoxybenzovorans, Clostridium methylpentosum, Clostridium nitrophenolicum, Clostridium novyi, Clostridium oceanicum, Clostridium oroticum, Clostridium oxalicum, Clostridium tertium, Clostridium tetani, Clostridium tetanomorphum, Clostridium thermaceticum, Clostridium thermautotrophicum, Clostridium thermoalcaliphilum, Clostridium thermobutyricum, Clostridium thermocellum, Clostridium thermocoprae, Clostridium thermohydrosulfuricum, Clostridium thermolacticum, Clostridium thermopalmarium, Clostridium thermopapyrolyticum, Clostridium thermosaccharolyticum, Clostridium thermosulfirigenes, Clostridium tyrobutyricum, Clostridium uliginosum, Clostridium ultunense, Clostridium villosum, Clostridium viride, Clostridium xylanolyticum, Clostridium xylanovorans, Clostridium bifermentans, y Clostridium sporogenes.

15 En ciertas realizaciones, el PheS es de un microorganismo seleccionado del grupo de Clostridium spp. o es su variante funcionalmente equivalente. En una realización, el PheS es de un microorganismo seleccionado de un grupo de Clostridium spp. acetogénico o es su variante funcionalmente equivalente. En una realización particular, PheS es de un microorganismo seleccionado del grupo de organismos carboxidotrópicos acetogénicos que comprenden las especies Clostridium autoethanogenum, Clostridium ljungdahlii, Clostridium ragsdalei, Clostridium carboxidivorans, Clostridium drakei, Clostridium scatologenes, Clostridium aceticum, Clostridium formicoaceticum, Clostridium magnum, y Clostridium coskatii, o es una variante funcionalmente equivalente de uno o más de los mismos.

25 En una realización, PheS se selecciona de un grupo de clostridios carboxidotrópicos que comprenden las especies C. autoethanogenum, C. ljungdahlii y "C. ragsdalei" y aislados relacionados, que se han descrito anteriormente en este documento. Solo a modo de ejemplo, las proteínas PheS naturales apropiadas (y las correspondientes secuencias de ácido nucleico) incluyen aquellas descritas en las bases de datos públicas, como GenBank, como se indica a continuación: PheS de Clostridium ljungdahlii DSM13528 (GenBank ADK16487.1); Clostridium carboxidivorans P7 (GenBank: EET86555.1); subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium kluveri] WP\_012620882.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium perfringens cepa 13] GenBank: BAB81592.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium botulinum A cepa ATCC 3502] GenBank: YP\_001255621.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium sporogenes] GenBank: WP\_003495653.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium beijerinckii NCIMB 8052] GenBank: YP\_001308703.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium acetobutylicum ATCC 824] GenBank: NP\_348973.1; subunidad alfa de fenilalanil-ARNt sintetasa [Clostridium thermocellum ATCC 27405] GenBank: YP\_001036648.1. En una realización, la PheS es de C. autoethanogenum y tiene la secuencia de aminoácidos de la SEQ ID No.21. Las variantes funcionalmente equivalentes de estas proteínas ejemplares también pueden ser útiles.

35 En una realización, la al menos una alteración en PheS en comparación con una PheS natural está localizada dentro de una región que comprende el sitio de especificidad de sustrato. En una realización, el sitio de especificidad del sustrato está ubicado entre los aminoácidos 306 y 313, leído en relación con el PheS natural de C. autoethanogenum SEQ ID No. 21 (o con el Phe-S natural de C. ljungdahlii DSM13528 (GenBank: ADK16487.1)).

40 Debe apreciarse que la ubicación precisa del sitio de especificidad del sustrato puede variar de una proteína PheS particular a otra. De acuerdo con esto, la invención debe tomarse para incluir proteínas PheS que se han alterado fuera del sitio definido por los aminoácidos 306 a 313 anteriores, y confiere a la fenilalanina ARNt sintetasa la capacidad de aminoacilar ARNt utilizando análogos de fenilalanina. Los expertos en la materia a los que se refiere la invención podrán identificar fácilmente el sitio apropiado basándose en la alineación de la secuencia de aminoácidos con la de C. autoethanogenum SEQ ID 21 (o con el Phe-S natural de C. ljungdahlii DSM13528 (GenBank ADK16487.1), descrito anteriormente. Sin embargo, a modo de ejemplo, en PheS en E. coli, el sitio de especificidad del sustrato se define por los aminoácidos en las posiciones 289 a 296.

45 En una realización, la al menos una alteración es una o más sustituciones, adición y/o eliminación de aminoácidos. En una realización, la al menos una alteración es una sustitución de aminoácido en la posición 311, leída con relación a la secuencia de aminoácidos de Clostridium autoethanogenum PheS (SEQ ID 21). En una realización, la al menos una alteración es la sustitución de Cys por Gly en el aminoácido 311. En otras realizaciones, la alternancia es una o una combinación de sustituciones de aminoácidos en las posiciones 310 y 312, leídas en relación con la secuencia de aminoácidos de PheS de Clostridium autoethanogenum (SEQ ID 21).

50 En una realización, la PheS alterada comprende o consiste en la secuencia de aminoácidos de la SEQ ID N° 53.

55 La invención también se refiere a ácidos nucleicos que codifican una PheS alterada de la invención. En una realización, la al menos una alteración está localizada dentro de una región del ácido nucleico que codifica el sitio de especificidad del sustrato, como se menciona en este documento anteriormente.

En una realización, la región que codifica el sitio de especificidad del sustrato está ubicada entre las bases 918 y 939, leída en relación con el ácido nucleico que codifica PheS en C. autoethanogenum (SEQ ID 12 en este

documento) (o al Phe-S natural de *C. ljungdahlii* DSM13528 (GenBank: ADK16487.1). Sin embargo, el sitio puede diferir de un ácido nucleico a otro (por ejemplo, ácidos nucleicos de diferentes especies u organismos), para reflejar la ubicación precisa del sitio de especificidad del sustrato en una proteína PheS natural, como se describe anteriormente en este documento. Las personas capacitadas podrán identificar fácilmente la región apropiada en ácidos nucleicos alternativos a través de alineamientos de secuencia estándar.

En una realización, la al menos una alteración es una o más sustituciones, adiciones y/o eliminaciones de nucleótidos.

En una realización, la al menos una alteración es una sustitución de nucleótido en la base 932 leída con relación al gen de *Clostridium autoethanogenum* que codifica PheS (SEQ ID 12). En una realización, la al menos una alteración es la sustitución de C por G en la base 932. En una realización, el ácido nucleico comprende o consiste en la secuencia de la SEQ ID N° 14.

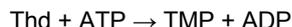
Los ácidos nucleicos que codifican un PheS alterado de acuerdo con la invención pueden generarse utilizando cualquier número de métodos conocidos en la técnica, basándose en la información del presente documento, la secuencia de aminoácidos (y/o la secuencia de ácido nucleico que codifica la secuencia de aminoácidos) de las proteínas PheS naturales ilustrativas y el código genético, por ejemplo. Sin embargo, a modo de ejemplo, pueden producirse mediante síntesis química o mediante técnicas recombinantes estándar.

A modo de ejemplo, los ácidos nucleicos ejemplares que codifican PheS naturales se proporcionan en el presente documento y en las bases de datos disponibles públicamente, tales como GenBank de la siguiente manera: *E. coli* K12 (NC\_000913.2), Gene ID: 946223, EcoGene: EG10709; *Clostridium ljungdahlii* DSM 13528, GenBank: CP001666.1, GI: 300433347.

Debe apreciarse que un ácido nucleico que codifica una PheS alterada puede ser un codón optimizado para la *Clostridium* spp. particular en la que ha de expresarse. Esto se puede lograr utilizando técnicas de optimización de codones estándar.

ThiK

ThiK es una proteína que funciona para catalizar la reacción:



donde Thd es desoxitimidina, ATP es adenosina 5'-trifosfato, TMP es desoxitimidina 5'-fosfato y ADP es adenosina 5'-difosfato. HSV-TK, por ejemplo, cataliza la fosforilación de la desoxitimidina.

La ThiK de uso en la invención se puede derivar de cualquier organismo apropiado. Sin embargo, a modo de ejemplo, ThiK puede provenir del Virus del Herpes Simple 1 o Virus del Herpes Simple 2 (HSV-TK), VZV, CMV, HHV7, HHV7, HHV8, EBV. Alternativamente, podrían usarse variantes funcionalmente equivalentes de ThiK de HSV-TK.

Solo a modo de ejemplo, las proteínas ThiK incluyen aquellas descritas en las bases de datos públicas como el GenBank de la siguiente manera: AB009254.2. Las variantes funcionalmente equivalentes de esta proteína ejemplar también pueden ser útiles.

En una realización, la ThiK comprende la secuencia de aminoácidos de la SEQ ID No. 54.

La invención también se refiere a ácidos nucleicos que codifican un ThiK. Las personas expertas apreciarán fácilmente la secuencia de nucleótidos apropiada de los ácidos nucleicos que codifican ThiK, teniendo en cuenta la secuencia de aminoácidos de las proteínas ThiK ejemplares proporcionadas en el presente documento, y el código genético. Sin embargo, a modo de ejemplo, se proporcionan ácidos nucleicos ejemplares que codifican ThiK en las bases de datos públicas tales como GenBank AB009254.2, JQ895546.1, AY575235.1, AF243479.1, AY575236.2, HQ123159.1

Sin embargo, a modo de ejemplo, los ácidos nucleicos pueden tener una secuencia de nucleótidos de SEQ ID 19 o SEQ ID 22 como se describe en el presente documento.

En una realización, el ácido nucleico que codifica ThiK tiene la secuencia de nucleótidos de la SEQ ID N° 19.

Los ácidos nucleicos que codifican un ThiK de acuerdo con la invención pueden generarse utilizando cualquier número de métodos conocidos en la técnica, basándose en la información del presente documento, la secuencia de aminoácidos (y/o la secuencia de ácidos nucleicos que codifican la secuencia de aminoácidos) de las proteínas ThiK ejemplares, y el código genético, por ejemplo. Sin embargo, a modo de ejemplo, pueden producirse mediante síntesis química o mediante técnicas recombinantes estándar.

Debe apreciarse que un ácido nucleico que codifica ThiK puede ser un codón optimizado para el *Clostridium* spp. particular en el que se va a expresar. Esto se puede lograr utilizando técnicas de optimización de codones estándar.

## Vectores de ácidos nucleicos

La invención también proporciona un vector de ácido nucleico que comprende un ácido nucleico que codifica ThiK y/o un PheS alterado de acuerdo con la invención. El vector puede ser de cualquier naturaleza o natural, como entenderán los expertos en la materia a los que se refiere la invención, incluidos, por ejemplo, aquellos adecuados para la clonación, la expresión y la transformación.

En una realización, el vector de ácido nucleico es uno adecuado para generar un microorganismo recombinante de la invención. En esta realización, el vector de ácido nucleico es un plásmido que comprende al menos un ácido nucleico que codifica un PheS alterado y/o un ThiK como se describe en este documento anteriormente. En una realización particular, el vector comprende además al menos:

(a) al menos una secuencia de ácido nucleico que codifica al menos un marcador de selección positiva;

(b) dos secuencias de ácido nucleico homólogas a las regiones seleccionadas alrededor de una ubicación diana o secuencia de ácido nucleico dentro del genoma de un microorganismo progenitor, que permite la recombinación del plásmido con el genoma del microorganismo progenitor.

En una realización, el vector comprende además al menos un ácido nucleico de interés que se desea insertar o integrar en el genoma de un microorganismo progenitor.

En una realización, el ácido nucleico que codifica el marcador de selección positivo se coloca en el vector plasmídico fuera de los brazos de homología. En otra realización, el ácido nucleico que codifica el marcador de selección positiva está ubicado entre los brazos de homología.

Cuando el vector se va a usar para producir un microorganismo recombinante de acuerdo con la invención, se adaptará en uso para permitir la expresión de los ácidos nucleicos que codifican uno o más marcadores de selección. Por consiguiente, también incluirá al menos un promotor capaz de dirigir la expresión de los marcadores de selección contenidos en el plásmido. El al menos un promotor puede comprender una parte de la al menos una secuencia de ácido nucleico que codifica al menos un marcador de contra-selección o una parte de la al menos una secuencia de ácido nucleico que codifica al menos un marcador de selección positiva, o puede ser un ácido nucleico contenido dentro del plásmido, que está separado del(de los) ácido(s) nucleico(s) que codifican el uno o más marcadores de selección mediante nucleótidos intermedios. En una realización, el promotor puede ser inducible. En otra realización, el promotor es constitutivo.

Las personas expertas apreciarán fácilmente los promotores de uso en un plásmido de la invención. Sin embargo, a modo de ejemplo, estos pueden incluir el promotor Ppta-ack (descrito en el presente documento en la sección de Ejemplos), el promotor lac, el sistema ara, tet o T7.

En una realización particular, el plásmido incluye un promotor fuerte que es capaz de dirigir la expresión del marcador(s) de selección. En una realización, el plásmido incluye un promotor fuerte para dirigir la expresión de al menos un ácido nucleico que codifica un marcador de contra-selección. Este es particularmente el caso en el que se utiliza un PheS alterado para la contra-selección y el genoma del huésped incluye un ácido nucleico que codifica PheS. Idealmente, el promotor fuerte será suficiente para dirigir la expresión de la PheS alterada en al menos el mismo nivel, y preferiblemente en un nivel aumentado, en comparación con la expresión de un ácido nucleico que codifica PheS que está presente en el genoma del huésped. Alternativamente, podrían incluirse uno o más elementos reguladores, como un operador y/o potenciador, en el plásmido además de un promotor, para aumentar la expresión de uno o más marcadores de selección. Los ejemplos de promotores fuertes de uso en la invención incluyen, por ejemplo, el promotor T3, el promotor T7, el promotor PARNr, el promotor Ptrc, o aquellos ejemplificados en la sección de Ejemplos a continuación.

El al menos un marcador de selección positiva puede elegirse entre cualquier número de marcadores de selección positiva conocidos que serán fácilmente apreciados por los expertos en el área de la tecnología a la que se refiere la invención. Sin embargo, a modo de ejemplo, se podría usar CatP (cloranfenicol acetiltransferasa), ErmB o TetA [Heap et al., 2009, J Microbial Methods; Jul; 78 (1); 78-85]. Las personas expertas apreciarán fácilmente la secuencia de nucleótidos para los ácidos nucleicos que codifican estos marcadores de selección positiva, según la información publicada y el código genético. Sin embargo, a modo de ejemplo, GenBank: WP\_002570989 (CatP), YP\_007078965 (ErmB), NP\_957551.1 (TetA); CatP (SEQ ID 23); ErmB (SEQ ID 24); TetA (SEQ ID 25).

Los brazos de homología permiten la recombinación homóloga del vector con el genoma del huésped. Si bien puede preferirse que los brazos tengan una complementariedad del 100% con la región en el genoma al que se dirigen, esto no es necesario, siempre que la secuencia sea lo suficientemente complementaria para permitir la recombinación dirigida con la región genética de interés. Típicamente, los brazos tendrán un nivel de homología que permitirá la hibridación con una región objetivo en condiciones rigurosas, como se define en Sambrook et al (Molecular Cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Como apreciarán los expertos en la técnica, los brazos de homología pueden diseñarse para hibridarse a secuencias de ácido nucleico dentro del genoma que están adyacentes entre sí o separadas entre sí por uno o más nucleótidos.

Las personas capacitadas podrán diseñar fácilmente brazos de homología suficientes para permitir la recombinación homóloga dirigida teniendo en cuenta la información de secuencia disponible públicamente para un microorganismo progenitor determinado. Se proporciona información ilustrativa en la sección de ejemplos que figura a continuación.

5 Un plásmido también puede comprender uno o más elementos adicionales que incluyen uno o más elementos reguladores, uno o más orígenes de replicación, uno o más sitios de multiclonación, entre otros elementos. En una realización particular, los plásmidos se adaptan para permitir la interrupción de un gen nativo de (o al menos ya presente en) un microorganismo progenitor, por ejemplo. En otra realización, los plásmidos se adaptan para permitir la integración y expresión de uno o más genes codificados por el plásmido. Los plásmidos pueden estar en forma de ácidos nucleicos desnudos, así como ácidos nucleicos formulados con uno o más agentes para facilitar el suministro a una célula (por ejemplo, un ácido nucleico conjugado con un liposoma, un organismo en el que está contenido el ácido nucleico).

15 Como se describe en este documento anteriormente, la fenilalanina ARNt sintetasa está formada por dos subunidades, de las cuales una es PheS. La segunda subunidad puede estar presente en el genoma de un microorganismo progenitor a transformar. Por consiguiente, en presencia de un vector de la invención, un microorganismo es capaz de producir fenilalanina tRNA sintetasa, aunque esté alterado y sea capaz de aminoacilar el ARNt usando análogos de fenilalanina. Además, aunque la PheS alterada pueda basarse en la PheS de cualquier organismo, como se describe anteriormente en este documento, en una realización preferida debe ser una que sea compatible con la otra subunidad expresada por el microorganismo progenitor. Si las subunidades no son compatibles, no formarán una enzima funcional. Las personas capacitadas podrán identificar fácilmente si las subunidades de tRNA sintetasa de fenilalanina son o no compatibles usando ensayos estándar para probar la actividad y la función de la enzima, como se describe en este documento anteriormente. Sin embargo, los inventores contemplan que un PheS de cualquier *Clostridium* spp. será compatible con una subunidad PheT de cualquier otra *Clostridium* spp. En una realización particular, la subunidad PheS es de la misma especie de *Clostridia* que el microorganismo progenitor. En una realización particular, la PheS alterada se basa en la PheS de la fenilalanina tRNA sintetasa expresada por el microorganismo progenitor a transformar. En otra realización, el plásmido de la invención también puede incluir un ácido nucleico que codifica una subunidad PheT, que junto con la PheS alterada, forma una fenilalanina tRNA sintetasa activa, aunque una alterada capaz de aminoacilar el ARNt usando análogos de fenilalanina. Esto puede ser útil, por ejemplo, cuando el microorganismo progenitor no contenga un ácido nucleico que codifique la subunidad PheT.

20 Un plásmido puede ser replicante o no replicante.

Los vectores de ácido nucleico de uso en la invención pueden construirse usando cualquier número de técnicas estándar en la técnica. Por ejemplo, pueden usarse síntesis química o técnicas recombinantes. Tales técnicas se describen, por ejemplo, en Sambrook et al (*Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Técnicas ejemplares adicionales se describen en la sección de Ejemplos en este documento a continuación. Esencialmente, los ácidos nucleicos individuales, incluido un ácido nucleico que codifica un marcador de preselección, un ácido nucleico que codifica un marcador de selección positiva, brazos de homología, un ácido nucleico de interés y, opcionalmente, otros ácidos nucleicos se unirán operativamente entre sí para que puedan realizar su función deseada.

35 Cualquiera de los numerosos vectores de plásmidos conocidos en la técnica puede ser adecuado para su uso en la invención. Sin embargo, a modo de ejemplo, los vectores de la serie PMTL80000 serían adecuados. Se proporcionan ejemplos específicos en la sección de ejemplos que se incluye a continuación.

40 En ciertas realizaciones de la invención, un vector de ácido nucleico es uno adecuado para generar o clonar un ácido nucleico que codifique un ThiK o un PheS alterado de la invención. En este caso, el vector no necesita ser adaptado para expresar el ThiK o un PheS alterado. Se puede usar cualquier número de vectores de ácido nucleico conocidos, incluyendo plásmidos y vectores virales. Dichos vectores pueden incluir uno o más elementos reguladores, un origen de replicación, un sitio de multiclonación y/o un marcador seleccionable, entre otros elementos, sitios y marcadores, como conocerán los expertos en la técnica.

#### Células

45 La invención también proporciona una célula que comprende un ácido nucleico de la invención, un vector, una PheS y/o ThiK de acuerdo con la invención. La célula puede ser de cualquier origen y puede incluir aquellas de uso en la clonación o preparación de un vector de acuerdo con la invención. En una realización, la célula es *E. coli* o *Clostridium* spp.

#### Métodos

50 Como se describe en este documento anteriormente, la invención proporciona un método para la producción de un microorganismo recombinante a partir de un microorganismo progenitor. El método generalmente comprende al menos las etapas de: transformar un microorganismo progenitor con un plásmido como se describe en este documento anteriormente, seleccionar uno o más microorganismos que expresen al menos un marcador de selección positiva y seleccionar uno o más microorganismos que no expresen el al menos un marcador de contra-

selección.

5 Un microorganismo progenitor puede transformarse con un plásmido de la invención usando cualquier número de técnicas conocidas en la técnica para producir microorganismos recombinantes. Solo a modo de ejemplo, la transformación (incluida la transducción o la transfección) se puede lograr mediante electroporación, electrofusión, ultrasonificación, transformación mediada por polietilenglicol, competencia química o natural, transformación de protoplastos, inducción de profagos o conjugación. Las técnicas de transformación adecuadas se describen, por ejemplo, en Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989.

10 Solo a modo de ejemplo, se ha descrito la electroporación de varios acetógenos carboxidotróficos, como *C. ljungdahlii* (Köpke et al. 2010, *Proc. Nat. Acad. Sci. U.S.A.* 107: 13087-92; Leang et al, 2012, *Appl. Environ. Microbiol.*; PCT/NZ2011/000203; documento WO2012/053905) y *C. autoethanogenum* (PCT/NZ2011/000203; documento WO2012/053905) y es un método estándar usado en muchos *Clostridia* como *C. acetobutylicum* (Mermelstein et al., 1992, *Biotechnology*, 10, 190-195), *C. cellulolyticum* (Jennert et al., 2000, *Microbiology*, 146: 3071 - 3080) o *C. thermocellum* (Tyurin et al., 2004, *Appl. Environ. Microbiol.* 70: 883-890).

15 A modo de ejemplo adicional, se ha descrito la electrofusión para el *Clostridium* sp. MT351 acetogénico (Tyurin y Kiriukhin, 2012, *J Biotech*: 1-12).

Otra técnica ejemplar incluye la inducción de profagos descrita para el acetógeno carboxidotrófico como *C. scatologenes* (Prasanna Tamarapu Parthasarathy, 2010, *Development of a Genetic Modification System in Clostridium scatologenes ATCC 25775 for Generation of Mutants*, Masters Project Western Kentucky University).

20 A modo de ejemplo adicional, pueden ser utilizados los métodos de conjugación de Herbert et al., 2003, (*FEMS Microbiol. Lett.* 229: 103-110) y Williams et al, 1990 (*J. Gen. Microbiol.* 136: 819-826).

Debe apreciarse que el plásmido puede suministrarse a un microorganismo progenitor como ácido nucleico desnudo o puede formularse con uno o más agentes para facilitar el procedimiento de transformación (por ejemplo, ácido nucleico conjugado con liposomas, un organismo en el que el ácido nucleico está contenido).

25 En ciertas realizaciones, debido a los sistemas de restricción que son activos en el microorganismo a transformar, es necesario metilar cualquier ácido nucleico (por ejemplo, un plásmido de la invención) para introducirlo en el microorganismo. Esto se puede hacer utilizando una variedad de técnicas, incluidas las que se describen a continuación.

30 A modo de ejemplo, en una realización, un microorganismo recombinante de la invención se produce mediante un método que comprende los siguientes pasos:

- introducir en un microorganismo lanzadera de (i) al menos un plásmido que se introduce en el microorganismo progenitor como se describe en el presente documento y (ii) una construcción/vector de metilación que comprende un gen de metiltransferasa;

- expresar el gen de la metiltransferasa;

35 - aislar el al menos un plásmido del microorganismo lanzadera; e,

- introducir el al menos un plásmido en un microorganismo diana.

En una realización, el gen de la metiltransferasa se expresa de manera constitutiva. En otra realización, se induce la expresión del gen de la metiltransferasa.

40 El microorganismo lanzadera es un microorganismo, preferiblemente un microorganismo de restricción negativa, que facilita la metilación de las secuencias de ácido nucleico que forman un plásmido de la invención. En una realización particular, el microorganismo lanzadera es una restricción negativa de *E. coli*, *Bacillus subtilis* o *Lactococcus lactis*.

El constructo/vector de metilación comprende una secuencia de ácido nucleico que codifica una metiltransferasa.

45 Una vez que el uno o más plásmidos y el constructo/vector de metilación se introducen en el microorganismo lanzadera, se induce el gen de metiltransferasa presente en el constructo/vector de metilación. La inducción puede ser por cualquier sistema promotor adecuado, aunque, en una realización particular de la invención, el constructo/vector de metilación comprende un promotor lac inducible y se induce mediante la adición de lactosa o un análogo de la misma, más preferiblemente isopropil-β-D-tio-galactósido (IPTG). Otros promotores adecuados incluyen el sistema ara, tet o T7. En una realización adicional de la invención, el constructor de metilación/promotor de vector es un promotor constitutivo.

50 En una realización particular, el constructo/vector de metilación tiene un origen de replicación específico a la identidad del microorganismo lanzadera, de modo que cualquier gen presente en el constructo/vector de metilación se expresa en el microorganismo lanzadera.

La expresión de la enzima metiltransferasa da como resultado la metilación de los genes presentes en uno o más plásmidos a ser introducidos en un microorganismo progenitor. El plásmido puede entonces aislarse del microorganismo lanzadera de acuerdo con uno cualquiera de una serie de métodos conocidos. Por ejemplo, los kits disponibles en el mercado, como Qiagen o Zymo, pueden usarse de acuerdo con las instrucciones del fabricante.

- 5 En una realización particular, tanto el constructo de metilación/vector como el uno o más plásmidos de la invención se aíslan simultáneamente.

El uno o más plásmidos destinados al microorganismo progenitor pueden introducirse en el microorganismo usando cualquier número de métodos conocidos. Sin embargo, a modo de ejemplo, se puede utilizar la metodología descrita anteriormente en este documento, o en la sección de ejemplos que figura a continuación.

- 10 Se prevé que un gen de metiltransferasa puede introducirse en un microorganismo lanzadera y sobreexpresarse. Por lo tanto, en una realización, la enzima metiltransferasa resultante puede recogerse usando métodos conocidos y usarse in vitro para metilar uno o más plásmidos a ser introducidos en el microorganismo progenitor. El uno o más plásmidos pueden entonces introducirse en el microorganismo diana (progenitor). En otra realización, el gen de la metiltransferasa se introduce en el genoma del microorganismo lanzadera seguido de la introducción del uno o más plásmidos destinados para el microorganismo progenitor en el microorganismo lanzadera, el aislamiento del uno o más plásmidos del microorganismo lanzadera y luego la introducción de el uno o más plásmidos en el microorganismo diana (progenitor).

- 20 Se prevé que el uno o más plásmidos destinados para el microorganismo progenitor y el constructo/vector de metilación, tal como se definió anteriormente, se pueden combinar para proporcionar una composición de materia. Dicha composición tiene una utilidad particular para eludir los mecanismos de barrera de restricción para producir los microorganismos recombinantes de la invención.

En una realización particular, el constructo/vector de metilación es un plásmido.

- 25 Las personas expertas apreciarán una serie de metiltransferasas adecuadas de uso en la producción de microorganismos de acuerdo con la invención. Sin embargo, a modo de ejemplo, se pueden usar la metiltransferasa del fago Bacillus subtilis ΦT1 y la metiltransferasa descrita en el documento WO2012053905. Los ácidos nucleicos que codifican las metiltransferasas adecuadas se apreciarán fácilmente teniendo en cuenta la secuencia de la metiltransferasa deseada y el código genético.

- 30 Se puede usar cualquier número de construcciones/vectores adaptados para permitir la expresión de un gen de metiltransferasa para generar el constructo/vector de metilación. Sin embargo, a modo de ejemplo, se pueden usar los mencionados en el documento WO2012053905.

Una vez que se ha introducido un plásmido en un microorganismo progenitor deseado, se produce una primera selección. Esto implica seleccionar uno o más microorganismos que expresen al menos un marcador de selección positiva.

- 35 Dichos microorganismos pueden identificarse y seleccionarse utilizando cualquier número de técnicas conocidas, teniendo en cuenta el marcador de selección positivo que se esté utilizando. Sin embargo, a modo de ejemplo general, los microorganismos pueden cultivarse en un medio, o sobre éste, que contiene una toxina que destruiría cualquier microorganismo que no expresara el marcador de selección positiva. A modo de ejemplo específico, los microorganismos pueden crecer en presencia de un antibiótico tóxico, incluyendo el plásmido de la invención un ácido nucleico que codifica un producto que confiere resistencia antibiótica al microorganismo. Los microorganismos en los que está presente el plásmido sobrevivirán y los que no lo hagan morirán.

Se describen otros ejemplos de metodología y condiciones de uso en la selección de microorganismos que expresan un marcador de selección positivo, por ejemplo, en Sambrook et al, 1989 (como se describió anteriormente en este documento). Se proporcionan ejemplos adicionales en la sección de ejemplos que se incluye a continuación.

- 45 Los métodos de la invención también incluyen una segunda selección. Esto implica seleccionar uno o más microorganismos que no expresen al menos un marcador de contra-selección.

- 50 En el caso del uso de ThiK como marcador de contra-selección, la selección de uno o más microorganismos que no expresan este marcador de contra-selección implica el cultivo de los microorganismos en un medio, o sobre éste, que contenga un análogo de guanósina. En una realización particular, el análogo de guanósina es ganciclovir. Los microorganismos que contienen y expresan un ácido nucleico que codifica el marcador de contra-selección ThiK no sobrevivirán en presencia del análogo de guanósina. Por consiguiente, los microorganismos que sobrevivan se seleccionarán como que han sufrido el evento de recombinación de doble entrecruzamiento deseado.

- 55 En el caso del uso de un PheS alterado como marcador de contra-selección, la selección de uno o más microorganismos que no expresen este marcador de contra-selección implica cultivar los microorganismos en o sobre un medio que contenga un análogo de fenilalanina. En una realización particular, el análogo de fenilalanina es como se describe en este documento antes. Los microorganismos que contienen y expresan un ácido nucleico que

codifica el marcador de contra-selección de PheS alterado no sobrevivirán en presencia del análogo de fenilalanina. Por consiguiente, los microorganismos que sobreviven se seleccionan como que han sufrido el evento de recombinación de doble entrecruzamiento deseado.

5 Cuando se utiliza un PheS alterado como marcador de contra-selección, puede ser necesario incluir también fenilalanina en los medios.

10 Los métodos de la invención incluyen pasos de selección simultáneos y consecutivos. Por ejemplo, se podrían seleccionar microorganismos para eventos de entrecruzamiento únicos utilizando el marcador de selección positiva y, posteriormente, seleccionando microorganismos para eventos de doble entrecruzamiento utilizando el marcador de contra-selección. Alternativamente, la selección positiva y la contra-selección pueden ocurrir simultáneamente. A modo de ejemplo, cuando el ácido nucleico que codifica el marcador de selección positiva se coloca en el vector plasmídico fuera de los brazos de homología, se puede seleccionar consecutivamente para eventos de un solo entrecruzamiento, y luego volver a seleccionar, seleccionando los eventos de doble entrecruzamiento. A modo de ejemplo adicional, cuando el ácido nucleico que codifica el marcador de selección positiva está ubicado entre los brazos de homología (y, por lo tanto, está integrado en el genoma del microorganismo progenitor), la selección positiva y la contra-selección pueden ocurrir simultáneamente; cualquier célula que tenga el marcador de selección positiva integrado en el genoma y sea resistente al marcador de contra-selección, habrá tenido un evento de doble entrecruzamiento.

20 Cualquier medio adecuado para el cultivo de uno o más microorganismos puede usarse en un método de la invención. Las personas expertas apreciarán fácilmente los medios apropiados en función de la información publicada y teniendo en cuenta la naturaleza de la invención y los microorganismos progenitores descritos en el presente documento. Preferiblemente, los medios serán un medio en el que hay poca o ninguna fenilalanina presente, o al menos un nivel de fenilalanina que no supere a los análogos de fenilalanina durante la contraselección. A modo de ejemplo, cualquier medio mínimo apropiado sería adecuado, como: Medio mínimo de clostridia, Medio definido mínimo (MDM), Medio definido complementado (SDM) y Medio definido completo (MDL). Se proporcionan ejemplos específicos en este documento después en la sección de ejemplos.

Una vez que se selecciona uno o más microorganismos de acuerdo con la invención, pueden cultivarse y almacenarse opcionalmente para su uso futuro utilizando la metodología conocida.

La invención se describirá ahora, solo a modo de ejemplo, con referencia a los siguientes ejemplos.

### Ejemplos

30 Los siguientes ejemplos describen la construcción de plásmidos para los marcadores contraseleccionables HSV-Tk y PheS\*, la funcionalidad HSV-Tk y PheS\* para la contraselección en Clostridium autoethanogenum y el uso de HSV-Tk y PheS\* para facilitar la sustitución de genes de recombinación homóloga en el genoma de Clostridium autoethanogenum. El mismo principio también se puede aplicar a otros miembros de la familia Clostridium, ya que no existe un homólogo del gen HSV-Tk en ningún Clostridium secuenciado y los genes pheS o las especies de Clostridia están altamente conservados.

40 Se utilizaron técnicas de ADN recombinante estándar y clonación molecular en esta invención y se describen en Sambrook et al, 1989 y Ausubel et al, 1987. Se utilizaron la cepa TOP10 de E. coli (Life Technologies) y Clostridium autoethanogenum DSM10061 y DSM23693 (un derivado de DSM10061). Se cultivó E. coli en medio LB y SOB como describen Sambrook et al, 1989 y Ausubel et al, 1987, mientras que se cultivó Clostridium autoethanogenum en medio PETC anaeróbico (Tabla 1).

Tabla 1: Medio PETC (medio ATCC 1754; [atcc.org/Attachments/2940.pdf](http://atcc.org/Attachments/2940.pdf))

Componente del medio	Concentración por 1,0 L de medio
NH <sub>4</sub> Cl	1 g
KCl	0,1 g
MgSO <sub>4</sub> ·7H <sub>2</sub> O	0,2 g
NaCl	0,8 g
KH <sub>2</sub> PO <sub>4</sub>	0,1 g
CaCl <sub>2</sub>	0,02 g
Solución de traza de metal	10 ml
Solución de vitaminas de Wolfe	10 ml

ES 2 699 479 T3

Extracto de levadura	1 g
Resazurin (solución madre 2 g/L)	0,5 ml
MES	2 g
Agente reductor	0,006-0,008% (v/v)
Agua destilada	Hasta 1 l, pH 5,5 (ajustado con HCl)
Solución de vitaminas de Wolfe	por litro de solución madre
Biotina	2 mg
Ácido fólico	2 mg
Clorhidrato de piridoxina	10 mg
Tiamina.HCl	5 mg
Riboflavina	5 mg
Ácido nicotínico	5 mg
D-(+)-pantotenato de calcio	5 mg
Vitamina B <sub>12</sub>	0,1 mg
Acido p-aminobenzoico	5 mg
Ácido tióctico	5 mg
Agua Destilada	A 1 L
Solución de metal traza	por litro de solución madre
Acido nitrilotriacetico	2 g
MnSO <sub>4</sub> .H <sub>2</sub> O	1 g
Fe(SO <sub>4</sub> ) <sub>2</sub> (NH <sub>4</sub> ) <sub>2</sub> .6H <sub>2</sub> O	0,8 g
COCl <sub>2</sub> .6H <sub>2</sub> O	0,2 g
ZnSO <sub>4</sub> .7H <sub>2</sub> O	0,2 mg
CuCl <sub>2</sub> .2H <sub>2</sub> O	0,02 g
NaMoO <sub>4</sub> .2H <sub>2</sub> O	0,02 g
Na <sub>2</sub> SeO <sub>3</sub>	0,02 g
NiCl <sub>2</sub> .6H <sub>2</sub> O	0,02 g
Na <sub>2</sub> WO <sub>4</sub> .2H <sub>2</sub> O	0,02 g
Agua Destilada	Hasta 1 L
Solución madre del agente reductor	por 100 ml de solución madre
NaOH	0,9 g
Cisteína.HCl	4 g
Na <sub>2</sub> S	4 g
Agua destilada	Hasta 100 ml

Ejemplo 1: Funcionalidad de HSV-Tk y PheS\* para la contraselección en *Clostridium autoethanogenum*

## Construcción de plásmidos:

La construcción de plásmidos que contienen Hsv-tk: la secuencia de ADN de la timidina quinasa del virus del herpes humano (Hsv-tk) se obtuvo de NCBI (ácido nucleico y aminoácido). Los codones en el gen Hsv-tk se optimizaron para adaptarse a *C. autoethanogenum* y se sintetizaron por GeneArt y se administraron en su vector estándar pMK-RQ (Seq. ID. 1 - pMK-RQ-HSV-tk).

HSV-tk se liberó del vector pMK-RQ mediante la digestión con las enzimas de restricción NdeI y NheI (New England Biolabs) y se clonó en una versión modificada del vector lanzadera de *E. coli*-*Clostridium* pMTL83151 (FJ797651.1; Nigel Minton, Universidad de Nottingham; Heap et al, 2009) que se conoce como pMTL83155 (SEQ ID 3), entre los mismos sitios en la cepa de *E. coli* TOP10 (Life Technologies) para crear pMTL83155-Hsv-tk (Seq. ID. 2, Figura 1). El plásmido pMTL83155 contiene la secuencia promotora del gen de la fosfato acetil transferasa de *C. autoethanogenum* entre los sitios NotI y NdeI (Seq. ID. 3).

Construcción de pheS\* mutado: la pheS nativa se identificó en el genoma a partir de la secuencia de *C. autoethanogenum* DSM 10061 (Seq. ID.12). Comparando la secuencia de *E. coli* MG1655 (Seq. ID. 13) y pheS de *C. autoethanogenum* por alineamiento de secuencia, el sitio de especificidad de sustrato putativo se identificó basándose en la homología de la secuencia de aminoácidos de los aminoácidos entre G284 y G298 de *E. coli* (la secuencia de aminoácidos de PheS de MG1655 de *E. coli* es la SEQ ID 20) y los aminoácidos G301 y G315 de *C. autoethanogenum* (Figura 3). Se introdujo una única mutación puntual en la base 932, sustituyendo C por G, dando como resultado el codón que codificaba la glicina en lugar de la alanina (Seq. ID. 14).

Construcción de plásmidos que contienen pheS\*: la pheS\* modificada se acopló transcripcionalmente a un promotor sintético y un sitio RBS (PpheS\*; Seq. ID. 17) cadena arriba del codón de inicio para permitir una alta expresión constitutiva. La construcción también estaba flanqueada por sitios de restricción PmeI para permitir la clonación del gen en vectores alternativos. La síntesis y la subclonación en el vector pMTL85151 utilizando la enzima de restricción PmeI se llevó a cabo por GeneArt dando como resultado el vector final pMTL85151-pheS\* (Seq. ID. 18; Figura 4).

## Pruebas de sensibilidad:

Pruebas de toxicidad de DL-4-clorofenilalanina: Se cultivaron *E. coli* Top10 y *C. autoethanogenum* DSM23693 que albergan el vector vacío pMTL85151 inicialmente en presencia de DL-4-clorofenilalanina en placas y en medio líquido para determinar si el marcador de contra-selección tenía algún efecto sobre el crecimiento de los organismos. Se observó que el producto químico no impidió el crecimiento de ninguno de los organismos en medios líquidos o en placas y las colonias crecieron al mismo tamaño después de 24/48 horas para *E. coli* y *C. autoethanogenum*, respectivamente.

Prueba de pheS\* en *E. coli*: para probar la capacidad del marcador de contra-selección para funcionar en *E. coli*, el plásmido pMTL85151-pheS\* se transformó en TOP10 y se cultivó solo con selección de cloranfenicol. Una vez que el cultivo alcanzó una OD de 0,5, representando una fase de crecimiento exponencial, se colocaron 100 ul en placas LB que contenían cloranfenicol y DL-4-clorofenilalanina, así como cloranfenicol solo como control, por triplicado, y se incubaron durante 24 horas a 37°C. Después de 24 horas, se inspeccionaron las placas y se observó que en las placas de cloranfenicol solo, había un jardín de colonias grandes como se esperaba, sin embargo, en las placas que contenían cloranfenicol y DL-4-clorofenilalanina, había un ligero sombreado de pequeñas colonias, lo que sugiere que la DL-4-clorofenilalanina estaba afectando el crecimiento de la *E. coli* que albergaba pMTL85151-pheS\*. Después de 36 horas, las placas de selección doble habían crecido al mismo nivel que las placas de cloranfenicol solo.

Transformación de *C. autoethanogenum* y confirmación de plásmidos:

Transformación de *C. autoethanogenum*: los plásmidos pMTL83155, pMTL83155-Hsv-tk y pMTL85155-PheS se introdujeron en *C. autoethanogenum* DSM23693 (un derivado de DSM10061) como se describe en el documento WO2012053905. El crecimiento se realizó en caldo de PETC y se extendió en medio de agar PETC suplementado con 15 µg/ml de tiamfenicol (Sigma) y 10 µg/ml de trimetoprim (Sigma). Se observaron colonias después de 3 días de incubación a 37°C en recipientes de gas a presión con 20 psi de una mezcla de gases de 48% de CO<sub>2</sub>, 2% de H<sub>2</sub>, 20% de CO<sub>2</sub>, 30% de N<sub>2</sub>. Se hicieron estrías de colonias individuales en medios de PETC-agar que contenían 15 µg/ml de tiamfenicol.

Selección de transformantes para detectar la presencia de plásmidos: 2 colonias de LZ-pMTL83155 y LZ-pMTL83155-hsv-tk trans-conjugantes se seleccionaron al azar para detectar la presencia de los plásmidos pMTL83155 y pMTL83155-Hsv-tk por PCR utilizando los cebadores repHf (Seq. ID. 4) y catr (Seq. ID. 5) abarcando el replicón Gram-positivo y el marcador de selección positivo catP en pMTL83155 y pMTL83155-Hsv-tk. Se utilizó *C. autoethanogenum* no modificado como control en estas PCR. El kit Maxime PCR PreMix se utilizó para la PCR. También se amplificó por PCR 16s rDNA a partir de estos transformantes usando los cebadores fDI (Seq. ID. 6) y rP2 (Seq. ID.7) y Maxime PCR PreMix Kit.

La PCR con cebadores repHF y catR amplificó bandas ~1,5 kb de LZ-pMTL83155-1 y -2 y LZ-pMTL83155-hsv-tk-1 y -2 (Figura 2). No se detectó amplificación en la muestra de *C. autoethanogenum* no modificada, lo que confirma la presencia de plásmidos solo en los transformantes. La secuenciación del ARNr 16s de LZ-pMTL83155-1 (Seq. ID. 8) y -2 (Seq. ID. 9) y LZ-pMTL83155-hsv-tk-1 (Seq. ID. 10) y -2 (Seq. ID. 11) confirmó además que los clones eran *C. autoethanogenum*.

La confirmación de los transformantes de *C. autoethanogenum* que albergan el plásmido pMTL85151-pheS\* se llevó a cabo en tres colonias independientes utilizando cebadores de PCR específicos para el plásmido (Seq. ID. 15 y 16), y los cebadores fDI (Seq. ID. 6) y rP2 (Seq. ID.7) para secuenciar el ARNr 16s.

Funcionalidad de HSV-Tk y PheS\* como marcadores contraseleccionables en *Clostridium autoethanogenum*:

10 Sensibilidad de los transformantes de *C. autoethanogenum* a ganciclovir: La sensibilidad de LZ-pMTL83155 y LZ-pMTL83155-hsv-tk a ganciclovir se probó en placas en agar de PETC que contenían ganciclovir 20 nM solo y medio agar PETC que contenía 20 nM ganciclovir y 15 µg/ml de tiamfenicol. Las colonias en las placas de ganciclovir se observaron solo con transformantes LZ-pMTL83155 y no con LZ-pMTL83155-hsv-tk (Tabla 2). La presencia del gen Hsv-tk confiere toxicidad al ganciclovir.

15 Sensibilidad de *C. autoethanogenum* que alberga el plásmido pMTL85151-pheS\* a DL-4-clorofenalanina: Los tres transformantes independientes de *C. autoethanogenum* DSM23693 que albergan pMTL85151-pheS\*, así como *C. autoethanogenum* DSM23693 que contiene pMTL85251 se cultivaron en medio PECT líquido con tiamfenicol y se cultivaron a 37°C bajo condiciones de CO solo. Después de 24 horas, 100 µl de cada uno de los tres independientes, así como el control pMTL85151, se sembraron en agar PETC-MES suplementado con tiamfenicol  
20 solo o con tiamfenicol y DL-4-clorofenilalanina y se incubaron durante 48 horas. Después de 48 horas, se inspeccionaron las placas y las placas que contenían solo tiamfenicol tenían un jardín de colonias para las 4 cepas, mientras que las placas que contenían la selección doble solo tenían 3, 4 y 7 colonias para cada uno de los transformantes independientes que contenían pheS\*, en contraste, los transformantes de *C. autoethanogenum* que albergaban pMTL85151 mostraron los mismos resultados que la placa de tiamfenicol solo, lo que sugiere que la DL-4-clorofenilalanina no tenía efecto sobre esta cepa.  
25

Tabla 2: Sensibilidad de los transconjugantes de *C. autoethanogenum* a diferentes profármacos en presencia del CSM correspondiente

CSM	Profármaco	Concentración	LZ-pMTL83155	LZ-pMTL83155-hsv-tk
Hsv-tk	Ganciclovir	20 nM	colonias	0-5
PheS	DL-4-Clorofenilalanina	0,20%	colonias	7-Mar

30 Esto demostró que tanto HSV-Tk como PheS\* en combinación con los profármacos Ganciclovir y DL-4-Clorofenalanina son efectivos para la contraselección en *C. autoethanogenum*

Ejemplo 2 - Uso de PheS\* para facilitar el reemplazo del gen de recombinación homóloga en el genoma de *Clostridium autoethanogenum*

35 Este ejemplo describe la sustitución de un gen nativo de 2,3-butanodiol deshidrogenasa R-específico de *Clostridium autoethanogenum* con un gen de 2,3-butanodiol deshidrogenasa S-específico de *Klebsiella pneumoniae* a través de la recombinación homóloga facilitada por PheS\* como un contra-marcador seleccionable en *Clostridium autoethanogenum* DSM23693 que tiene una alcohol deshidrogenasa secundaria inactivada.

Construcción de la cepa DSM23693 de *C. autoethanogenum* que tiene una alcohol deshidrogenasa secundaria inactivada

40 Se construyó una cepa de *Clostridium autoethanogenum* DSM23693 que tiene una alcohol deshidrogenasa secundaria inactivada (SEQ ID NO: 55) utilizando el sistema ClosTron. (Heap et al., 2007). La herramienta de diseño de intrones alojada en el sitio web de ClosTron.com se utilizó para diseñar una región de reconocimiento de 344 pb (SEQ ID NO: 56), así como para identificar seis sitios diana (Fig. 8) en las cadenas sentido y antisentido. La región de reconocimiento se sintetizó químicamente en el vector pMTL007C-E2 que contenía un marcador ermB activado por retro-transposición (RAM).

45 Los vectores se introdujeron en *C. autoethanogenum* DSM23693 como se describe en el documento WO2012/053905. Se inocularon en estrías colonias individuales crecidas en PETC MES con 15 µg/ml de tiamfenicol en PETC MES con 5 µg/ml de clarotromicina. Las colonias de cada diana se seleccionaron al azar y se seleccionaron para la inserción utilizando los cebadores flanqueantes 155F (SEQ ID NO: 57) y 939R (SEQ ID NO: 58). La amplificación se realizó utilizando la premezcla de PCR iNtron Maxime. Un producto PCR de 783 pb indicó  
50 un genotipo natural, mientras que un tamaño de producto de aproximadamente 2,6 kb sugería la inserción del intrón

del grupo II en el sitio diana (Fig. 9). El plásmido de pérdida se comprobó mediante la amplificación del marcador de resistencia (catP) y el origen de replicación gram positivo (pCB 102). Esta cepa se usó como cepa de base que reemplaza un gen de 2,3-butanodiol deshidrogenasa R-específico de *Clostridium autoethanogenum* con el gen de 2,3-butanodiol deshidrogenasa S-específico de *Klebsiella pneumoniae* a través de la recombinación homóloga facilitada por PheS\*.

Construcción del plásmido pPheS-ErmB. El fragmento que contenía el casete PheS\* (Seq. ID 27) y ColE1 con traJ (Seq ID No. 26) se amplificó de pMTL85151-PheS\* (Seq ID No. 18) con los cebadores PheS-repH-F (SEQ ID No. 28) y traJ-ermB-R (SEQ ID No. 29). El fragmento que contenía el origen de replicación de pCB 102 (Seq ID No. 30) se amplificó de la serie pMTL80000 (Heap et al., 2009) utilizando los cebadores RepH-ermB-F (SEQ ID No. 31) y RepH-pheS-R (SEQ ID No. 32). El casete de resistencia a la eritromicina (SEQ ID No. 33) se amplificó de la serie pMTL80000 (Heap et al., 2009) utilizando los cebadores ermB-traJ-F (SEQ ID No. 34) y ermB-repH-R (SEQ ID No. 35). Los productos de PCR descritos contenían superposiciones para facilitar el ensamblaje sin problemas. Fueron ensamblados utilizando el kit de ensamblaje y clonación GENEART Seamless de Life Technologies. El plásmido resultante, pPheS\*-ErmB se verificó mediante digestión de restricción y análisis de fragmentos.

Amplificación por PCR de las partes del plásmido para recombinación homóloga. La estructura del vector que contenía PheS\* (SEQ ID No. 27) se amplificó a partir de pPheS\*-ErmB (SEQ ID No. 36) utilizando los cebadores AM015 (SEQ ID No. 37) y AM035 (SEQ ID No. 38). El brazo de homología en sentido ascendente (SEQ ID No. 39) se amplificó a partir del ADN genómico de *C. autoethanogenum* utilizando los cebadores AM016 (SEQ ID No. 40) y AM017 (SEQ ID No. 41). El gen de la butanodiol deshidrogenasa de *K. pneumoniae* (SEQ ID No. 42) se amplificó, utilizando los cebadores AM018 (SEQ ID No. 43) y AM019 (SEQ ID No. 44), de pMTL85141-P-alsS-budA-budA (Köpke et al., 2014). El casete de expresión de cloranfenicol acetiltransferasa (SEQ ID No. 45) se amplificó de la serie pMTL80000 (Heap et al., 2009) utilizando los cebadores AM020 (SEQ ID No. 46) y AM021 (SEQ ID No. 47). El brazo de homología en sentido descendente (SEQ ID No. 48) se amplificó a partir del ADN genómico de *C. autoethanogenum* utilizando los cebadores AM022 (SEQ ID No. 49) y AM036 (SEQ ID No. 50).

Construcción del plásmido que contiene pheS\* para recombinación homóloga. Los productos PCR anteriores contenían superposiciones para facilitar el ensamblaje sin problemas. Fueron ensamblados utilizando el kit de ensamblaje y clonación GENEART Seamless de Life Technologies. El plásmido resultante, pPheS\*-CaBDHXXKpBDH se verificó mediante digestión de restricción y análisis de fragmentos.

Introducción de plásmidos, contra-selección y rastreo para la integración. Se introdujo pPheS\*-CaBDHXXKpBDH como se describe en el documento WO2012053905 en una cepa de *C. autoethanogenum* DSM23693 que tenía una alcohol deshidrogenasa secundaria inactivada por CloStron en la posición 287 como se ha descrito anteriormente. Los transformantes se seleccionaron por su capacidad de crecer en medio de agar PETC suplementado con 15 µg/ml de tiamfenicol (Sigma) y 10 µg/ml de trimetoprima (Sigma). Para seleccionar el doble entrecruzamiento de recombinación homóloga exitosa, las colonias se reiniciaron en medio de agar PETC suplementado con 15 µg/ml de tiamfenicol y 2 mg/ml de p-clorofenilalanina. Las colonias que crecieron se seleccionaron para la integración por PCR con los cebadores AM041 (SEQ ID No. 51) y AM042 (SEQ ID No. 52) que bordeaban el sitio de integración fuera de los brazos de homología. La integración exitosa se identificó como la producción de un producto PCR de 3570 pares de bases de longitud en comparación con 3137 pares de bases para el natural (Figura 7). El producto de la PCR se secuenció por fidelidad mediante la secuenciación de Sanger y se encontró que era exactamente la secuencia de inserción esperada, lo que confirmaba la integración exitosa del fragmento facilitado por el marcador contraseleccionable pheS\*.

Tabla 3 - Cebadores utilizados en el Ejemplo 2

Nombre de cebador	Secuencia	SEQ ID
AM015	CTTGCCTTGCTCGTCGGT	37
AM016	GACGAGCAAGGCAAGCAATTATAGTGAAAGATGTGAAGG	40
AM017	ACCTTTTTTCATAATTATCTCTCTTTTTTTATAATAGTATGG	41
AM018	AGAGATAATTATGAAAAAGGTTGCATTAGTTAC	43
AM019	CCTTACAATTTAATTAATACCATACCACCGTC	44
AM020	GTATTTAATTAATTGTAAGGATCCTAGTCAG	46
AM021	GTACTTTTTATGAGCTCTTAACTATTATATCAATTC	47
AM022	AGAGCTCATAAAAAGTACTCATAGAATTGATTAATAAATG	49
AM035	AAGTGATAGTCAAAGGCATAACAGTG	38

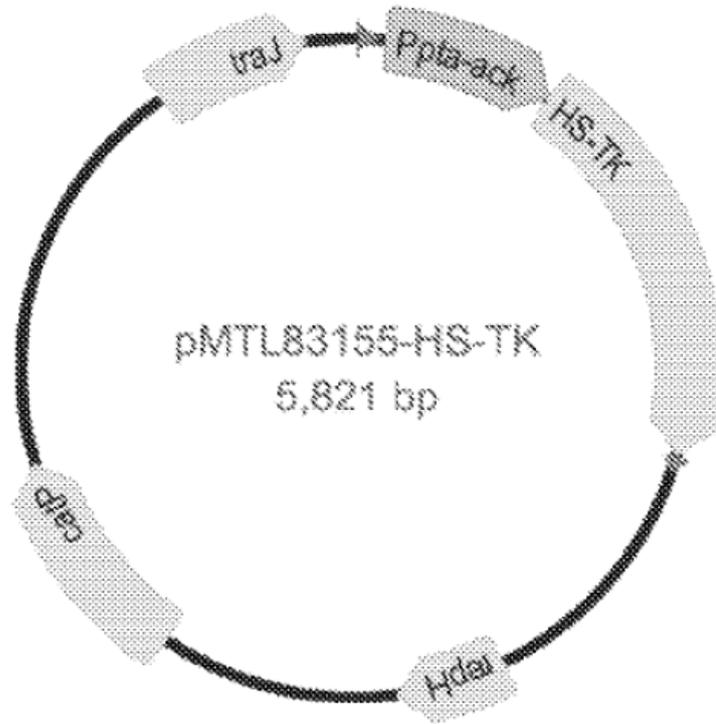
AM036	ATGCCTTTTGA CTATCACTTATACATCTCCTTTAAATCCATTTG	50
AM041	CTGGAAAAGAACTCTTAGC	51
AM042	TGCGGTGGAATACAATGG	52
PheS-repH-F	GCAAGTTGAAAAATTCACGAAAGTTACACGTTACTAAAGG	28
traJ-ermB-R	CACTATCAACACACTCTTAAGCTTGCCTTGCTCGTCGGTG	29
RepH-ermB-F	GCTTTTGTAAATTTGCATAAAAAATAAGAAGCCTGCATTTG	31
RepH-pheS-R	TTTAGTAACGTGTAACCTTCGTGAATTTTCAACTTGCC	32
ermB-traJ-F	CACCGACGAGCAAGGCAAGCTTAAGAGTGTGTTGATAGTG	34
ermB-repH-R	GCTTCTTATTTTTATGCAAATTTACAAAAGCGACTCATAG	35

La misma estrategia y el mismo plásmido también se pueden aplicar a *C. ljungdahlii* o *C. ragsdalei*. Se han descrito los protocolos de transformación (documento WO2012/053905) (Leang, Ueki, Nevin, y Lovley, 2012).

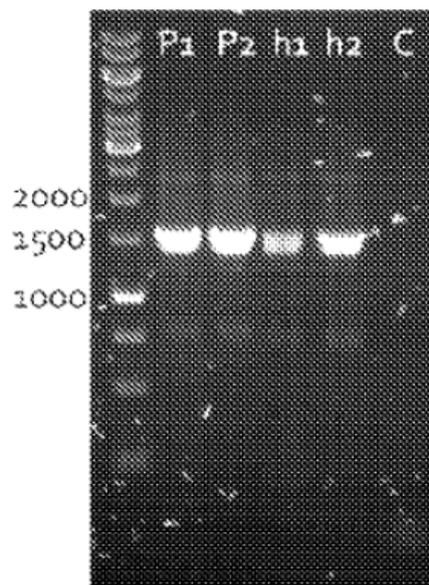
- 5 La invención se ha descrito en este documento, con referencia a ciertas realizaciones preferidas, con el fin de permitir al lector poner en práctica la invención sin realizar una experimentación excesiva. Sin embargo, cualquier persona con experiencia ordinaria en la técnica reconocerá fácilmente que muchos de los componentes y parámetros pueden variarse o modificarse en cierta medida o sustituirse por equivalentes conocidos sin apartarse del alcance de la invención. Debe apreciarse que dichas modificaciones y equivalentes se incorporan en este documento como si se expusieran individualmente.
- 10 A lo largo de esta memoria descriptiva y de las siguientes reivindicaciones, a menos que el contexto requiera lo contrario, las palabras "comprende", "que comprende" y similares, deben interpretarse en un sentido inclusivo en lugar de un sentido exclusivo, es decir, en sentido de "que incluye, entre otros".

**REIVINDICACIONES**

1. Una fenilalanina tRNA sintetasa modificada (PheS), en donde la PheS modificada tiene la secuencia de la SEQ ID N° 21 que comprende al menos una alteración en las posiciones de los aminoácidos 306 a 313.
- 5 2. La PheS modificada de la reivindicación 1, en la que la PheS modificada aminoacila el ARNt con un análogo de fenilalanina que se selecciona del grupo que consiste en clorofenilalanina, fluorofenilalanina y bromofenilalanina.
3. La PheS modificada de la reivindicación 2, en la que el análogo de fenilalanina se selecciona del grupo que consiste en DL-4-clorofenilalanina, p-clorofenilalanina, p-fluoro-L-fenilalanina, p-fluoro-DL-fenilalanina y p-bromo -L-fenilalanina.
- 10 4. La PheS modificada de la reivindicación 1, en la que la PheS modificada comprende una alteración en la posición del aminoácido 311 de la SEQ ID N° 21.
5. La PheS modificada de la reivindicación 4, en la que la PheS modificada comprende glicina en lugar de alanina en la posición de aminoácido 311 de la SEQ ID N° 21.
6. Un ácido nucleico que codifica la PheS modificada de la reivindicación 1.
- 15 7. Un método para producir un microorganismo recombinante, que comprende introducir en un microorganismo progenitor un vector que comprende un ácido nucleico que codifica la PheS modificada de la reivindicación 1.
8. El método de la reivindicación 7, en el que el método comprende además la selección frente a la expresión de la PheS modificada.
- 20 9. El método de la reivindicación 8, en el que la selección frente a la expresión de la PheS modificada implica cultivar el microorganismo recombinante en presencia de un análogo de fenilalanina seleccionado del grupo que consiste en clorofenilalanina, fluorofenilalanina y bromofenilalanina.
10. El método de la reivindicación 9, en el que el análogo de fenilalanina se selecciona del grupo que consiste en DL-4-clorofenilalanina, p-clorofenilalanina, p-fluoro-L-fenilalanina, p-fluoro-DL-fenilalanina y p-bromo L-fenilalanina.
- 25 11. El método de la reivindicación 7, en el que el vector comprende además brazos de homología que permiten la recombinación homóloga entre el vector y el genoma del microorganismo progenitor, un ácido nucleico de interés y/o un marcador de selección positiva.
12. El método de la reivindicación 7, en el que el método comprende además seleccionar la expresión del marcador de selección positiva, preferiblemente en el que el marcador de selección positiva se selecciona del grupo que consiste en CatP, ErmB o TetA.
13. El uso del PheS de la reivindicación 1 como un marcador de contra-selección.
- 30 14. El método de la reivindicación 9, en el que la PheS modificada comprende una alteración en la posición del aminoácido 311 de la SEQ ID N° 21.
15. El método de la reivindicación 14, en el que la PheS modificada comprende glicina en lugar de alanina en la posición de aminoácido 311 de la SEQ ID N° 21.

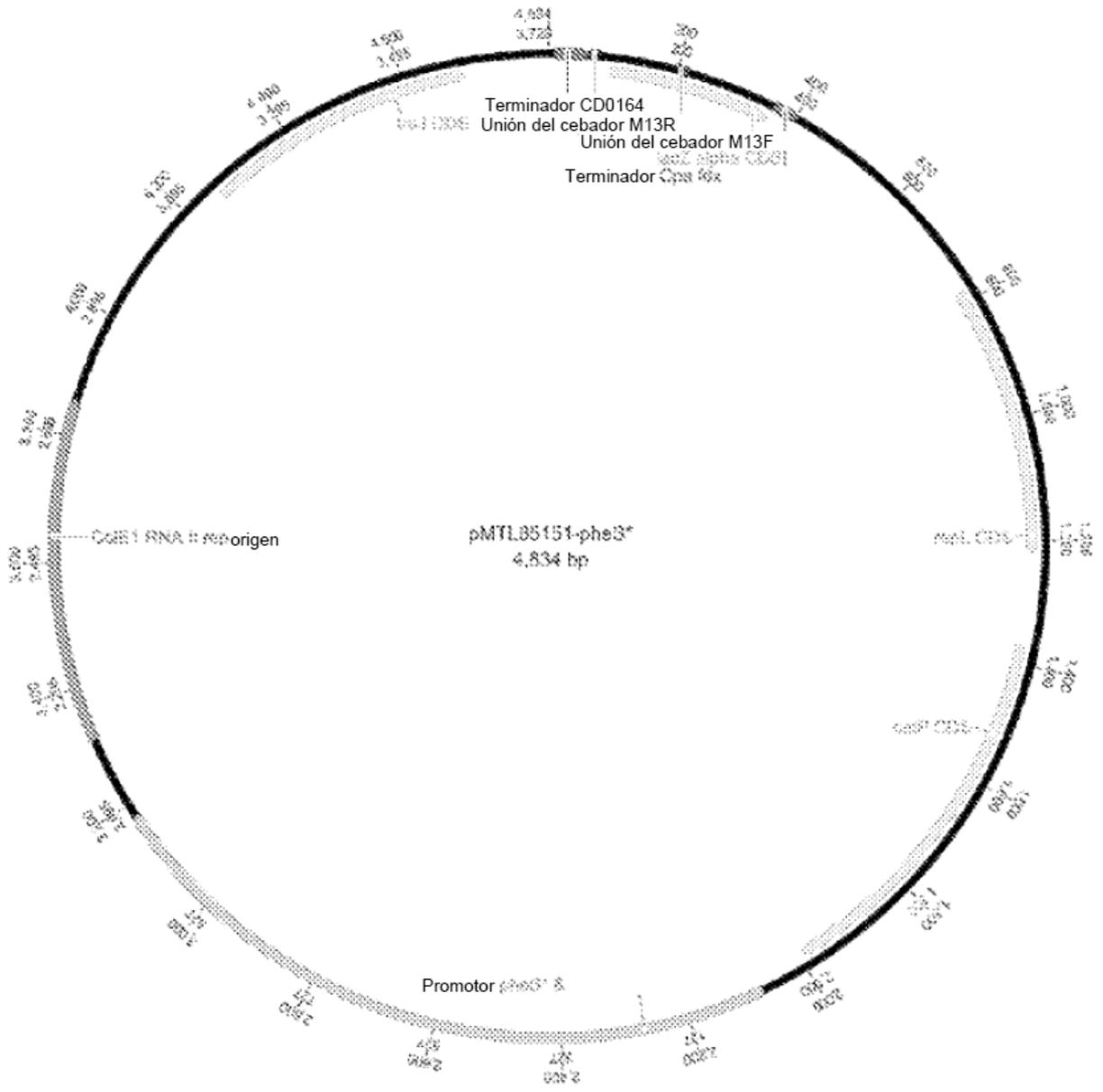


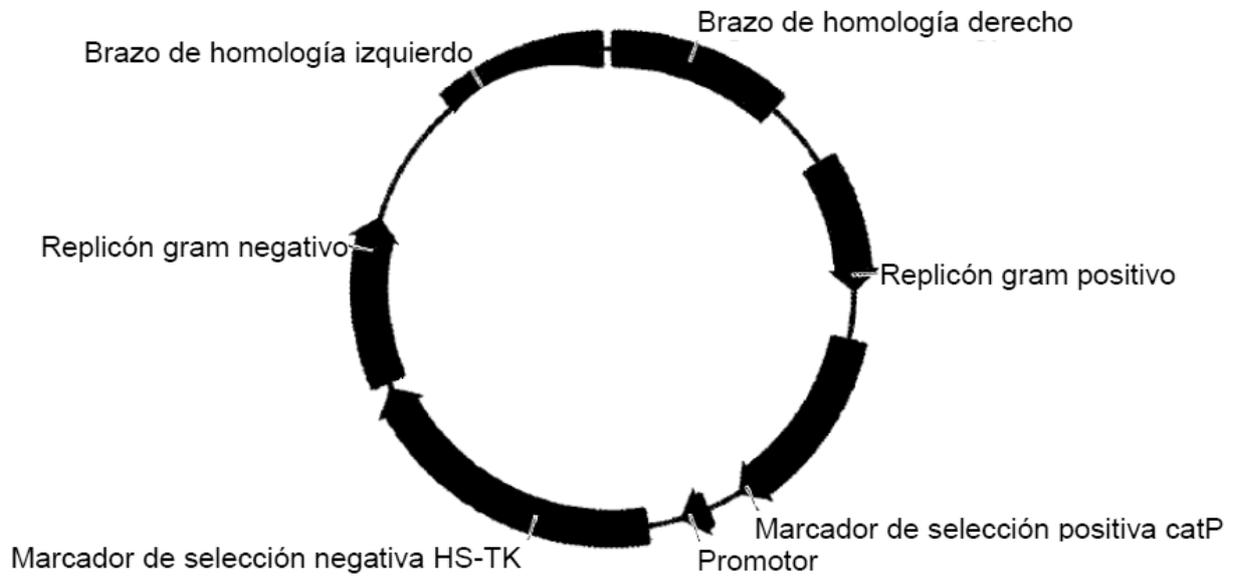
**FIG. 1**



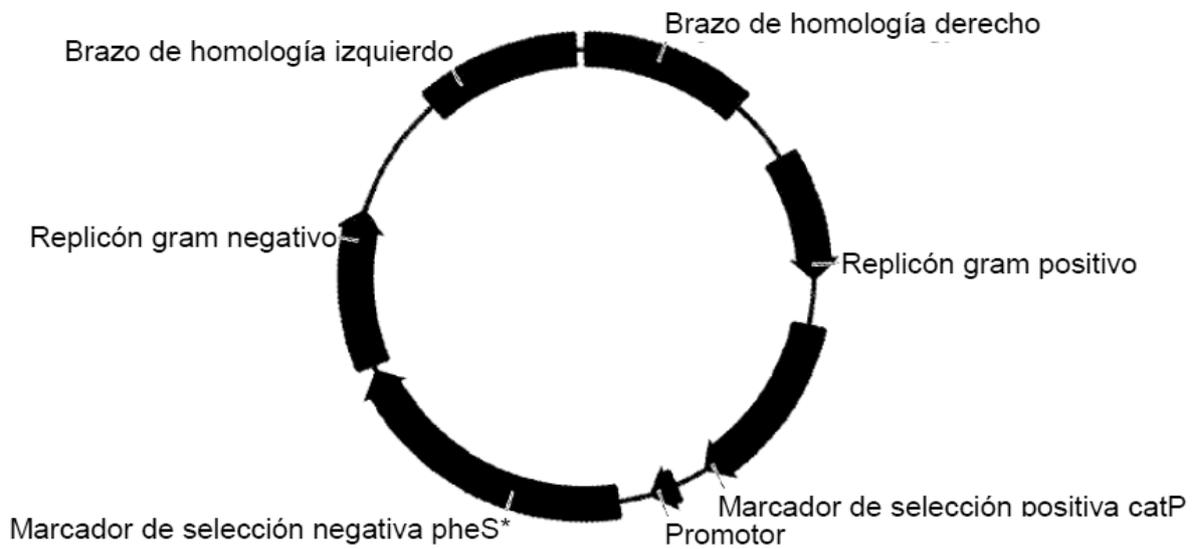
**FIG. 2**







**FIG. 5**



**FIG. 6**

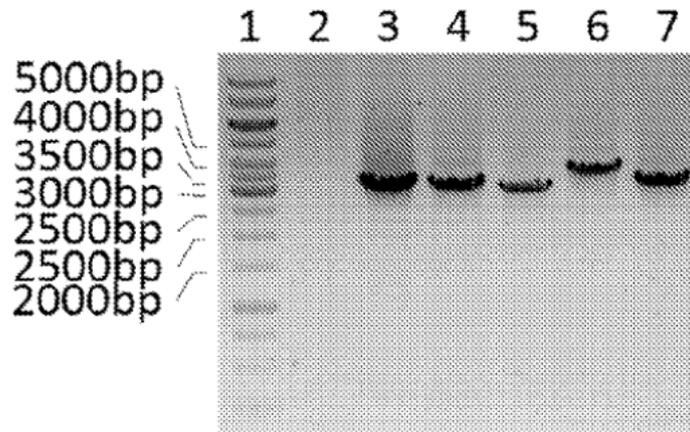


FIG. 7

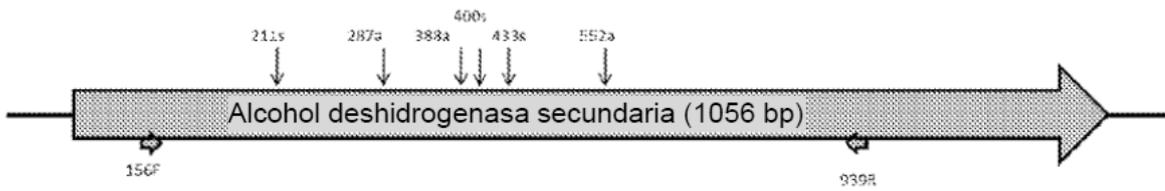
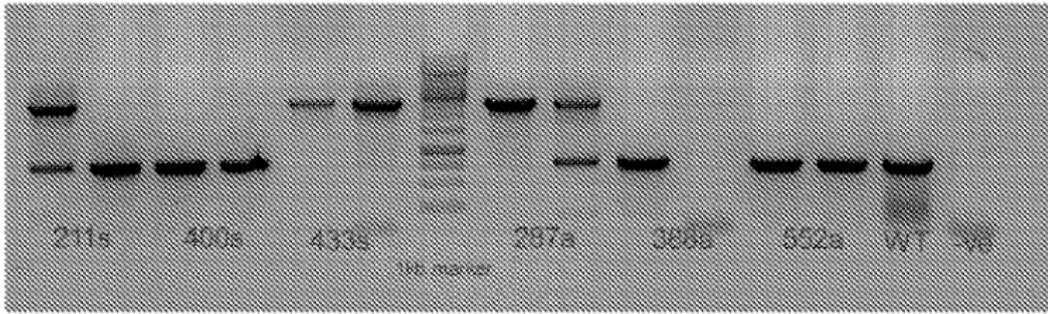
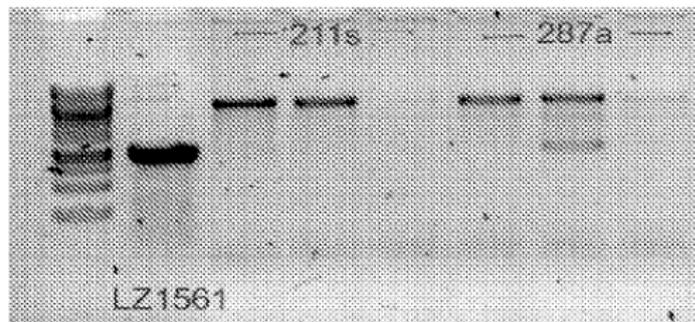


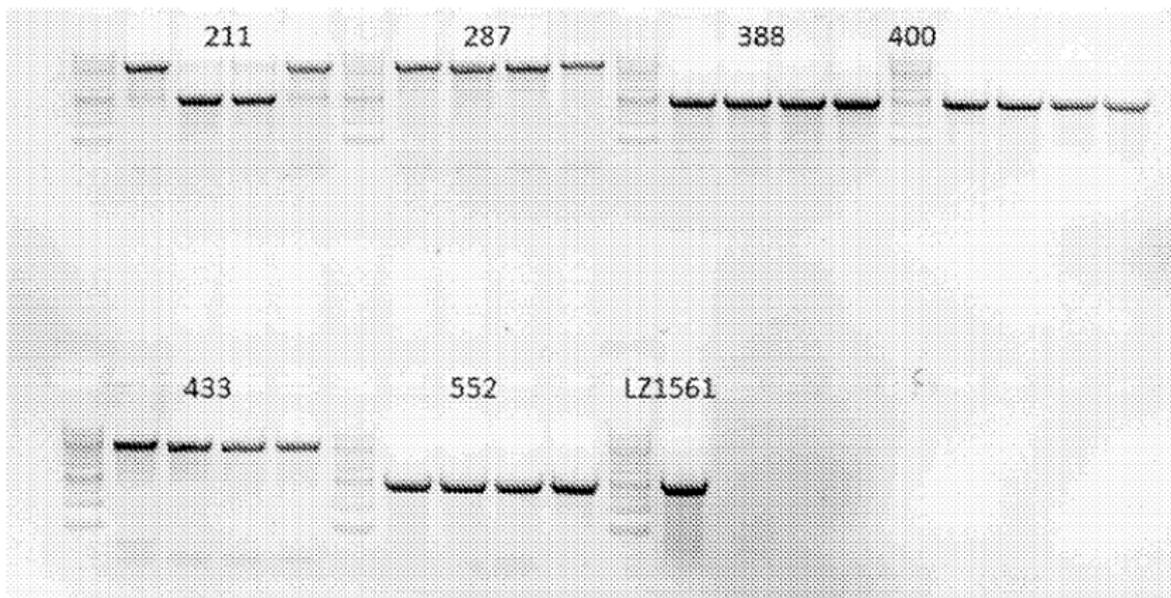
FIG. 8



**FIG. 9A**



**FIG. 9B**



**FIG. 9C**