

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 728 307**

51 Int. Cl.:

**C12N 15/70** (2006.01)

**C12N 15/74** (2006.01)

**C12P 5/00** (2006.01)

**C12P 7/06** (2006.01)

**C12N 9/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.06.2014 PCT/US2014/043424**

87 Fecha y número de publicación internacional: **24.12.2014 WO14205355**

96 Fecha de presentación y número de la solicitud europea: **20.06.2014 E 14737481 (3)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019 EP 3011018**

54 Título: **Composiciones y métodos para la transformación clostridial**

30 Prioridad:

**21.06.2013 US 201361838224 P**

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

**23.10.2019**

73 Titular/es:

**DANISCO US INC. (100.0%)  
925 Page Mill Road  
Palo Alto, CA 94304-1013, US**

72 Inventor/es:

**SCOTCHER, MILES, C. y  
WELLS, DEREK, H.**

74 Agente/Representante:

**MARTÍN BADAJOZ, Irene**

**Observaciones:**

**Véase nota informativa (Remarks, Remarques o  
Bemerkungen) en el folleto original publicado por  
la Oficina Europea de Patentes**

ES 2 728 307 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

Composiciones y métodos para la transformación clostridial

5 **Referencia cruzada a solicitud relacionada**

Esta solicitud reivindica prioridad con respecto a la solicitud de patente provisional estadounidense n.º 61/838.224, presentada el 21 de junio de 2013; cuyo contenido se incorpora en el presente documento como referencia en su totalidad.

10

**Campo de la invención**

La invención proporciona composiciones y métodos para la ingeniería genética de bacterias clostridiales para producir y/o mejorar la eficacia de producción de bioproductos industriales.

15

**Antecedentes de la invención**

Los sistemas de restricción/modificación (R-M) bacterianos tienen una especificidad y estrategia diversas, pero su función general es proteger a las bacterias del ADN foráneo, tal como el ADN de bacteriófagos. Los sistemas de R-M pueden consistir en una ADN metiltransferasa y una endonucleasa de restricción. Las ADN metiltransferasas catalizan la transferencia de un grupo metilo del donador S-adenosil-L-metionina (también conocido como "SAM" o "AdoMet") a residuos de adenina o citosina dentro de secuencias de ADN particulares de la bacteria huésped, que se denominan secuencias de reconocimiento. Existen tres clases principales de ADN metiltransferasas, clasificadas según la naturaleza del producto que producen. La primera clase consiste en amino-metiltransferasas que catalizan la metilación del grupo amino exocíclico de la adenina para formar el producto N6-metiladenina. La segunda clase consiste en amino-metiltransferasas que catalizan la formación del grupo amino exocíclico de la citosina para formar el producto N4-metilcitosina, mientras que la tercera clase consiste en metiltransferasas que metilan el átomo de carbono-5 cíclico de la citosina para formar la 5-metilcitosina. Estas bases metiladas desempeñan funciones importantes en los sistemas de R-M bacterianos, ya que protegen el cromosoma huésped contra la acción, por lo demás perjudicial, de la enzima de restricción asociada, que escinde el ADN de la secuencia de reconocimiento no metilado, pero ignora el ADN completamente metilado. Por tanto, es la acción combinada de la ADN metiltransferasa y su endonucleasa de restricción relacionada la que protege a la bacteria huésped de cualquier ADN foráneo no modificado. Mientras que los sistemas de R-M realizan una importante función protectora, también inhiben la transferencia de plásmidos entre especies bacterianas e incluso entre cepas de la misma especie de bacterias, ya que múltiples sistemas de R-M dentro de una sola cepa bacteriana pueden participar en la barrera de restricción. Por lo tanto, los sistemas de R-M actúan como barrera para la manipulación genética de muchas bacterias, incluyendo el género de importancia biotecnológica *Clostridium*.

El género *Clostridium* consiste en un gran número de especies con una amplia gama de rasgos bioquímicos y fisiológicos. Véase Cato *et al.*, 1986, Genus *Clostridium*, págs. 1141-1200, en P.H. Sneath *et al.* (eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 2, Williams y Wilkins, Baltimore, MD. Existen cuatro criterios que es necesario cumplir para que un aislado se asigne al género *Clostridium*: (1) la capacidad para formar endosporas, (2) metabolismo energético anaerobio, (3) la incapacidad de reducción de sulfato disimilatoria y (4) la posesión de una pared celular Gram-positiva. Véase Andresson *et al.*, 1989, *Introduction to the physiology and biochemistry of the genus Clostridium*, págs. 27-62, en Minton y Clarke (eds.), *Clostridia*, Plenum Press, Nueva York. Las bacterias acetogénicas del género *Clostridium* usan gas de síntesis (sintegas) como fuente de carbono y poder de reducción para el crecimiento en condiciones anaerobias. El sintegas se compone de una mezcla de H<sub>2</sub>, CO y CO<sub>2</sub>, que se produce mediante gasificación de cualquier material orgánico, desde residuos municipales hasta residuos agrícolas. El uso de sintegas como materia prima para la producción biológica de productos químicos y enzimas de mercancía es atractivo debido a su bajo coste y la amplitud y flexibilidad de las fuentes de las que se deriva. Sin embargo, los acetógenos dentro del género *Clostridium* están relativamente sin caracterizar, y la capacidad para manipular genéticamente estos organismos, particularmente a través de la introducción de ácidos nucleicos heterólogos que son estables y no se escinden por endonucleasas de restricción clostridiales, está en gran medida sin desarrollar. La capacidad para transformar bacterias clostridiales es una primera etapa necesaria y fundamental para su uso eficaz en la producción de bioproductos industriales (por ejemplo, isopreno, butadieno y etanol).

Los esfuerzos para superar los sistemas de R-M en *Clostridium* han implicado normalmente la metilación *in vivo* del ADN heterólogo antes de su transformación para protegerlo de la degradación por endonucleasas de restricción en las células huésped; por ejemplo, puede realizarse metilación *in vivo* transformando plásmidos lanzadera dentro de una cepa (por ejemplo, *E. coli*) que expresa una o más metiltransferasas heterólogas (por ejemplo, una metiltransferasa del fago de *Bacillus subtilis* Φ 3T). Después de aislar el ADN metilado, puede transformarse dentro de células anaerobias huésped (por ejemplo, células de *Clostridium acetivum*) mediante electroporación, transformación de protoplastos, transformación conjugativa, pistola génica u otro método conocido en la técnica.

Otros métodos para superar los sistemas de R-M clostridiales implican la metilación de ADN heterólogo *in vitro* usando una o más enzimas metiltransferasa purificadas disponibles para su adquisición a proveedores comerciales

(por ejemplo, New England BioLabs), o implican la creación y el uso de células huésped clostridiales deficientes en al menos un gen de endonucleasa de restricción en su sistema de restricción-modificación. Véase, por ejemplo, Dong *et al.*, PLoS ONE 2010 5 (2): e9038. En Dong *et al.* (2010), una supuesta endonucleasa de restricción de tipo II (Cac824I), identificada a partir del genoma disponible públicamente de *Clostridium acetobutylicum* ATCC 824, se alteró usando el sistema de inactivación génica basado en inserción de intrones del grupo II de ClosTron. El sistema ClosTron, similar a la mayoría de los enfoques de intrones del grupo II, usa un elemento derivado del intrón LI.LtrB de intervalo de huésped amplio de *Lactococcus lactis*. Véase, por ejemplo, Kuehne *et al.*, 2011, ClosTron-mediated engineering of *Clostridium*. *Methods in Molecular Biology*, vol. 765: 389-407. Las células resultantes deficientes en Cac824I podrían transformarse con ADN no metilado (por ejemplo, ADN plasmídico no metilado) a través de electroporación. Butkus *et al* (1985, *Nucleic Acids Res.*, 13 (16): 5727-5746) describen una enzima de restricción-modificación MvaI.

Sin embargo, estos procesos para superar los sistemas de restricción-modificación en bacterias clostridiales dependen de la identificación de las metiltransferasas y endonucleasas de restricción específicas presentes en las bacterias clostridiales de interés. Por ejemplo, con el fin de transformar una especie bacteriana clostridial con un plásmido de interés, tratar el plásmido deseado *in vivo* o *in vitro* con una metiltransferasa heteróloga (por ejemplo, con la metiltransferasa del fago de *Bacillus subtilis*  $\Phi$ 3T) solo protegerá el plásmido de la escisión si la endonucleasa de restricción dentro de la célula huésped tiene la misma secuencia de reconocimiento de ADN que la metiltransferasa heteróloga. Para mejorar la efectividad de un enfoque de este tipo, pueden usarse múltiples metiltransferasas heterólogas, cada una con diferentes secuencias de reconocimiento de ADN; sin embargo, esto aumenta el tiempo y el costo de cada intento de transformación. Si las metiltransferasas usadas no reconocen la misma secuencia que la endonucleasa de restricción presente dentro de la célula clostridial de interés, el ADN heterólogo no estará protegido de la escisión.

Por consiguiente, sigue existiendo la necesidad de identificar y evadir los sistemas de restricción-modificación en bacterias clostridiales para facilitar su uso en la producción de bioproductos industriales incluyendo, pero sin limitarse a, isopreno, butadieno y etanol.

A lo largo de la memoria descriptiva, se divulgan diversas publicaciones (incluyendo secuencias), patentes y solicitudes de patentes.

#### Breve resumen de la invención

La presente divulgación proporciona, entre otros, la dilucidación de un sistema de restricción-modificación específico en bacterias clostridiales (por ejemplo, *Clostridium acetivum*) que escinde en el sitio CCWGG (W puede ser A o T) y metiltransferasas que pueden usarse para proteger frente a la escisión, tal como se describe adicionalmente en el presente documento. El conocimiento sobre este sistema de restricción-modificación permite modificar por ingeniería genética bacterias clostridiales, lo que permite la producción biológica de diversos productos industriales (por ejemplo, bioproductos).

La presente invención proporciona:

[1]. Un polinucleótido aislado que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en el que el polinucleótido codifica para un polipéptido con actividad metiltransferasa.

[2]. El polinucleótido de [1],

(a) en el que el polinucleótido es SEQ ID NO: 2; o

(b) en el que el polipéptido codificado metila un polinucleótido en una secuencia que comprende CCWGG, en el que opcionalmente la secuencia que comprende CCWGG se selecciona del grupo que consiste en CCAGG (SEQ ID NO: 9) y/o CCTGG (SEQ ID NO: 10); o

(c) en el que el polipéptido codificado metila un polinucleótido a SEQ ID NO: 9 y/o SEQ ID NO: 10.

[3]. Un plásmido que comprende el polinucleótido de uno cualquiera de [1]-[2], operativamente unido a una o más secuencias de control de manera que el polipéptido codificado es capaz de expresarse en un huésped de expresión.

[4]. El plásmido de [3],

(a) en el que el huésped de expresión es *E. coli*; y/o

(b) en el que dicho plásmido comprende además SEQ ID NO: 14.

[5]. El plásmido de uno cualquiera de [3]-[4], en el que dicho plásmido se transforma dentro de una célula de *E. coli* S17-1.

[6]. Una célula huésped recombinante que comprende

- (a) uno cualquiera de los polinucleótidos de uno cualquiera de [1]-[2] como ácido nucleico heterólogo; o
- (b) uno cualquiera de los plásmidos de uno cualquiera de [3]-[4].

[7]. Un polipéptido aislado que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 3, en el que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende CCWGG.

[8]. El polipéptido de [7],

- (a) en el que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10; o
- (b) en el que dicho polipéptido es capaz de metilar un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10; o
- (c) en el que dicho polipéptido es SEQ ID NO 3.

[9]. Un polipéptido aislado producido por uno cualquiera de los polinucleótidos de uno cualquiera de [1]-[2], en el que el polipéptido tiene actividad metiltransferasa.

[10]. Un método de producción de una ADN metiltransferasa, que comprende: (a) cultivar una célula huésped recombinante que comprende el polinucleótido de uno cualquiera de [1]-[2] en condiciones adecuadas para la producción de la ADN metiltransferasa codificada, y (b) recuperar la ADN metiltransferasa.

[11]. Un método de producción de un transformante bacteriano de *Clostridium* recombinante, que comprende:

- a) introducir el polinucleótido de uno cualquiera de [1]-[2] o el plásmido de uno cualquiera de [3]-[4] que codifica para una ADN metiltransferasa en una célula huésped bacteriana de *Escherichia*,
- b) cultivar la célula huésped bacteriana de *Escherichia* en condiciones adecuadas para la expresión de la ADN metiltransferasa, y
- c) transferir un polinucleótido metilado de la célula huésped bacteriana de *Escherichia* a una célula huésped bacteriana de *Clostridium*, en el que las bacterias transformadas usando este método se seleccionan del grupo que consiste en *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum* y *Clostridium autoethanogenum*.

[12]. El método de [11], en el que el polinucleótido metilado codifica para una enzima isopreno sintasa.

[13]. El método de uno cualquiera de [11]-[12], en el que el polinucleótido metilado comprende uno o más polinucleótidos heterólogos que codifican para una o más enzimas de la ruta del isopreno expresadas en una cantidad suficiente para producir isopreno, y en el que las bacterias clostridiales contienen una o más rutas para la producción de isopreno.

[14]. El método de [12], en el que la célula huésped bacteriana de *Clostridium* se ha modificado por ingeniería genética para producir isopreno a partir de sintegas y/o a partir de hidratos de carbono o mezclas de los mismos.

Por consiguiente, en un aspecto, la invención proporciona polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en la que los polinucleótidos codifican para un polipéptido con actividad metiltransferasa. En cualquiera de las realizaciones descritas en el presente documento, el polinucleótido es SEQ ID NO: 2. En cualquiera de las realizaciones descritas en el presente documento, el polipéptido codificado metila un polinucleótido en una secuencia que comprende CCWGG. En cualquiera de las realizaciones descritas en el presente documento, la secuencia que comprende CCWGG se selecciona del grupo que consiste en CCAGG (SEQ ID NO: 9) y/o CCTGG (SEQ ID NO: 10). En cualquiera de las realizaciones descritas en el presente documento, el polipéptido codificado metila un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10.

En otro aspecto, la invención proporciona plásmidos que comprenden uno o más polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, operativamente unidos a una o más secuencias de control de manera que el polipéptido codificado puede expresarse en un huésped de expresión. En cualquiera de las realizaciones descritas en el presente documento, el huésped de expresión es *E. coli*. En cualquiera de las realizaciones descritas en el presente documento, el plásmido comprende además SEQ ID NO: 14. En cualquiera de las realizaciones descritas en el presente documento, el plásmido se transforma en una células de *E. coli* S17-1.

En otro aspecto, la invención proporciona células huésped recombinantes que comprenden polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en la que los polinucleótidos heterólogos codifican para un polipéptido con actividad metiltransferasa.

5 En otro aspecto, la invención proporciona células huésped recombinantes que comprenden plásmidos que comprenden uno o más polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, operativamente unidos a una o más secuencias de control de manera que el polipéptido codificado puede expresarse en un huésped de expresión.

10 En otro aspecto, la invención proporciona polipéptidos aislados que comprenden una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 3, en la que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende CCWGG. En cualquiera de las realizaciones descritas en el presente documento, el polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10. En cualquiera de las realizaciones descritas en el presente documento, el polipéptido es capaz de metilar un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10.

15 En otro aspecto, la invención proporciona polipéptidos aislados que comprenden una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 3, en la que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende CCWGG. En cualquiera de las realizaciones descritas en el presente documento, el polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10. En cualquiera de las realizaciones descritas en el presente documento, el polipéptido es SEQ ID NO: 3.

20 En otro aspecto, la invención proporciona polipéptidos aislados producidos por polinucleótidos que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en la que el polipéptido tiene actividad metiltransferasa.

25 En otro aspecto, la invención proporciona métodos de producción de una ADN metiltransferasa, que comprende: (a) cultivar una célula huésped recombinante que comprende polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en los que los polinucleótidos codifican para un polipéptido con actividad metiltransferasa, en los que la célula huésped se cultiva en condiciones adecuadas para la producción de la ADN metiltransferasa codificada, y (b) recuperar la ADN metiltransferasa.

30 En otro aspecto, la invención proporciona métodos de producción de un transformante bacteriano de *Clostridium* recombinante, que comprende: introducir un polinucleótido que codifica para una ADN metiltransferasa en una célula huésped bacteriana de *Escherichia*, (a) cultivar la célula huésped bacteriana de *Escherichia* en condiciones adecuadas para la expresión de la ADN metiltransferasa, (b) transferir el polinucleótido metilado desde la célula huésped bacteriana de *Escherichia* hasta una célula huésped bacteriana de *Clostridium*, en los que las bacterias transformadas usando este método se seleccionan del grupo que consiste en *Clostridium acetivum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum* y *Clostridium autoethanogenum*.

35 En otro aspecto, la presente divulgación proporciona polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 4, en la que el polinucleótido codifica para un polipéptido con actividad endonucleasa. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el polipéptido codificado es capaz de escindir un polinucleótido en una secuencia que comprende CCWGG. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el polipéptido codificado es capaz de escindir un polinucleótido en una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el polipéptido codificado es capaz de escindir un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el polinucleótido es SEQ ID NO: 4.

40 En otro aspecto, la presente divulgación proporciona plásmidos que comprenden polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 4, en la que el polinucleótido codifica para un polipéptido con actividad endonucleasa, y en la que el plásmido está operativamente unido a una o más secuencias de control de manera que el polipéptido codificado puede expresarse en un huésped de expresión. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el polipéptido codificado es capaz de expresarse en un huésped de expresión de *E. coli*.

45 En otro aspecto, la divulgación proporciona células huésped recombinantes que comprenden polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 4, en la que el polinucleótido codifica para un polipéptido con actividad endonucleasa.

50 En otro aspecto, la divulgación proporciona células huésped recombinantes que comprenden plásmidos que comprenden polinucleótidos aislados que tienen una identidad de secuencia de al menos el 90% con SEQ ID NO: 4, en la que el polinucleótido codifica para un polipéptido con actividad endonucleasa, y en la que el plásmido está

operativamente unido a una o más secuencias de control de manera que el polipéptido codificado puede expresarse en un huésped de expresión.

5 En otro aspecto, la divulgación proporciona un método de reducción de la escisión por endonucleasa de un ácido nucleico heterólogo en una célula huésped de *Clostridium*, comprendiendo el método metilar una secuencia que comprende CCWGG. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el método comprende metilar una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10 en el ácido nucleico heterólogo. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, el método comprende metilar SEQ ID NO: 9 y/o SEQ ID NO: 10. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, la endonucleasa tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 5. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, la endonucleasa es SEQ ID NO: 5. En cualquiera de las realizaciones de la divulgación descrita en el presente documento, la metiltransferasa es SEQ ID NO: 3.

15 En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pDW280 (SEQ ID NO: 15).

En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS537 (SEQ ID NO: 16).

20 En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS200 (SEQ ID NO: 17).

En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS201 (SEQ ID NO: 18).

25 En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS444 (SEQ ID NO: 19).

30 En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS445 (SEQ ID NO: 20).

En otro aspecto, la presente divulgación proporciona un plásmido lanzadera que comprende pMCS94 (SEQ ID NO: 22).

35 En otro aspecto, la presente divulgación proporciona un plásmido que comprende pMCS466 (SEQ ID NO: 23).

En otro aspecto, la presente divulgación proporciona métodos para el suministro de uno o más ácidos nucleicos de interés al interior de una célula bacteriana de *Clostridium*, comprendiendo los métodos las etapas de:

40 transformar conjuntamente una célula de *E. coli* con:

un plásmido que comprende un polinucleótido que codifica para un polipéptido con actividad metiltransferasa, y

45 al menos un plásmido lanzadera seleccionado del grupo de pDW280, pMCS537, pMCS200, pMCS201, pMCS444 o pMCS445, en el que el plásmido lanzadera comprende además el uno o más ácidos nucleicos de interés;

50 cultivar la célula de *E. coli* de la etapa (a) con una célula bacteriana de *Clostridium* en condiciones que permiten la transferencia conjugativa de (a)(1) y (a)(2), suministrando de ese modo uno o más ácidos nucleicos al interior de una célula bacteriana de *Clostridium*.

55 En cualquier realización de la divulgación descrita en el presente documento, la célula bacteriana de *Clostridium* se selecciona del grupo que consiste en: *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum* y *Clostridium autoethanogenum*. En cualquier realización de la divulgación descrita en el presente documento, la célula de *E. coli* es de la cepa S17-1.

En otro aspecto, la presente divulgación proporciona células bacterianas de *Clostridium* recombinantes que comprenden:

60 a) un plásmido que comprende pDW268 (SEQ ID NO: 14), y

b) al menos un plásmido lanzadera seleccionado del grupo de pDW280 (SEQ ID NO: 15), pMCS537 (SEQ ID NO: 16), pMCS200 (SEQ ID NO: 17), pMCS201 (SEQ ID NO: 18), pMCS444 (SEQ ID NO: 19) o pMC4245 (SEQ ID NO: 20), en el que el plásmido lanzadera comprende además uno o más ácidos nucleicos de interés.

65 En otro aspecto, la presente divulgación proporciona células bacterianas de *Clostridium* recombinantes producidas:

(a) transformando conjuntamente una célula de *E. coli* con: (1) un plásmido que comprende un polinucleótido que codifica para un polipéptido con actividad metiltransferasa y (2) al menos un plásmido lanzadera seleccionado del grupo de pDW280, pMCS537, pMCS200, pMCS201, pMCS444 o pMCS445, en el que el plásmido lanzadera comprende además el uno o más ácidos nucleicos de interés; (b) cultivando la célula de *E. coli* de la etapa (a) con una célula bacteriana de *Clostridium* en condiciones que permiten la transferencia conjugativa de (a)(1) y (a)(2), suministrando de ese modo uno o más ácidos nucleicos al interior de una célula bacteriana de *Clostridium*.

En otro aspecto, la presente divulgación proporciona sistemas de expresión de *Clostridium* para la expresión de uno o más ácidos nucleicos de interés, comprendiendo el sistema:

a) un plásmido que comprende pDW268 (SEQ ID NO: 14),

b) un plásmido lanzadera seleccionado del grupo de pDW280 (SEQ ID NO: 15), pMCS537 (SEQ ID NO: 16), pMCS200 (SEQ ID NO: 17), pMCS201 (SEQ ID NO: 18), pMCS444 (SEQ ID NO: 19) o pMC4245 (SEQ ID NO: 20), en el que el plásmido lanzadera comprende además uno o más ácidos nucleicos de interés para la expresión,

c) una célula bacteriana de *Escherichia* capaz de interactuar con una célula bacteriana de *Clostridium* para permitir la transferencia de (a) y (b); y

d) una célula bacteriana de *Clostridium* capaz de interactuar con una célula bacteriana de *Escherichia* de manera que el uno o más ácidos nucleicos se expresan en la célula bacteriana de *Clostridium*.

En cualquier realización de la divulgación descrita en el presente documento, la célula bacteriana de *Clostridium* se selecciona del grupo que consiste en *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum* y *Clostridium autoethanogenum*. En cualquier realización de la divulgación descrita en el presente documento, la célula bacteriana de *Clostridium* es *Clostridium aceticum*.

#### Breve descripción de los dibujos

La figura 1 muestra la secuencia de ADN con codones optimizados (1422 pb) de una ADN metiltransferasa de *Clostridium aceticum* (M.Cacl), optimizada para la expresión en *E. coli* (SEQ ID NO: 1).

La figura 2 muestra la secuencia de ADN de tipo natural (1425 pb) para una metiltransferasa de *Clostridium aceticum* (M.Cacl, RYBO02455) (SEQ ID NO. 2).

La figura 3 muestra la secuencia de aminoácidos deducida (474 aa) para una ADN metiltransferasa de *Clostridium aceticum* (M.Cacl) (SEQ ID NO. 3).

La figura 4A muestra la secuencia de ADN de tipo natural (714 pb) de una endonucleasa de restricción de la cepa de *Clostridium aceticum* ATCC35044 (Cacl, RYBO02454) (SEQ ID NO. 4). La figura 4B y la figura 4C muestran la ubicación genómica y las anotaciones de M.Cacl (figura 4B, RYBO02455 - SEQ ID NO. 2) y Cacl (figura 4C, RYBO02454; SEQ ID NO. 4), respectivamente, en la cepa de *Clostridium aceticum* ATCC35044. M.Cacl y Cacl están ubicadas adyacentes entre sí, pero en hebras opuestas del cromosoma de *C. aceticum* (figura 4B - C). La figura 4C también muestra que Cacl (flecha dentro del círculo) se anotó erróneamente como glicosil hidrolasa mediante la base de datos de Genbank.

La figura 5 muestra la secuencia de aminoácidos deducida (237 aa) para la endonucleasa de restricción de *Clostridium aceticum* Cacl (SEQ ID NO. 5).

La figura 6 muestra el mapa de plásmido para pCA1.

Las figuras 7A - B muestran la secuencia de ADN de pCA1 (5720 pb) (SEQ ID NO. 6).

La figura 8 muestra el mapa de plásmido para pMCS203.

La figura 9 muestra la secuencia de ADN de pMCS203 (3729 pb) (SEQ ID NO. 7).

La figura 10A muestra el mapa de plásmido para pMCS244. La figura 10B muestra el mapa de plásmido de pMCS244 con flechas que indican las ubicaciones aproximadas de sus cuatro sitios de restricción Cacl (marcados con flechas en negrita). Dos sitios Cacl en la región de ARN de ColE1 11 del plásmido están ubicados próximos entre sí, y se representan mediante sólo una flecha.

La figura 11 muestra la secuencia de ADN de pMCS244 (3270 pb) (SEQ ID NO. 8).

La figura 12A muestra los resultados de un ensayo de endonucleasas de restricción usando 500 ng de pMCS244 tratado con 1 µl de lisado de *Clostridium aceticum*, 1 µl de la endonucleasa de restricción HindIII, 1 µl de la

endonucleasa de restricción *Apa*LI, o las combinaciones indicadas de las mismas. De izquierda a derecha, carril 1: marcador de peso molecular X de ADN de Roche, carril 2: pMCS244 sin cortar, carril 3: pMCS244 y lisado de *Clostridium acetikum*, carril 4: pMCS244 y *Hind*III, carril 5: pMCS244 y *Apa*LI, carril 6: pMCS244 con lisado de *Clostridium acetikum* y *Hind*III, carril 7: pMCS244 con lisado de *Clostridium acetikum* y *Apa*LI, carril 8: pMCS244 con lisado de *Clostridium acetikum*, carril 9: pMCS44 con lisado de *Clostridium acetikum*, *Hind*III y *Apa*LI combinados.

La figura 12B muestra los resultados de un ensayo de mapeo de precisión. Carril 1: 500 ng de pMCS244 y 1 µl de lisado de *Clostridium acetikum*, carril 2: marcador de peso molecular X de ADN de Roche, carril 3: 100 ng de producto de PCR lineal generado a partir de pMCS244 usando los cebadores M13R y oMCS25; carril 4: 100 ng de producto de PCR lineal generado a partir de pMCS244 usando los cebadores M13R y oMCS25 y 1 µl de *Hind*III, carril 5: 100 ng de producto de PCR lineal generado a partir de pMCS244 usando los cebadores M13R y oMCS25 y 1 µl de lisado de *Clostridium acetikum*.

La figura 13 muestra la secuencia de reconocimiento de endonucleasa de restricción de tipo II CCWGG (W = T o A) que es proximal al sitio de escisión *Hind*III en un producto de PCR lineal generado a partir de pMCS244 usando los cebadores M13R y oMCS25. Tanto CCAGG (SEQ ID NO. 9) como CCTGG (SEQ ID NO. 10) se reconocen por *Cacl* y *M.Cacl*.

La figura 14 muestra el mapa de plásmido para pDW265, con las ubicaciones de los cuatro sitios de reconocimiento de ADN de *Cacl* mutados indicados en el mismo.

La figura 15 muestra la secuencia de ADN (3270 pb) para pDW265 (SEQ ID NO. 11).

La figura 16 muestra los resultados de un ensayo de endonucleasas de restricción usando 500 ng del plásmido de control pMCS244 o 500 ng del plásmido pDW265 (que tiene los cuatro sitios de reconocimiento de ADN de *Cacl* mutados) tratados con 1 µl de lisado de *Clostridium acetikum*, 1 µl de *Hind*III, o ambos. Carril 1: Marcador de peso molecular X de ADN de Roche, carril 2: plásmido de control pMCS244; carril 3: plásmido pDW265 no tratado; carril 4: pMCS244 de control tratado con lisado de *C. acetikum*; carril 5: plásmido pDW265 tratado con lisado de *C. acetikum*; carril 6: pMCS244 tratado con *Hind*III; carril 7: pDW265 tratado con *Hind*III; carril 8: plásmido pMCS244 tratado con tanto lisado de *C. acetikum* como *Hind*III; carril 9: pDW265 con tanto lisado de *C. acetikum* como *Hind*III.

La figura 17A muestra los resultados cuando pMCS244, que contiene 4 sitios de secuencia de reconocimiento de *Cacl*, se incubaba con lisado de *Clostridium acetikum* y luego se transforma en *E. coli*. La figura 17B muestra los resultados cuando pDW265, que es idéntico a pMCS244 excepto porque los cuatro sitios *Cacl* se han mutado, se incubaba con lisado de *Clostridium acetikum* y luego se transforma en *E. coli*.

La figura 18 muestra productos de PCR amplificados a partir de plásmidos aislados de una cepa de *Clostridium acetikum* transformada de manera conjugativa usando los cebadores oMCS418 a oMCS423 (tabla 4), que confirman la presencia de toda la secuencia heteróloga (en pDW280), el origen de replicación de *Clostridium acetikum* y el casete de resistencia a eritromicina, respectivamente.

La figura 19 muestra el mapa de plásmido para pDW263.

Las figuras 20A - C muestran la secuencia de ADN (8285 pb) para pDW263 (SEQ ID NO. 12).

La figura 21 muestra el mapa de plásmido para pDW264.

Las figuras 22A - C muestran la secuencia de ADN (8285) para pDW264 (SEQ ID NO 13).

La figura 23 muestra el mapa de plásmido para pDW268.

Las figuras 24A - C muestran la secuencia de ADN (6758 pb) para pDW268 (SEQ ID NO 14).

La figura 25 muestra los resultados cuando el plásmido pDW265 se incubaba con lisado de *Clostridium acetikum* y se transforma dentro de *E. coli*.

La figura 26 muestra los resultados cuando se incubaba pMCS244 sin metilar con lisado de *Clostridium acetikum* y se transforma dentro de *E. coli*.

La figura 27 muestra un mapa de plásmido para pDW280.

Las figuras 28A - C muestran la secuencia de ADN (8398 pb) para pDW280 (SEQ ID NO 15).

La figura 29 muestra bacterias de *Clostridium acetikum* sometidas a múltiples pases que se hacen crecer sobre medios AcM con ácido nalidíxico 10 µg/ml y eritromicina 20 µg/ml después de la conjugación satisfactoria con células de *E. coli* S17-1 que albergan los plásmidos pDW268 y pDW280.

La figura 30 muestra un mapa de plásmido para pMCS537.

Las figuras 31A-B muestran la secuencia de ADN para pMCS537 (SEQ ID NO. 16).

La figura 32 muestra el mapa de plásmido para pMCS200, también denominado pMTL82151.

Las figuras 33A-B muestran la secuencia de ADN (5254 pb) para pMCS200 (SEQ ID NO. 17).

La figura 34 muestra el mapa de plásmido para pMCS201, también denominado pMTL83151.

Las figuras 35A-B muestran la secuencia de ADN (4476 pb) para pMCS201 (SEQ ID NO. 18).

Las figuras 36A-B muestran los resultados de ensayos para determinar la concentración inhibitoria mínima de los antibióticos tiamfenicol (Thi) y eritromicina (Em) para *Clostridium acetivum* que se hace crecer en cultivo líquido (figura 36A) o para *Clostridium acetivum* que se hace crecer sobre placas de medios de crecimiento de *Clostridium acetivum* (medios AcM) (figura 36B).

La figura 37 muestra los resultados de la titulación de fructosa para *Clostridium acetivum*, demostrando que 10 g/l de fructosa no era limitativo, y que la fructosa solo se vuelve limitativa en concentraciones de menos de ~ 1,5 g/l.

La figura 38A-B muestra los resultados de ensayos para determinar la concentración inhibitoria mínima de los antibióticos tiamfenicol (Thi) y eritromicina (Em) para *Clostridium ljungdahlii* que se hace crecer en cultivo líquido (figura 38A) o para *Clostridium ljungdahlii* hecha crecer sobre las placas (figura 38b).

La figura 39 muestra los combustibles microbianos que pueden producirse a partir de sintegas por medio de rutas celulares.

La figura 40 muestra las rutas de MVA clásica y modificada. 1, acetil-CoA acetiltransferasa (AACT); 2, HMG-CoA sintasa (HMGS); 3, HMG-CoA reductasa (HMGR); 4, mevalonato cinasa (MVK); 5, fosfomevalonato cinasa (PMK); 6, difosfomevalonato descarboxilasa (MVD o DPMDC); 7, isopentenil difosfato isomerasa (IDI); 8, fosfomevalonato descarboxilasa (PMDC); 9, isopentenil fosfato cinasa (IPK). La ruta clásica de MVA avanza de la reacción 1 a la reacción 7 a través de las reacciones 5 y 6, mientras que una ruta de MVA modificada pasa a través de las reacciones 8 y 9. P y PP en la fórmula estructural son fosfato y pirofosfato, respectivamente. Esta figura se tomó de Koga y Morii, *Microbiology and Mol. Biology Reviews*, 71: 97-120, 2007, a la que se hace referencia, particularmente con respecto a ácidos nucleicos y polipéptidos de la ruta de MVA modificada. La ruta de MVA modificada está presente, por ejemplo, en algunos organismos arqueas, tales como *Methanosarcina mazei*.

La figura 41 muestra una representación esquemática de un anaerobio obligado que expresa (a) un polipéptido de IspS heterólogo, (b) un polipéptido de DXS heterólogo y (c) un polipéptido de IDI heterólogo para aumentar el flujo de la ruta de DXP y la producción de isopreno.

La figura 42 muestra una representación esquemática de un anaerobio obligado modificado por ingeniería genética con *mvaE* y *mvaS* para expresar la ruta de MVA superior.

La figura 43 muestra una representación esquemática de la expresión de la ruta de MVA inferior en un anaerobio obligado que incluye expresar (a) un polipéptido de MVK heterólogo, (b) un polipéptido de PMK heterólogo y (c) un polipéptido de MVD heterólogo en las células que expresan un polipéptido de IDI heterólogo y polipéptido de IspS heterólogo con el propósito de aumentar la producción de isopreno.

La figura 44 muestra una representación esquemática de la expresión de la ruta completa de MVA en un anaerobio obligado mediante la introducción de *mvaE* y *mvaS* en las células que expresan (a) un polipéptido de MVK heterólogo, (b) un polipéptido de PMK heterólogo, (c) un polipéptido de MVD heterólogo, (d) un polipéptido de IDI heterólogo y (e) un polipéptido de IspS heterólogo con el fin de aumentar la producción de isopreno.

La figura 45 muestra una representación esquemática de la redirección del flujo de carbono lejos del acetato reduciendo la expresión de *ack* y *adhE* para reducir la pérdida de carbono a los productos secundarios. Las flechas junto a *Ack* o *AdhE* usados en la producción de acetato y etanol, respectivamente, indican una reducción de la actividad o expresión enzimática para rutas que conducen a productos de fermentación tales como acetato, etanol o cualquier otro alcohol o producto final que contenga carbono. El propósito es maximizar la canalización de carbono hacia el isopreno a través de manipulación genética.

La figura 46 muestra rutas a modo de ejemplo para la producción de butadieno a partir de acetil-CoA, glutaconil-CoA, glutaril-CoA, 3-aminobutiril-CoA o 4-hidroxiobutiril-CoA a través de alcohol crofilico. Las enzimas para la transformación de los sustratos identificados en productos incluyen: A. acetil-CoA:acetil-CoA aciltransferasa, B. acetoacetil-CoA reductasa, C. 3-hidroxiobutiril-CoA deshidratasa, D. crotonil-CoA reductasa (que forma aldehído), E.

crotonaldehído reductasa (que forma alcohol), F. crotil alcohol cinasa, G. 2-butenil-4-fosfato cinasa, H. butadieno sintasa, I. crotonil-CoA hidrolasa, sintetasa, transferasa, J. crotonato reductasa, K. crotonil-CoA reductasa (que forma alcohol), L. glutaconil-CoA descarboxilasa, M., glutaril-CoA deshidrogenasa, N. 3-aminobutiril-CoA desaminasa, O. 4-hidroxiobutiril-CoA deshidratasa, P. crotil alcohol difosfocinasa.

5 La figura 47 muestra rutas a modo de ejemplo para la producción de butadieno a partir de eritrosa-4-fosfato. Las enzimas para la transformación de los sustratos identificados en productos incluyen: A. eritrosa-4-fosfato reductasa, B. eritritol-4-fosfato citidililtransferasa, C. 4-(citidina 5'-difosfo)-eritritol cinasa, D. eritritol 2,4-ciclodifosfato sintasa, E. 1-hidroxi-2-butenil 4-difosfato sintasa, F. 1-hidroxi-2-butenil 4-difosfato reductasa, G. butenil 4-difosfato isomerasa, H. butadieno sintasa, I. eritrosa-4-fosfato cinasa, J. eritrosa reductasa, K. eritritol cinasa.

15 La figura 48 muestra una ruta a modo de ejemplo para la producción de butadieno a partir de malonil-CoA más acetil-CoA. Las enzimas para la transformación de los sustratos identificados en productos incluyen: A. malonil-CoA:acetil-CoA aciltransferasa, B. 3-oxoglutaril-CoA reductasa (que reduce cetona), C. 3-hidroxioglutaril-CoA reductasa (que forma aldehído), D. 3-hidroxi-5-oxopentanoato reductasa, E. 3,5-dihidroxipentanoato cinasa, F. 3H5PP cinasa, G. 3H5PDP descarboxilasa, H. butenil 4-difosfato isomerasa, I. butadieno sintasa, J. 3-hidroxioglutaril-CoA reductasa (que forma alcohol), K. 3-oxoglutaril-CoA reductasa (que forma aldehído), L. 3,5-dioxopentanoato reductasa (que reduce cetona), M. 3,5-dioxopentanoato reductasa (que reduce aldehído), N. 5-hidroxi-3-oxopentanoato reductasa, O. 3-oxo-glutaril-CoA reductasa (que reduce CoA y que forma alcohol). Las abreviaturas de los compuestos incluyen: 3H5PP = 3-hidroxi-5-fosfonatooxipentanoato y 3H5PDP = 3-hidroxi-5-[hidroxi(fosfonooxi)fosforil]oxipentanoato.

La figura 49 muestra el mapa de plásmido para el plásmido pMCS444.

25 La figura 50 muestra la secuencia de ADN (5367 pb) para pMCS444.

La figura 51 muestra el mapa de plásmido para el plásmido pMCS445.

30 La figura 52 muestra la secuencia de ADN (4589 pb) para pMCS445.

La figura 53 muestra el mapa de plásmido para el plásmido PMCljs.

La figura 54 muestra la secuencia de ADN para pMCljs (7571 pb).

35 La figura 55 muestra el mapa de plásmido para pMCS94.

La figura 56 muestra la secuencia de ADN para pMCS94 (5056 pb).

40 La figura 57 muestra el mapa de plásmido para pMCS466.

La figura 58 muestra la secuencia de ADN para pMCS466 (6334 pb).

### Descripción detallada de la invención

45 La presente divulgación proporciona, entre otros, la dilucidación de un sistema de restricción-modificación específico en bacterias clostridiales (por ejemplo, *Clostridium acetivum*) que escinde en el sitio CCWGG (W puede ser A o T) y metiltransferasas que pueden usarse para proteger frente a la escisión, tal como se describe adicionalmente en el presente documento. El conocimiento sobre este sistema de restricción-modificación permite la modificación por ingeniería genética de bacterias clostridiales, lo que permite la producción biológica de diversos productos industriales (por ejemplo, bioproductos).

#### Técnicas generales

55 La práctica de la presente invención empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular (incluyendo técnicas recombinantes), microbiología, biología celular, bioquímica e inmunología, que están entro del conocimiento de la técnica. Tales técnicas se explican completamente en la bibliografía, tal como Handbook on Clostridia (P. Durre, ed., 2004), Biotechnology: A Textbook of Industrial Microbiology (Brock, Sinauer Associates, Inc., segunda edición, 1989), Molecular Cloning: A Laboratory Manual (Sambrook *et al.*, 1989, 2ª ed.); Oligonucleotide Synthesis (M.J. Gait, ed., 1984); Methods in Enzymology (Academic Press, Inc.); Current Protocols in Molecular Biology (F.M. Ausubel *et al.*, eds., 1987, y actualizaciones periódicas); PCR: The Polimerase Chain Reaction (Mullis *et al.*, eds., 1994), Dictionary of Microbiology and Molecular Biology (Singleton *et al.*, 2ª ed., J. Wiley y Sons, Nueva York, NY, 1994); y Advanced Organic Chemistry Reactions, Mechanisms and Structure (marzo, 4ª ed., John Wiley y Sons, Nueva York, NY, 1992), que proporcionan a un experto en la técnica orientación general con respecto a muchos de los términos y métodos usados en la presente divulgación.

65 A menos que se defina otra cosa, los términos técnicos y científicos usados en el presente documento tienen el

mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta invención.

#### Definiciones

- 5 “Isopreno” se refiere a 2-metil-1,3-butadieno (n.º CAS 78-79-5). Puede referirse al producto de hidrocarburo C5 volátil directo y final de la eliminación de pirofosfato del pirofosfato de 3,3-dimetilalilo (DMAPP). Puede que no implique la unión o polimerización de una o más moléculas de difosfato de isopentenilo (IPP) a una o más moléculas de DMAPP. El isopreno no está limitado por el método de su fabricación.
- 10 Los “bioproductos industriales” pueden incluir, pero no se limitan a, isopreno, isoprenoides, precursores de isoprenoides, butadieno y etanol. Los productos industriales también pueden incluir, pero no se limitan a, bioproductos derivados directa o indirectamente de 2-cetoácidos, malonil-CoA y acetoacetil-CoA. Los bioproductos industriales también pueden incluir, pero sin limitarse a, monoterpenos, diterpenos, triterpenos, tetraterpenos, sesquiterpeno, politerpeno, abietadieno, amorfadieno, careno,  $\alpha$ -farneseno,  $\beta$ -farneseno, farnesol, geraniol, geranilgeraniol, linalool, limoneno, mirceno, nerolidol, ocimeno, patchulol,  $\beta$ -pineno, sabineno,  $\gamma$ -terpineno, terpendeno, valenceno. Los bioproductos industriales pueden incluir además, pero sin limitarse a, alcoholes no fermentativos (por ejemplo, 1-propanol, 1-butanol, isobutanol, 2-metil-1-butanol, 3-metil-1-butanol, 3-metil-1-pentanol, 4-metil-1-pentanol y 1-hexanol), hidrocarburos derivados de ácidos grasos (alcoholes grasos, ésteres grasos, olefinas y alcanos) y alcoholes fermentativos (por ejemplo, butanol).
- 15
- 20 Un “ácido nucleico” o “polinucleótido” se refiere a dos o más desoxirribonucleótidos y/o ribonucleótidos en forma o bien monocatenaria o bien bicatenaria.
- 25 Un “ácido nucleico de interés” se refiere a un polinucleótido que codifica para un polipéptido que forma parte de la ruta de síntesis para cualquier producto industrial.
- Un “ácido nucleico endógeno” es un ácido nucleico cuya secuencia de ácido nucleico se encuentra de manera natural en la célula huésped. En algunos aspectos, un ácido nucleico endógeno es idéntico a un ácido nucleico de tipo natural que se encuentra en la célula huésped en la naturaleza. En algunos aspectos, se introducen una o más copias de ácidos nucleicos endógenos en una célula huésped.
- 30
- Un “ácido nucleico heterólogo” puede ser un ácido nucleico cuya secuencia de ácido nucleico es de otra especie distinta de la célula huésped u otra cepa de la misma especie que la célula huésped. En algunos aspectos, la secuencia no es idéntica a la de otro ácido nucleico que se encuentra de manera natural en la misma célula huésped. En algunos aspectos, un ácido nucleico heterólogo no es idéntico a un ácido nucleico de tipo silvestre que se encuentra en la misma célula huésped en la naturaleza. En diversas realizaciones de la invención, un ácido nucleico heterólogo codifica para uno o más bioproductos industriales.
- 35
- “Polipéptidos” incluye polipéptidos, proteínas, péptidos, fragmentos de polipéptidos, polipéptidos de fusión y variantes.
- 40
- Un “polipéptido endógeno” es un polipéptido cuya secuencia de aminoácidos se encuentra de manera natural en la célula huésped. En algunos aspectos, un polipéptido endógeno es idéntico a un polipéptido de tipo silvestre que se encuentra en la célula huésped en la naturaleza.
- 45
- Un “polipéptido heterólogo” es un polipéptido codificado por un ácido nucleico heterólogo. En algunos aspectos, la secuencia no es idéntica a la de otro polipéptido codificado por un ácido nucleico que se encuentra de manera natural en la misma célula huésped.
- 50
- A menos que se defina lo contrario en el presente documento, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto en la técnica a la que se refiere esta invención.
- 55
- Tal como se usa en el presente documento, los términos singulares “un”, “una” y “el/la” incluyen la referencia en plural a menos que el contexto indique claramente lo contrario.

#### Identificación de un sistema de restricción-modificación clostridial

- 60 Los inventores han descubierto un sistema de restricción-modificación (R-M) específico en bacterias clostridiales. En un aspecto, el sistema de R-M está en *Clostridium acetivum* que reconoce la secuencia CCWGG donde W puede ser A o T. Antes de este descubrimiento, este sistema de R-M era una barrera importante para la introducción de ácidos nucleicos heterólogos en bacterias clostridiales (por ejemplo, *Clostridium acetivum*). Los ácidos nucleicos heterólogos pueden codificar para la producción de productos industriales deseados en las bacterias clostridiales. Sin embargo, algunos de los desafíos de tratar de producir biológicamente productos industriales en bacterias clostridiales fueron que los ácidos nucleicos heterólogos se digirieron por endonucleasas endógenas en la célula bacteriana clostridial o se vieron afectados adversamente de otra forma de manera que el bioproducto industrial
- 65

deseado no pudo producirse. La presente divulgación proporciona la identificación del sitio de restricción para una endonucleasa, endonucleasas que pueden unirse al sitio de restricción y metiltransferasas que pueden proteger frente a la escisión no deseada de ácidos nucleicos de interés. Debe entenderse que las composiciones y/o sistemas, métodos de preparación y uso de estos aspectos y/o realizaciones se abarcan dentro del alcance de la presente divulgación.

#### Composiciones y métodos de uso

Como resultado de este descubrimiento, los inventores han creado (y describen en el presente documento) polinucleótidos, polipéptidos, plásmidos, vectores, sistemas de expresión, células huésped, etc. basándose en los componentes de este sistema de restricción-metilación clostridial, así como métodos de preparación y uso de estos componentes para facilitar la manipulación genética de bacterias clostridiales (por ejemplo, *Clostridium acetivum*, *Clostridium acetobutylicum*, *Clostridium ljungdahlii* y *Clostridium autoethanogenum*) para producir bioproductos industriales tales como (pero sin limitarse a) isopreno, butadieno y etanol.

#### Endonucleasas de restricción

La presente divulgación proporciona composiciones de endonucleasas de restricción específicas que actúan en células clostridiales escindiendo ácidos nucleicos y métodos de identificación y uso de las mismas. Se describen varias endonucleasas de restricción a modo de ejemplo en el presente documento y también en la sección de ejemplos (por ejemplo endonucleasa de restricción *Cacl*). Estas endonucleasas de restricción reconocen secuencias CCWGG (en donde W puede ser A o T). En una realización de la presente divulgación el polinucleótido y secuencia de aminoácidos divulgados de la endonucleasa de restricción *Cacl* puede usarse para identificar otras endonucleasas de restricción relacionadas con homología con *Cacl* que tienen la misma funcionalidad. En otra realización de la divulgación, la secuencia de ácido nucleico o secuencia de aminoácidos de *Cacl* puede usarse para diseñar una sonda de ácido nucleico para identificar y clonar ADN que codifica para polipéptidos que tienen actividad de endonucleasa de restricción a partir de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica.

Estos homólogos identificados pueden inactivarse entonces para facilitar la introducción de uno o más polinucleótidos de interés en la célula huésped. Tal como se usa en el presente documento, "homología" se refiere a similitud o identidad de secuencia, prefiriéndose la identidad. Esta homología se determina usando técnicas convencionales conocidas en la técnica (véase por ejemplo, Smith y Waterman, *Adv Appl Math*, 2:482, 1981; Needleman y Wunsch, *J Mol Biol*, 48:443, 1970; Pearson y Lipman, *Proc Natl Acad Sci USA*, 85:2444, 1988; programas tales como GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, Madison, WI; y Devereux *et al.*, *Nucl Acid Res*, 12:387-395, 1984).

La inactivación de endonucleasas de restricción puede lograrse a través de métodos bien conocidos en la técnica, tales como inserciones, interrupciones, reemplazos o deleciones de la totalidad o un segmento del/de los gen(es) de endonucleasa de restricción present(es) en la célula (por ejemplo, mediante técnicas de alteración genética para eliminar o reducir la expresión del gen, tal como el método Clostron basado en inserción de intrones del grupo II). Véase, por ejemplo, Dong *et al.*, *PLoS ONE* 2010 5(2):e9038. En Dong *et al.* (2010), una supuesta endonucleasa de restricción de tipo II (*Cac824I*), identificada a partir del genoma públicamente disponible de *Clostridium acetobutylicum* ATCC 824, se alteró usando el sistema de inactivación génica basado en inserción de intrones del grupo II Clostron. Las células resultantes deficientes en *Cac824I* pudieron transformarse con ADN no metilado (por ejemplo, ADN de plásmido no metilado) por medio de electroporación. El sistema Clostron, similar a la mayoría de los enfoques de intrones del grupo II, usa un elemento derivado del intrón LI.LtrB de intervalo de huésped amplio de *Lactococcus lactis*. Véase, por ejemplo, Kuehne *et al.*, 2011, Clostron-mediated engineering of Clostridium. *Methods in Molecular Biology*, vol. 765:389-407.

Puede usarse un enfoque de alteración génica similar para inactivar el gen *Cacl* en otras bacterias del género *Clostridium*, facilitando así la elusión de su(s) sistema(s) de restricción-modificación. Usando métodos bien conocidos en la técnica, (por ejemplo, programas de alineación de secuencias tales como BLAST o CLUSTAL W) pueden encontrarse homólogos de *Cacl* en otras bacterias clostridiales e inactivarse usando el sistema Clostron o un sistema de selección como diana de genes similar. La porción del gen inactivado puede ser, por ejemplo, la región codificante o un elemento regulador requerido para la expresión de la región codificante. Un ejemplo de una secuencia reguladora de este tipo puede ser una secuencia de promotor o parte funcional de la misma, por ejemplo, una parte que es suficiente para afectar a la expresión de la secuencia de nucleótidos. Otras secuencias de control para su posible modificación incluyen, pero no se limitan a, una secuencia líder, secuencia de propéptido, secuencia señal, terminador de la transcripción y activador transcripcional.

La inactivación de una endonucleasa de restricción puede lograrse también mediante mutagénesis al azar o específica usando mutagénesis química (véase, por ejemplo, Hopwood, *The Isolation of Mutants, Methods of Microbiology* (J.R. Norris y D.W. Ribbons, eds., págs. 363-433, Academic Press, Nueva York, 1970) y transposición (por ejemplo, Youngman *et al.*, 1983, *PNAS* 80: 2305-2309). La modificación del gen de endonucleasa de restricción puede realizarse sometiendo la célula original a mutagénesis y seleccionando células mutantes en las que la

expresión de la endonucleasa de restricción se ha reducido o eliminado. La mutagénesis, que puede ser específica o al azar, puede realizarse mediante, por ejemplo, el uso de un agente de mutagenización física o química adecuado, el uso de un oligonucleótido adecuado o sometiendo la secuencia de ADN a mutagénesis generada por PCR. Además, la mutagénesis puede realizarse mediante el uso de cualquier combinación estos métodos de mutagenización.

En otro aspecto, la endonucleasa clostridial puede usarse como diana para unir moléculas, tales como anticuerpos. Los anticuerpos frente a una endonucleasa clostridial pueden ser útiles como herramienta de investigación (por ejemplo, detección de la presencia de endonucleasa en lisados clostridiales), herramienta de laboratorio o herramienta médica.

#### *Modificación de sitios de reconocimiento de Cacl*

Pueden modificarse sitios de reconocimiento de *Cacl* de tal manera que las endonucleasas ya no los reconocen en células clostridiales. Estos sitios de reconocimiento de *Cacl* pueden estar en ácidos nucleicos de interés, por ejemplo, ácidos nucleicos heterólogos que codifican para diversos bioproductos industriales. En algunas realizaciones de la presente divulgación, la introducción de un polinucleótido de interés en una célula de *Clostridium* puede lograrse modificando el polinucleótido de interés para mutar o delecionar cualquier sitio de reconocimiento de ADN específico de *Cacl* (por ejemplo, mediante la mutación de cualquier secuencias de reconocimiento de ADN de *Cacl* CCWGG), de modo que el polinucleótido introducido no se degrada por la endonucleasa de restricción de la célula huésped bacteriana. En otras realizaciones de la presente divulgación, el polinucleótido de interés se modifica para mutar o delecionar una o más secuencias de reconocimiento de ADN de *Cacl* CCWGG. En otras realizaciones de la presente divulgación, el polinucleótido de interés se modifica para mutar o delecionar uno o más sitios CCAGG (SEQ ID NO: 9). En otras realizaciones de la presente divulgación, el polinucleótido de interés se muta para delecionar uno o más sitios CCTGG (SEC ID NO: 10).

La presencia de cualquier sitio *Cacl* en un polinucleótido de interés (por ejemplo, un plásmido lanzadera para su uso entre *E. coli* y una o más especies de *Clostridium* que contiene genes de la ruta de DXP para la síntesis de isopreno) puede determinarse usando métodos de secuenciación conocidos en la técnica o divulgados en el presente documento. La modificación del polinucleótido de interés puede lograrse mediante mutagénesis usando métodos bien conocidos en la técnica, que incluyen, pero no se limitan a, mutagénesis dirigida al sitio o mutagénesis generada por PCR. Véase, por ejemplo, Shimada, 1996, *Methods in Molecular Biology*, vol. 57: 157-165, que se incorpora por el presente documento en su totalidad, particularmente en lo que se refiere a mutagénesis dirigida al sitio.

El polinucleótido modificado puede contener una inserción, sustitución o delección de uno o más nucleótidos presentes en la secuencia de reconocimiento de ADN CCWGG. En algunas realizaciones de la divulgación, el polinucleótido de interés modificado puede contener una inserción, sustitución o delección de uno o más nucleótidos presentes en la secuencia de reconocimiento de ADN CCAGG (SEQ ID NO: 9). En algunas realizaciones de la divulgación, el polinucleótido de interés modificado puede contener una inserción, sustitución o delección de uno o más nucleótidos presentes en la secuencia de reconocimiento de ADN CCTGG (SEQ ID NO: 10). En algunas realizaciones, el polinucleótido de interés modificado puede contener una inserción, sustitución o delección de uno o más nucleótidos presentes en la secuencia de reconocimiento de ADN CCAGG (SEQ ID NO: 9) y puede contener una inserción, sustitución o delección de uno o más nucleótidos presentes en la secuencia de reconocimiento de ADN CCTGG (SEQ ID NO: 10), por ejemplo, como en el plásmido resistente a *Cacl* pDW265 divulgado en el ejemplo 6 de la presente solicitud. Además, la mutagénesis puede realizarse usando cualquier combinación de métodos de mutagenización.

#### *Metiltransferasas*

La divulgación también proporciona composiciones de metiltransferasas específicas que actúan en células clostridiales para proteger a los ácidos nucleicos de la escisión por endonucleasas así como métodos de identificación y uso de los mismos. En una realización de la divulgación, la secuencia de ácido nucleico o secuencia de aminoácidos de *M.Cacl* puede usarse para diseñar una sonda de ácido nucleico para identificar y clonar ADN que codifica para polipéptidos que tienen actividad metiltransferasa a partir de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica.

Pueden obtenerse metiltransferasas de la invención a partir de diversas especies clostridiales, por ejemplo, *C. aceticum* y *C. ljungdhalii*. En particular, tales sondas pueden usarse para la hibridación con el ADN genómico del género o especie de interés, seguido por procedimientos de transferencia de tipo Southern convencionales, con el fin de identificar y aislar el gen correspondiente en el mismo. Tales sondas puede ser considerablemente más cortas que las secuencias completas, pero deben ser de al menos 14, preferiblemente al menos 25, más preferiblemente al menos 35, y lo más preferiblemente al menos 70 nucleótidos de longitud. Pueden usarse tanto sondas de ADN como de ARN, y las sondas pueden marcarse para detectar el gen correspondiente (por ejemplo, con 32P, 3H, 35S, biotina o avidina). Tales sondas se abarcan en la presente divulgación.

Puede usarse metilación de diversos modos, por ejemplo, metilación *in vitro* o metilación *in vivo*.

#### Metilación *in vitro*

- 5 La elusión de un sistema de restricción-modificación clostridial puede lograrse usando metilación *in vitro* de uno o más polinucleótidos de interés seguido por su introducción en una célula huésped clostridial.

En primer lugar se analiza un polinucleótido de interés para confirmar la presencia de una o más secuencias de reconocimiento de ADN de endonucleasa de restricción *CacI*, CCWGG. En algunas realizaciones, el polinucleótido de interés comprende una o más secuencias de reconocimiento de ADN CCAGG (SEQ ID NO. 9). En algunas realizaciones, el polinucleótido comprende una o más secuencias de reconocimiento de ADN CCTGG (SEQ ID NO: 10). En algunas realizaciones, el polinucleótido de interés comprende una o más secuencias de reconocimiento de ADN CCAGG (SEQ ID NO: 9) y CCTGG (SEQ ID NO: 10).

15 Los ejemplos no limitativos de métodos de análisis de secuencias incluyen secuenciación de Maxam-Gilbert, secuenciación de Sanger, secuenciación de ADN en alineamiento capilar, secuenciación de ciclo térmico (Sears *et al.*, *Biotechniques*, 13:626-633 (1992)), secuenciación en fase sólida (Zimmerman *et al.*, *Methods in Molecular Cell Biology*, 3:39-42 (1992)), secuenciación con espectrometría de masas tal como espectrometría de masas de desorción/ionización por láser asistida por matriz tiempo de vuelo (MALDI-TOF/MS; Fu *et al.*, *Naure Biotechnology*, 16:381-384 (1998)), y secuenciación mediante hibridación. Chee *et al.*, *Science*, 274:610-614 (1996); Drmanac *et al.*, *Science*, 260:1649-1652 (1993); Drmanac *et al.*, *Nature Biotechnology*, 16:54-58 (1998).

Una vez confirmada la presencia de una o más de las secuencias de reconocimiento de ADN de *CacI* en un polinucleótido de interés, se usa una metiltransferasa para metilar la secuencia CCWGG (W = T o A) *in vitro*. Esto puede lograrse, por ejemplo, transformando la secuencia codificante de una metiltransferasa (por ejemplo, una metiltransferasa con una identidad de secuencia de al menos el 90% con SEQ ID NO: 2) que reconoce la secuencia de reconocimiento de ADN CCWGG (W = T o A) en un vector capaz de expresión en una célula huésped recombinante (por ejemplo, un vector pBAD33 inducible por arabinosa capaz de expresión en *E. coli*). Este vector que comprende un polinucleótido que codifica para una metiltransferasa que reconoce específicamente CCWGG (W = T o A) puede transformarse en una célula huésped recombinante (por ejemplo, una célula de *E. coli*) y cultivarse en condiciones adecuadas (por ejemplo, tal como se describe en el ejemplo 4 de la presente solicitud) para la producción de la ADN metiltransferasa codificada. La ADN metiltransferasa producida puede entonces recuperarse y purificarse usando métodos bien conocidos en la técnica tales como cromatografía (por ejemplo, de intercambio iónico, de afinidad, hidrófoba, cromatoenfoco y de exclusión molecular), procedimientos electroforéticos (por ejemplo, isoelectroenfoco preparativo) y solubilidad diferencial (por ejemplo, precipitación con sulfato de amonio). Véanse, por ejemplo, *Protein Purification*, J.C. Janson y Lars Ryden, (eds), VCH Publishers, Nueva York, NY 1989; y Lodish *et al.* (eds.), 2000. *Purifying, Detecting, and Characterizing Proteins*, en *Molecular Biology of the Cell*, 4ª edición, incorporados por el presente documento en su totalidad, particularmente en lo que se refieren a la purificación de proteínas. La metiltransferasa purificada puede usarse entonces para metilar el polinucleótido de interés *in vitro* usando S-adenosil-L-metionina y protocolos de metilación del ADN que se conocen bien en la técnica, dando como resultado por tanto la formación de S-adenosil-L-homocisteína y polinucleótido metilado. La metilación del polinucleótido de interés puede confirmarse usando marcaje radiactivo con [<sup>3</sup>H]S-adenosilmetionina y mapeo y secuenciación de sitios de metilación individuales (por ejemplo, Bitinaite *et al.*, 1992, *Nucleic Acids Research*, vol. 20: 4981-4985), así como ensayos basados en secuenciación de Sanger (por ejemplo, Bart *et al.*, 2005, *Nucleic Acids Research*, vol. 33: e124) o secuenciación de ADN en una sola molécula, en tiempo real (SMRT) (por ejemplo, Clark *et al.*, 2012, *Nucleic Acids Research*, vol. 40, No. 4, e29). Todas las referencias citadas en el presente documento se incorporan por el presente documento en su totalidad, particularmente en lo que se refieren a ensayos de metilación y mapeo de sitios de metilación.

50 En algunas realizaciones de la presente invención, puede usarse un polinucleótido que codifica para una metiltransferasa con una identidad de secuencia de al menos el 90% con SEQ ID NO: 1 que reconoce específicamente sitios de reconocimiento de ADN CCWGG (W = T o A). En otras realizaciones, puede usarse un polinucleótido que codifica para una metiltransferasa con una identidad de secuencia de al menos el 90% con SEQ ID NO: 2 que reconoce específicamente sitios de reconocimiento de ADN CCWGG (W = T o A). En otras realizaciones, puede obtenerse un polinucleótido que codifica para una metiltransferasa con una identidad de secuencia de al menos el 90% con SEQ ID NO: 1 o SEQ ID NO: 2 a través de métodos de síntesis química (por ejemplo, DNA2.0) o crearse usando técnicas de biología molecular convencionales.

60 En algunas realizaciones de la presente divulgación, puede usarse un polinucleótido aislado que tiene una identidad de secuencia de ácido nucleico de al menos aproximadamente cualquiera del 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98%, el 99%, el 99,5%, el 99,9% o el 100% con SEQ ID NO. 1, en el que el polinucleótido codifica para un polipéptido con actividad metiltransferasa que reconoce específicamente CCWGG (W = T o A). En otras realizaciones de la presente divulgación, puede usarse un polinucleótido aislado que tiene una identidad de secuencia de ácido nucleico de al menos aproximadamente cualquiera del 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, el 91%, el 92%, el 93%, el 94%, el

95%, el 96%, el 97%, el 98%, el 99%, el 99,5%, el 99,9%, o el 100% con SEQ ID NO. 2, en el que el polinucleótido codifica para un polipéptido con actividad metiltransferasa que reconoce específicamente CCWGG (W = T o A).

5 En algunas realizaciones, la presente invención se refiere a un polipéptido aislado que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO. 3, en el que dicho polipéptido es capaz de metilar un polinucleótido en SEQ ID NO. 9 y/o SEQ ID NO. 10. En otras realizaciones, la presente divulgación se refiere a un polipéptido aislado que tiene una identidad de secuencia de aminoácidos de al menos aproximadamente cualquiera del 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98%, el 99%, el 99,5%, el 99,9% o el 100% con SEQ ID NO. 3, en el que dicho polipéptido es capaz de metilar un polinucleótido en SEQ ID NO. 9 y/o SEQ ID NO. 10. En todavía otras realizaciones, el polipéptido aislado que tiene actividad metiltransferasa que es capaz de metilar un polinucleótido en una secuencia que comprende CCWGG es SEQ ID NO. 3.

15 Una vez que se han metilado uno o más polinucleótidos de interés, estos polinucleótidos de interés pueden introducirse en células huésped clostridiales usando métodos de transformación tales como electroporación, conjugación, transformación de protoplastos, pistola génica u otro método de transformación conocido en la técnica o comentado en cualquiera de los ejemplos de la presente solicitud. Véanse, por ejemplo, Davis *et al.*, "Gene cloning in Clostridia" (P. Durre, P., ed. (2005) Handbook on Clostridia); Current Protocols in Molecular Biology (F.M. Ausubel *et al.* (eds) capítulo 9, 1987); Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor, 1989; Campbell *et al.*, Current Genetics, vol. 16: 53-56, 1989.

#### Metilación *in vivo* (vectores lanzadera)

25 En algunas realizaciones de la presente invención, la elusión de un sistema de restricción-modificación clostridial puede lograrse usando metilación *in vivo* y vectores lanzadera capaces de propagarse en dos o más especies huésped diferentes. Además de contener cualquier polinucleótido de interés (por ejemplo, polinucleótidos que codifican para la enzima isopreno sintasa y/o cualquier componente de la ruta de DXP), los vectores lanzadera pueden contener un polinucleótido que codifica para una metiltransferasa que reconoce específicamente CCWGG (W = T o A). Alternativamente, la metiltransferasa que reconoce específicamente CCWGG puede proporcionarse en un plásmido diferenciado (por ejemplo, tal como se describe en los ejemplos 7 - 10).

35 Vectores lanzadera a modo de ejemplo son capaces de replicarse en *E. coli* y en un anaerobio obligado, tal como *Clostridium acetivum*. Véase, por ejemplo, Heap *et al.*, 2009, Journal of Microbiological Methods, vol. 78: 79-85, al que se hace referencia por el presente documento, particularmente con respecto a la creación y los componentes de vectores lanzadera para su uso entre *E. coli* y especies bacterianas clostridiales.

40 Se conocen bien en la técnica métodos usados para ligar un constructo (por ejemplo constructo de ADN) que comprende un polinucleótido de interés (por ejemplo, un ácido nucleico de metiltransferasa o endonucleasa), un promotor, un terminador y otras secuencias y para insertarlos en un vector adecuado. Por ejemplo, pueden usarse enzimas de restricción para manipular genéticamente ácido nucleico de metiltransferasa o endonucleasa de manera que pueda ponerse en uno o más vectores. Entonces, los extremos compatibles del ácido nucleico de metiltransferasa o endonucleasa escindido y el vector escindido pueden ligarse. La unión se consigue generalmente mediante ligación en sitios de restricción convenientes. Si tales sitios no existen, se usan ligadores oligonucleotídicos sintéticos según la práctica convencional. Véase Sambrook *et al.*, (1989), Molecular Cloning: A Laboratory Manual (2ª ed., Cold Spring Harbor), al que se hace referencia por el presente documento, particularmente con respecto al aislamiento de ADN, la construcción de vectores y el uso de ligadores oligonucleotídicos. Adicionalmente, pueden construirse vectores usando técnicas de recombinación conocidas (por ejemplo, Invitrogen Life Technologies, tecnología Gateway), o pueden adquirirse de proveedores comerciales de polinucleótidos químicamente sintetizados (por ejemplo, DNA2.0). Los plásmidos lanzadera de la invención reivindicada pueden crearse usando cualquier combinación de métodos bien conocidos en la técnica, incluyendo los descritos en cualquiera de los ejemplos de la presente solicitud.

55 Por ejemplo, para transformar satisfactoriamente *C. acetivum* con ADN heterólogo, pueden construirse vectores lanzadera para la propagación en *E. coli* tal como se describe en el ejemplo 7 de la presente solicitud. En resumen, la construcción de una serie de vectores lanzadera modulares entre *E. coli* y diversas especies bacterianas clostridiales (conocida como "la serie pMTL80000") se describe en Heap *et al.*, 2009 Journal of Microbiological Methods, vol. 78: 79-85. Estos vectores pMTL80000 portan de uno a cuatro replicones Gram-positivos, un origen de replicación p15A o ColE1 en *E. coli*, un sitio de clonación múltiple con terminadores transcripcionales flanqueantes y un marcador resistente a antibióticos seleccionado del grupo de, *catP*, *ermB*, *aad9* o *tetA*. Algunos de los vectores también portan un promotor de ferredoxina de *C. sporogenes* (*Pfdx*) y un sitio de unión al ribosoma (RBS) o un promotor de tiolasa de *C. acetobutylicum* y RBS para la expresión génica.

65 Para crear el vector lanzadera pDW280, se amplificó la estructura principal de plásmido de pMCS203 (también conocido como plásmido pMTL85151) mediante PCR (PfuUltra II, Agilent Technologies) usando los pares de cebadores indicados en la tabla 4 (por ejemplo, GA CA1\_1 203 Dir y GA CA1\_1 203 Inv). El mapa de plásmido y la

5 secuencia de ADN para pMCS203 se proporcionan en la figura 8 y la figura 9A-B, respectivamente. El plásmido pCA1 se amplificó usando los pares de cebadores indicados (por ejemplo, GA CA1\_1 Plásmido Dir y GA CA1\_1 Plásmido Inv, tal como se enumeran en la tabla 4). El mapa de plásmido y la secuencia de ADN para pCA1 se proporcionan en la figura 6 y la figura 7A-B, respectivamente. Se purificaron productos de PCR del peso molecular apropiado mediante electroforesis en gel (Qiagen) y se combinaron usando el kit de clonación sin costuras GeneArt (Life Technologies). Entonces se transformaron estos productos de PCR en células *E. coli* TOP10 químicamente competentes (Life Technologies) según el protocolo recomendado por el fabricante. Se recuperaron células y se sembraron en placa sobre medio selectivo, y se seleccionaron transformantes resistentes a cloranfenicol para su análisis adicional. Se hicieron crecer colonias individuales durante la noche en medio LB selectivo, y el día siguiente se purificaron plásmidos (Qiagen) y se compararon los pesos moleculares con el del plásmido pCA1 original mediante electroforesis en gel. Esto dio como resultado el plásmido pDW264.

15 Tal como se indica en el mapa de plásmido pDW264 mostrado en la figura 20, el vector lanzadera pDW264 contiene el plásmido pCA1 de *Clostridium acetivum* nativo y casetes de ADN que permiten la replicación en *E. coli*, la transferencia conjugativa y la resistencia al antibiótico cloranfenicol. La secuencia de ADN para pDW264 se muestra en la figura 22A-C. A continuación, se cortó pDW264 con las enzimas de restricción *FseI* y *PmeI* (New England Biolabs), siguiendo el protocolo recomendado por el fabricante, para eliminar el casete de resistencia a cloranfenicol. Este vector se ligó entonces (ligasa de T4, NEB) a un casete de resistencia a eritromicina que se había aislado a partir del molde pDW265 mediante digestión por restricción con *FseI*, *PmeI* y *AscI*, y se transformó en células de *E. coli* químicamente competentes Top10 (Life Technologies), usando técnicas de biología molecular convencionales. El plásmido lanzadera conjugativo resultante, pDW280, contenía toda la secuencia nativa de pCA1 de *Clostridium acetivum*, un origen de transferencia, un origen de replicación en *E. coli* y el casete de resistencia a eritromicina. El mapa de plásmido y la secuencia para pDW280 se proporcionan en la figura 27 y la figura 28A-C, respectivamente.

25 El vector lanzadera resultante puede introducirse en una célula huésped que comprende una metiltransferasa que reconoce específicamente la secuencia de reconocimiento de ADN CCWGG (por ejemplo, una célula huésped de *E. coli* S17-1 que expresa metiltransferasa M.Cacl a partir de un plásmido pDW268) para el fin de metilar el vector lanzadera. En algunas realizaciones, el vector lanzadera puede metilarse en una secuencia que comprende CCWGG. En algunas realizaciones, el vector lanzadera puede metilarse en una secuencia que comprende CCAGG (SEQ ID NO: 9). En algunas realizaciones, el vector lanzadera puede metilarse en una secuencia que comprende CCTGG (SEQ ID NO: 10). En algunas realizaciones, el vector lanzadera puede metilarse en CCWGG. En algunas realizaciones, el vector lanzadera puede metilarse en la secuencia de reconocimiento de ADN CCAGG (SEQ ID NO: 9) y/o en la secuencia de reconocimiento de ADN CCTGG (SEQ ID NO: 10).

35 El vector lanzadera metilado puede aislarse entonces e introducirse en una célula huésped bacteriana de *Clostridium* para la expresión del polinucleótido de interés. La introducción del ADN metilado en la célula huésped bacteriana de *Clostridium* puede lograrse mediante los métodos descritos en cualquiera de los ejemplos de la presente solicitud (por ejemplo, conjugación tal como se describe en el ejemplo 10), o mediante el uso o la adaptación de otros métodos de transformación bien conocidos en la técnica. Véanse, por ejemplo, D. Parke, 1990. Construction of mobilizable vectors derived from plasmids RP4, pUC18 y pUC19. *Gene*, vol. 93: 135-137; Simon *et al.*, 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. *Bio-Technology*, vol. 1: 784-791; y McFalane *et al.*, A simplified method for conjugal gene transfer into the filamentous cyanobacterium *Anabaena* sp. ATCC 27893. *Journal of Microbiological Methods*, vol. 6: 301-305, todos los cuales se incorporan en el presente documento en su totalidad, particularmente con respecto a la conjugación, células *E. coli* S17-1 y la creación y el uso de vectores lanzadera bacterianos.

50 Puede usarse cualquier plásmido o vector lanzadera, tal como cualquiera de los plásmidos lanzadera descritos en la presente divulgación (por ejemplo, pDW280, pMCS537, pMCS244, pMCS245, pMCS200 o pMCS201) y/o cualquiera de los plásmidos lanzadera descritos en Heap *et al.*, (2009), *Journal of Microbiological Methods*, vol. 78: 79-85.

55 Puede usarse una diversidad de células huésped para contener, transferir o expresar las metiltransferasas. Las células huésped a modo de ejemplo incluyen, pero no se limitan a, cepas de *Escherichia* tales como células de *Escherichia coli* S17-1. En otras realizaciones, puede usarse cualquier especie bacteriana que pertenezca al género *Clostridium* para contener, transferir o expresar las metiltransferasas. En algunas realizaciones, las metiltransferasas se obtienen a partir de y/o se derivan de especies bacterianas clostridiales, tales como *C. acetivum* y/o *C. ljungdahlii*.

#### Ácidos nucleicos y polipéptidos a modo de ejemplo

60 Pueden usarse diversas metiltransferasas, endonucleasas de restricción y otros polipéptidos y ácidos nucleicos (o bien individualmente o en cualquier combinación) en las composiciones y métodos de la presente divulgación.

65 En algunas realizaciones, un ácido nucleico que codifica para una metiltransferasa o una endonucleasa de restricción está operativamente unido a otro ácido nucleico que codifica para una o más secuencias de control que facilitan la expresión de los polipéptidos codificados. "Operativamente unido" se refiere a uno o más genes que se han colocado bajo el control regulatorio de un promotor, que luego controla la transcripción y opcionalmente la traducción de esos genes. En la construcción de combinaciones de genes promotores/estructurales heterólogos, se

5 prefiere generalmente colocar el promotor o la secuencia genética a una distancia del sitio de inicio de la transcripción génica que es aproximadamente la misma que la distancia entre ese promotor o secuencia genética y el gen que controla en su entorno natural; es decir el gen del que se deriva el promotor o la secuencia genética. Tal como se conoce en la técnica, puede tener cabida algo de variación sin pérdida de función. De manera similar, la colocación preferida de un elemento de secuencia reguladora con respecto a un gen heterólogo que va a ponerse bajo su control se define mediante la colocación del elemento en su entorno natural; es decir, los genes de los que se deriva.

10 En algunas realizaciones, el ácido nucleico tiene una o más mutaciones en comparación con la secuencia de un ácido nucleico de metiltransferasa o endonucleasa de tipo natural (es decir, una secuencia que se produce en la naturaleza). En algunas realizaciones, el ácido nucleico tiene una o más mutaciones (por ejemplo, una mutación silenciosa) que aumentan la transcripción o translación del ácido nucleico. En algunas realizaciones, el ácido nucleico es una variante degenerada de cualquier ácido nucleico que codifica para una metiltransferasa o endonucleasa.

15 Tal como entenderán los expertos en la técnica, las secuencias de polinucleótido de esta invención pueden incluir secuencias genómicas, secuencias extragenómicas y codificadas por plásmidos y segmentos génicos diseñados por ingeniería genética más pequeños que expresan, o pueden adaptarse para expresar, proteínas, polipéptidos, péptidos y similares. Tales segmentos pueden aislarse de manera natural, o modificarse de manera sintética por la mano del hombre.

20 Los polinucleótidos pueden ser monocatenarios (codificantes o antisentido) o bicatenarios, y pueden ser moléculas de ADN (genómico, ADNc o sintético) o de ARN. Pueden estar presentes, pero no necesariamente, secuencias codificantes o no codificantes adicionales dentro de un polinucleótido de la presente invención, y un polinucleótido puede estar unido, pero no necesariamente, a otras moléculas y/o materiales de soporte.

25 Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena) o pueden comprender una variante, o un equivalente funcional biológico de una secuencia de este tipo. Las variantes de polinucleótido pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones. En algunas realizaciones, la actividad enzimática del polipéptido codificado no está sustancialmente disminuida en relación con el polipéptido no modificado. En algunas realizaciones, la actividad enzimática del polipéptido codificado mejora (por ejemplo, se optimiza) en relación con el polipéptido no modificado. En otras realizaciones, la actividad enzimática del polipéptido codificado está sustancialmente disminuida en relación con el polipéptido no modificado. El efecto sobre la actividad enzimática del polipéptido codificado puede evaluarse generalmente tal como se describe en el presente documento.

30 Tal como entenderán los expertos en la técnica, puede ser ventajoso en algunos casos producir secuencias de nucleótidos que codifican para polipéptidos que presentan codones que no se producen de manera natural. Por ejemplo, los codones preferidos por un huésped procarionta o eucariota particular pueden seleccionarse para aumentar la velocidad de expresión de la proteína o para producir un transcrito de ARN recombinante que tiene propiedades deseables, tales como una semivida que es más larga que la de un transcrito generado a partir de la secuencia que se produce de manera natural. Tales nucleótidos se denominan normalmente "de codones optimizados". Cualquiera de las secuencias de nucleótidos descritas en el presente documento puede utilizarse en una forma "de codones optimizados". Además, las secuencias de polinucleótido de la presente invención pueden diseñarse por ingeniería genética usando métodos generalmente conocidos en la técnica con el fin de alterar las secuencias codificantes de polipéptidos por una variedad de motivos, incluyendo pero sin limitarse a, alteraciones que modifican la clonación, el procesamiento, la expresión y/o la actividad del producto genético.

35 Los polinucleótidos pueden comprender un "ácido nucleico heterólogo", cuya secuencia es de otra especie distinta de la célula huésped u otra cepa de la misma especie de célula huésped. En algunas realizaciones, la secuencia no es idéntica a la de otro ácido nucleico que se encuentra de manera natural en la misma célula huésped. En algunas realizaciones, un ácido nucleico heterólogo no es idéntico a un ácido nucleico de tipo natural que se encuentra en la misma célula huésped en la naturaleza.

40 Los polinucleótidos de la presente invención, independientemente de la longitud de la propia secuencia codificante, pueden combinarse con otras secuencias de ADN, tales como promotores, sitios de enzimas de restricción adicionales, sitios de clonación múltiple, otros segmentos codificantes, y similares, de manera que su longitud global puede variar considerablemente. Por tanto, se contempla que pueda emplearse un fragmento de polinucleótido de casi cualquier longitud, estando la longitud total preferiblemente limitada por la facilidad de preparación y uso en el protocolo de ADN recombinante previsto.

45 Los polinucleótidos y las fusiones de los mismos pueden prepararse, manipularse y/o expresarse usando cualquiera de una variedad de técnicas bien establecidas conocidas y disponibles en la técnica. Por ejemplo, pueden usarse secuencias de polinucleótido que codifican para polipéptidos de la invención, o proteínas de fusión o equivalentes funcionales de las mismas en moléculas de ADN recombinante para dirigir la expresión de una enzima seleccionada en células huésped apropiadas. Debido a la degeneración inherente del código genético, pueden producirse otras

secuencias de ADN que codifican sustancialmente para la misma o una secuencia de aminoácidos funcionalmente equivalente y estas secuencias pueden usarse para clonar y expresar un polipéptido dado.

5 En algunas realizaciones, el polipéptido es un polipéptido aislado. Tal como se usa en el presente documento, un "polipéptido aislado" no es parte de una biblioteca de polipéptidos, tal como una biblioteca de 2, 5, 10, 20, 50 o más polipéptidos diferentes y está separado de al menos un componente con el que aparece en la naturaleza. Puede obtenerse un polipéptido aislado, por ejemplo, mediante la expresión de un ácido nucleico recombinante que codifica para el polipéptido.

10 En algunas realizaciones, el polipéptido es un polipéptido heterólogo. Por "polipéptido heterólogo" quiere decirse un polipéptido cuya secuencia de aminoácidos no es idéntica a la de otro polipéptido expresado de manera natural en la misma célula huésped. En particular, un polipéptido heterólogo no es idéntico a un polipéptido de tipo natural que se encuentra en la misma célula huésped en la naturaleza.

15 Con el fin de expresar un polipéptido deseado, puede insertarse una secuencia de nucleótidos que codifica para el polipéptido, o un equivalente funcional, en un vector de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia codificante insertada. Pueden usarse métodos que conocen bien los expertos en la técnica para construir vectores de expresión que contienen secuencias que codifican para un polipéptido de interés y elementos de control de la transcripción y traducción apropiados. Estos métodos incluyen técnicas de ADN recombinante *in vitro*, técnicas de síntesis y recombinación genética *in vivo*. Tales técnicas se describen en Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (1989), y Ausubel *et al.*, *Current Protocols in Molecular Biology* (1989).

25 "Polipéptido", "fragmento de polipéptido", "péptido" y "proteína" se usan indistintamente en el presente documento para referirse a un polímero de residuos de aminoácido y a variantes y análogos sintéticos del mismo. Por tanto, estos términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácido son aminoácidos sintéticos que no se producen de manera natural, tales como un análogo químico de un aminoácido que se produce de manera natural correspondiente, así como a polímeros de aminoácidos que se producen de manera natural. Los polipéptidos incluyen polipéptidos enzimáticos, o "enzimas", que normalmente catalizan (es decir, aumentan la velocidad de) diversas reacciones químicas, (por ejemplo, ADN metiltransferasas o endonucleasas de restricción).

La "identidad de secuencia", tal como se usa en este documento, se refiere al grado en que las secuencias son idénticas en una base de nucleótido por nucleótido o una base de aminoácido por aminoácido a lo largo de una ventana de comparación. Por lo tanto, un "porcentaje de identidad de secuencia" puede calcularse comparando dos secuencias alineadas de manera óptima a lo largo de la ventana de comparación, determinando el número de posiciones en las que aparece una base de ácido nucleico (por ejemplo, A, T, C, G, I) o un residuo de aminoácido (por ejemplo, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys y Met) idéntico en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la ventana de comparación (es decir, el tamaño de la ventana), y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia.

Los términos usados para describir relaciones de secuencia entre dos o más polinucleótidos o polipéptidos incluyen "secuencia de referencia", "ventana de comparación", "identidad de secuencia", "porcentaje de identidad de secuencia" e "identidad sustancial". Una "secuencia de referencia" tiene al menos 12 pero con frecuencia de 15 a 18 y a menudo al menos 25 unidades de monómero, incluidos nucleótidos y residuos de aminoácidos, de longitud. Debido a que dos polinucleótidos pueden comprender cada uno (1) una secuencia (es decir, sólo una porción de la secuencia de polinucleótido completa) que es similar entre los dos polinucleótidos, y (2) una secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencias entre dos (o más) polinucleótidos se realizan normalmente comparando secuencias de los dos polinucleótidos a lo largo de una "ventana de comparación" para identificar y comparar regiones locales de similitud de secuencia. Una "ventana de comparación" se refiere a un segmento conceptual de al menos 6 posiciones contiguas, habitualmente de aproximadamente 50 a aproximadamente 100, más habitualmente de aproximadamente 100 a aproximadamente 150 en el que se compara una secuencia con una secuencia de referencia del mismo número de posiciones contiguas después de que la dos secuencias se alineen de manera óptima. La ventana de comparación puede incluir adiciones o eliminaciones (es decir, huecos) de aproximadamente el 20% o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para una alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación puede realizarse mediante implementaciones computarizadas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en la versión 7.0 del paquete de software Wisconsin Genetics, Genetics Computer Group, 575 Science Drive Madison, WI, EE. UU.) o mediante inspección y la mejor alineación (es decir, que da como resultado el porcentaje de homología más alto a lo largo de la ventana de comparación) generada por cualquiera de los diversos métodos seleccionados.

También puede hacerse referencia a la familia de programas BLAST tal como divulgan por ejemplo Altschul *et al.*, 1997, *Nucl. Acids Res.* 25:3389. Una discusión detallada del análisis de secuencias puede encontrarse en la unidad 19.3 de Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc., 1994-1998, capítulo 15.

Además, pueden usarse programas convencionales de alineación de secuencias y/o predicción de estructuras para identificar polipéptidos y ácidos nucleicos de metiltransferasa o endonucleasa adicionales basándose en la similitud de su estructura secundaria polipeptídica predicha y/o primaria con la de polipéptidos y ácidos nucleicos de metiltransferasa o endonucleasa conocidos. Pueden usarse también bases de datos convencionales tales como la base de datos SwissProt-Treml (página web en “expasy.org”, Swiss Institute of Bioinformatics Swiss-Prot group CMU-1 rue Michel Servet CH-1211 Ginebra 4, Suiza) para identificar polipéptidos y ácidos nucleicos de metiltransferasa o endonucleasa. La estructura secundaria y/o terciaria de un polipéptido de metiltransferasa o endonucleasa puede predecirse usando los parámetros por defecto de programas de predicción de estructuras convencionales, tales como PredictProtein. Alternativamente, la estructura secundaria y/o terciaria real de un polipéptido de metiltransferasa o endonucleasa puede determinarse usando métodos convencionales.

#### *Métodos a modo de ejemplo para aislar ácidos nucleicos*

Pueden aislarse ácidos nucleicos que codifican para metiltransferasas o endonucleasas de restricción usando métodos convencionales. Métodos de obtención de ácidos nucleicos deseados a partir de un organismo fuente de interés (tal como un genoma bacteriano) son comunes y se conocen bien en la técnica de biología molecular (véase, por ejemplo, el documento WO 2004/033646 y referencias citadas en el mismo, al que se hace referencia por el presente documento, particularmente con respecto al aislamiento de ácidos nucleicos de interés). Por ejemplo, si la secuencia del ácido nucleico se conoce (tal como cualquiera de los ácidos nucleicos conocidos descritos en el presente documento), pueden crearse bibliotecas genómicas adecuadas mediante digestión por endonucleasas de restricción y pueden examinarse con sondas complementarias a la secuencia de ácido nucleico deseada. Una vez que la secuencia se aísla, el ADN puede amplificarse usando métodos de amplificación dirigida de cebadores convencionales tales como reacción en cadena de la polimerasa (PCR) (patente estadounidense n.º 4.683.202, a la que se hace referencia, particularmente con respecto a métodos de PCR) para obtener cantidades de ADN adecuadas para la transformación usando vectores apropiados.

Alternativamente, pueden sintetizarse químicamente polinucleótidos que codifican para metiltransferasas o endonucleasas que reconocen específicamente CCWGG (W = T o A) usando métodos convencionales (por ejemplo, DNA2.0).

#### *Vectores, promotores y otros elementos a modo de ejemplo*

##### *Vectores*

Cualquiera de los ácidos nucleicos de metiltransferasa o endonucleasa descritos en el presente documento (solo o en cualquier combinación) puede incluirse en uno o más vectores. Por consiguiente, la presente divulgación también presenta vectores con uno o más ácidos nucleicos que codifican para cualquiera de los polipéptidos de metiltransferasa o endonucleasa que se describen en el presente documento. Tal como se usa en el presente documento, un “vector” significa un constructo que es capaz de administrar, y expresar de manera deseable, uno o más ácidos nucleicos de interés en una célula huésped. Los ejemplos de vectores incluyen, pero no se limitan a, plásmidos, vectores virales, vectores de expresión de ADN o ARN, cósmidos y vectores de fagos. En algunas realizaciones, el vector contiene un ácido nucleico bajo el control de una secuencia de control de la expresión.

Tal como se usa en este documento, una “secuencia de control de la expresión” significa una secuencia de ácido nucleico que dirige la transcripción de un ácido nucleico de interés. Una secuencia de control de expresión puede ser un promotor, tal como un promotor constitutivo o inducible, o un potenciador. Un “promotor inducible” es un promotor que es activo bajo una regulación ambiental o de desarrollo, tal como un promotor inducible por arabinosa. La secuencia de control de la expresión está operativamente unida al segmento de ácido nucleico que va a transcribirse.

En algunas realizaciones, el vector contiene un marcador selectivo. El término “marcador selectivo” se refiere a un ácido nucleico capaz de expresarse en una célula huésped que permite una fácil selección de aquellas células huésped que contienen un ácido nucleico o vector introducido. Los ejemplos de marcadores seleccionables incluyen, pero no se limitan a, ácidos nucleicos de resistencia a antibióticos (por ejemplo, eritromicina, cloranfenicol, tiamfenicol, kanamicina, ampicilina, carbenicilina, gentamicina, higromicina, estreptomycin, fleomicina, bleomicina o neomicina), y/o ácidos nucleicos que confieren una ventaja metabólica, tal como una ventaja nutricional en la célula huésped. Vectores adecuados son aquellos que son compatibles con la célula huésped empleada. Los vectores adecuados pueden derivarse, por ejemplo, de una bacteria, un virus (como el bacteriófago T7 o un fago derivado de M-13), un cósmido, una levadura o una planta. Los expertos en la técnica conocen protocolos para obtener y usar tales vectores (véase, por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor, 1989, al que se hace referencia por el presente documento, particularmente con respecto al uso de vectores).

##### *Promotores*

Se usan promotores adecuados para expresar cualquiera de los ácidos nucleicos heterólogos descritos en el presente documento. Pueden usarse promotores adecuados para dirigir la producción de polipéptidos de metiltransferasa o endonucleasa, o para reducir la degradación de polipéptidos de metiltransferasa o endonucleasa en células huésped.

Pueden usarse promotores adecuados para optimizar la expresión de polipéptidos de metiltransferasa o endonucleasa en una célula huésped. Cualquiera de los ácidos nucleicos descritos en el presente documento (por ejemplo, un ácido nucleico que codifica para polipéptidos de metiltransferasa o endonucleasa) puede unirse operativamente a un promotor. Puede usarse cualquiera de los promotores descritos en el presente documento, tal como el promotor de *Clostridium acetivum* nativo contenido en el plásmido pCA1 (SEQ ID NO. 6).

Altos niveles de expresión en determinadas células clostridiales pueden provocar la degradación del/de los polipéptido(s) diseñados por ingeniería genética incluyendo metiltransferasas o las endonucleasas. Para mejorar la producción de metiltransferasa o endonucleasa, puede usarse un sistema de expresión inducible que permite tanto controlar el momento como la magnitud de la expresión del/de los polipéptido(s) diseñado(s) por ingeniería genética. El control más estricto puede facilitar la expresión del/de los polipéptido(s) diseñado(s) por ingeniería genética a una concentración y un período durante el crecimiento de las células que es tóxico para las células, y da como resultado la producción de mayores cantidades del polipéptido deseado.

Un promotor usado en cualquiera de las células descritas en el presente documento puede ser un promotor inducible. Puede usarse un sistema de expresión inducible por arabinosa; por ejemplo, el sistema inducible por arabinosa PBAD tal como se describe en Guzman *et al.*, "Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter". Journal of Bacteriology, vol. 177, n.º 14: 4121-4130 (julio de 1995), al que se hace referencia por el presente documento, particularmente con respecto a su divulgación de vectores pBAD que usan el promotor P<sub>BAD</sub> inducible por arabinosa. Alternativamente, puede usarse un sistema de expresión inducible por gluconato, por ejemplo, un sistema de expresión inducible por gluconato endógeno para *C. ljungdahlii*. Se predice que los ORF *clju19880* y *clju30510* codifican para factores de transcripción que reprimen la expresión de genes implicados en la importación y el metabolismo de gluconato. En presencia de gluconato, el gluconato se une a y reprime estos factores de transcripción, permitiendo por tanto la expresión de genes implicados en la importación y el metabolismo de gluconato. Se ha indicado el ORF *clju11610* como "gluconocinasa" en el genoma de *C. ljungdahlii*. En *Corynebacterium glutamicum*, el promotor de gluconato cinasa (nombre alternativo para gluconocinasa) presenta el aumento más fuerte en la expresión en respuesta a la inducción con gluconato (Frunzke *et al.* 2008, Mol Microbiol., 67(2):305-22). Por tanto, en algunos aspectos, el promotor puede ser un promotor inducible por gluconato. En algunos aspectos, el promotor puede ser de *C. acetobutylicum*, *C. ljungdahlii*, *C. autoethanogenum* o *C. acetivum*. En algunos aspectos, el promotor puede ser el promotor presente en el ORF *clju19880*, ORF *clju 11610* u ORF *clju30510* en una célula anaerobia (por ejemplo, *C. ljungdahlii*). En algunos aspectos, el promotor puede ser un promotor de *C. acetivum* nativo, tal como se encuentra en el plásmido pCA1 (SEQ ID NO. 6). En algunos aspectos, el promotor es un promotor presente en pCA1. En algunos aspectos, el promotor es un promotor inducible por arabinosa. En algunos aspectos, el promotor es un promotor inducible por gluconato tal como el promotor de gluconato cinasa. El promotor puede ser también un promotor que se induce cuando las células se cultivan en presencia de gas de síntesis, hidratos de carbono (por ejemplo, fructosa o glucosa), o cualquier combinación de los mismos.

Un promotor usado en cualquiera de las células descritas en el presente documento puede ser un promotor constitutivo. Los promotores constitutivos no requieren inducción por medios artificiales (tal como IPTG para la inducción del operón *lac*) y por tanto pueden dar como resultado una reducción de costes considerable para fermentaciones a gran escala. Pueden usarse promotores constitutivos que funcionan en anaerobios (por ejemplo, *C. acetobutylicum*, *C. acetivum* y *C. ljungdahlii*). Pueden ser deseables promotores que tienen una baja expresión en determinadas realizaciones. El promotor de *ptb* (fosfotransbutirilasa) de *C. acetobutylicum* es fuertemente activo durante la fase de crecimiento exponencial de cultivos de *C. acetobutylicum*. Los promotores que pueden usarse en la presente invención pueden tener menos actividad que el promotor de *ptb* (fosfotransbutirilasa). Se ha mostrado que el promotor de *spolIE* (proteína de esporulación de fase II E), también de *C. acetobutylicum*, es activo de manera transitoria en fase estacionaria media. El promotor de *spolIE* (proteína de esporulación de fase II E) puede usarse en la presente invención. Por tanto, en algunos aspectos, el promotor es el promotor de *spolIE* (por ejemplo, promotor de *spolIE* de *Clostridium acetobutylicum*). En algunos aspectos, el promotor tiene una fuerza que está a un nivel inferior a *ptb* (por ejemplo, el promotor tiene una capacidad reducida para dirigir la expresión en comparación con *ptb* tal como *ptb* de *Clostridium acetobutylicum*). En algunos aspectos, el promotor tiene una fuerza que está a un nivel similar a *spolIE* (por ejemplo, el promotor tiene una capacidad similar de dirigir la expresión en comparación con *spolIE*). En algunos aspectos, el promotor es activo tras la fase de crecimiento exponencial. En algunos aspectos, el promotor es activo durante la fase de crecimiento lineal. En algunos aspectos, el promotor es activo en fase estacionaria. En algunos aspectos, el promotor usado en cualquiera de las células descritas en el presente documento es sólo activo en presencia de sintegas. En algunos aspectos, el promotor expresa la metiltransferasa o endonucleasa a bajo nivel. En algunos aspectos, el promotor expresa la metiltransferasa o endonucleasa a un nivel tal que la metiltransferasa o endonucleasa no se escinde por una proteasa o un menor porcentaje de la metiltransferasa o endonucleasa se escinde por una proteasa. En algunos aspectos, el promotor produce una expresión de bajo nivel.

Puede usarse uno cualquiera de los promotores caracterizados o usados en los ejemplos de la presente divulgación.

5 Se conocen bien en la técnica promotores, y cualquier promotor que funcione en la célula huésped puede usarse para la expresión de un ácido nucleico de metiltransferasa o endonucleasa en la célula huésped. Las regiones o promotores de control de la iniciación, que son útiles para dirigir la expresión de polipéptidos en diversas células huésped, son numerosos y familiares para los expertos en la técnica (véase, por ejemplo, el documento WO 2004/033646 y referencias citadas en el mismo, a cada una de las que se hace referencia por el presente documento, particularmente con respecto a vectores para la expresión de ácidos nucleicos de interés).  
10 Prácticamente cualquier promotor capaz de dirigir estos ácidos nucleicos es adecuado para la presente invención incluyendo, pero sin limitarse a, lac, trp, T7, tac y trc (útiles para la expresión en *E. coli*).

#### Plásmidos

15 En diversas realizaciones, un ácido nucleico de metiltransferasa o endonucleasa está contenido en un plásmido de bajo número de copias (por ejemplo, un plásmido que se mantiene a de aproximadamente 1 a aproximadamente 4 copias por célula), plásmido de número de copias medio (por ejemplo, un plásmido que se mantiene a de aproximadamente 10 a aproximadamente 15 copias por célula) o plásmido de alto número de copia (por ejemplo, un plásmido que se mantiene a de aproximadamente 50 o más copias por célula). En algunas realizaciones, el ácido nucleico de metiltransferasa o endonucleasa está operativamente unido a un promotor P<sub>BAD</sub>. En algunas realizaciones, el ácido nucleico de metiltransferasa o endonucleasa operativamente unido a un promotor P<sub>BAD</sub> está contenido en un plásmido de número de copias medio o alto. En algunas realizaciones, el ácido nucleico de metiltransferasa o endonucleasa está operativamente unido a un promotor *de Clostridium acetikum* nativo, tal como está contenido en el plásmido pCA1. En algunas realizaciones, el ácido nucleico de metiltransferasa o endonucleasa operativamente unido a un promotor está contenido en un plásmido de número de copias medio o alto.  
20  
25

En algunas realizaciones, el vector es un plásmido de replicación que no se integra en un cromosoma en las células. En algunas realizaciones, parte o la totalidad del vector se integra en un cromosoma en las células. Se proporcionan ejemplos adicionales de vectores de expresión y/o integración adecuados en Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>a</sup> ed., Cold Spring Harbor, 1989 y Current Protocols in Molecular Biology (F. M. Ausubel *et al.* (Eds) 1987, suplemento 30, sección 7.7.18) a los que se hace referencia por el presente documento, particularmente con respecto a los vectores. Los vectores particularmente útiles incluyen pFB6, pBR322, PUC18, pUC100 y pENTR/D.  
30

#### 35 Otros elementos

También pueden usarse otros elementos de biología molecular, tales como secuencia de terminación, orígenes de replicación y similares.

40 En algunas realizaciones, el vector de expresión también incluye una secuencia de terminación. Las regiones de control de la terminación también pueden derivarse de diversos genes nativos para la célula huésped. En algunas realizaciones, la secuencia de terminación y la secuencia del promotor se derivan de la misma fuente. En otra realización, la secuencia de terminación es endógena para la célula huésped. Opcionalmente, puede incluirse un sitio de terminación. Para la expresión eficaz de los polipéptidos, el ADN que codifica para el polipéptido se une operativamente a través de codones de iniciación a regiones de control de la expresión seleccionadas de manera que la expresión da como resultado la formación del ARN mensajero apropiado.  
45

Puede incorporarse un ácido nucleico de metiltransferasa o endonucleasa en un vector, tal como un vector de expresión, usando técnicas convencionales (Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982, al que se hace referencia por el presente documento, particularmente con respecto a la selección de secuencias de ADN apropiadas y la construcción de vectores). Los métodos usados para ligar el constructo de ADN que comprende un ácido nucleico de interés (tal como un ácido nucleico de metiltransferasa o endonucleasa), un promotor, un terminador y otras secuencias, y para insertarlos en un vector adecuado, se conocen bien en la técnica. Por ejemplo, pueden usarse enzimas de restricción para escindir el ácido nucleico de metiltransferasa o endonucleasa y el vector. Luego, los extremos compatibles del ácido nucleico de metiltransferasa o endonucleasa escindido y el vector escindido pueden ligarse. La unión se realiza generalmente mediante ligación en sitios de restricción convenientes. Si tales sitios no existen, se usan ligadores oligonucleotídicos sintéticos según la práctica convencional (véanse Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2<sup>a</sup> ed., Cold Spring Harbor, 1989 y Bennett y Lasure, More Gene Manipulations in Fungi, Academic Press, San Diego, págs. 70-76, 1991, a los que se hace referencia por el presente documento, particularmente con respecto a los ligadores oligonucleotídicos). Además, pueden construirse vectores usando técnicas de recombinación conocidas (por ejemplo, Invitrogen Life Technologies, tecnología Gateway).  
50  
55  
60

Pueden usarse diferentes tipos de orígenes de replicación. Pueden usarse uno, dos o más orígenes de replicación. Los orígenes de replicación pueden ser de diferentes organismos y/o organismos Gram-positivos o Gram-negativos. En los ejemplos se describen adicionalmente usos a modo de ejemplo de orígenes de replicación para poner en  
65

práctica la invención.

*Métodos de transformación clostridial*

5 Actualmente, los métodos de transformación clostridial incluyen pero no se limitan a: (i) electroporación, mediante la cual se exponen células a campos eléctricos de alta intensidad que provocan que la membrana celular se vuelva transitoriamente porosa, permitiendo por tanto la entrada de ADN dentro de la célula; (ii) transferencia conjugativa (o conjugación) de ADN de plásmido de un organismo donador tal como *E. coli*, mediante la cual se transfiere ADN de la célula donadora a una célula receptora a través de contacto de célula a célula; (iii) transformación de protoplastos, mediante lo cual la pared de células clostridiales se elimina enzimática o químicamente para formar protoplastos que incorporan plásmidos en su citoplasma cuando se incuban con ADN; y/o (iv) pistola génica (sistema de administración de partículas biolísticas), mediante la cual una pequeña partícula de metal pesado se recubre con ADN de plásmido y posteriormente se impulsa a alta velocidad hacia la célula bacteriana. Estas y otras técnicas de transformación se describen en la técnica, véanse por ejemplo, Davis *et al.*, "Gene cloning in Clostridia" (P. Durre, P., ed. 2005) Handbook on Clostridia); Current Protocols in Molecular Biology (F.M. Ausubel *et al.* (eds) capítulo 9, 1987); Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor, 1989; Campbell *et al.*, Current Genetics, vol. 16: 53-56, 1989.

15 Pueden generarse cebadores, oligonucleótidos y polinucleótidos empleados en la presente invención usando técnicas convencionales conocidas en la técnica.

*Sistemas de expresión clostridial*

25 La presente divulgación proporciona sistemas de expresión de *Clostridium* para la producción de uno o más bioproductos industriales (por ejemplo, isopreno, butadieno o etanol). El sistema de expresión puede incluir cualquier combinación de elementos requeridos para la producción de uno o más bioproductos industriales. En algunas realizaciones de la divulgación, el sistema puede incluir uno o más de: (a) una metiltransferasa (por ejemplo, un plásmido que comprende pDW268 o pMCS466), (b) un plásmido lanzadera (por ejemplo, pDW280, pMCS537, pMCS200, pMCS201, pMCS444 o PMCS445), (c) una célula bacteriana de *E. coli* capaz de interactuar con una célula bacteriana de *Clostridium* para permitir la transferencia de (a) y (b); y (d) una célula bacteriana de *Clostridium* capaz de interactuar con una célula bacteriana de *Escherichia* de manera que el uno o más ácidos nucleicos se expresan en la célula bacteriana de *Clostridium*. En algunas realizaciones de la divulgación, la célula de bacterias de *E. coli* capaz de interactuar con una célula bacteriana de *Clostridium* es una célula de *E. coli* S17-1. En algunas realizaciones de la divulgación la célula bacteriana de *Clostridium* capaz de interactuar con una célula bacteriana de *Escherichia* se selecciona del grupo de *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum* o *Clostridium acetobutylicum*. En algunas realizaciones de la divulgación el sistema proporciona la expresión de uno o más ácidos nucleicos de interés (por ejemplo, ácidos nucleicos que codifican para isopreno sintasa o enzimas implicadas en la producción de etanol a partir de acetyl-CoA).

40 *Células huésped para la producción de bioproductos industriales*

Diversos tipos de células bacterianas clostridiales pueden usarse como células huésped para producir bioproductos industriales. Las células huésped a modo de ejemplo incluyen, pero no se limitan a, especies del género *Clostridium* tales como *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum*, *Clostridium autoethanogenum*. Las células huésped a modo de ejemplo también incluyen, pero no se limitan a especies del género *Clostridium* tales como *Clostridium carboxydvorans*, *Clostridium difficile*, *Clostridium botulinum*, *Clostridium tetani*, *Clostridium perfringens*, *Clostridium thermoaceticum* (también conocida como *Moorella thermoacetica*), *Clostridium aminobutyricum*, *Clostridium beijerinckii*, *Clostridium beijerinckii* NCIMB 8052, *Clostridium beijerinckii* NRRL B593, *Clostridium kluyveri*, *Clostridium kluyveri* DSM 555, *Clostridium novyi* NT, *Clostridium propionicum* y *Clostridium saccharoperbutylacetonicum*.

*Parámetros de crecimiento y/o producción*

55 Las células clostridiales y composiciones de las mismas pueden modificarse por ingeniería genética para producir un bioproducto industrial en un sistema de fermentación. En una realización el sistema está sustancialmente libre de oxígeno. En algunas realizaciones, el sistema de fermentación contiene un hidrato de carbono como fuente de energía y/o carbono. En algunas realizaciones, el sistema de fermentación contiene hidrato de carbono e hidrógeno como fuente de energía y/o carbono. Las composiciones y los métodos de la invención utilizan condiciones sustancialmente libres de oxígeno. En un aspecto, condiciones sustancialmente libres de oxígeno son condiciones en las que organismos anaerobios pueden crecer y/o producir los productos deseados. Las condiciones pueden referirse al sistema de fermentación (por ejemplo, biorreactor) además de al medio de cultivo. En otros aspectos, condiciones sustancialmente libres de oxígeno se refieren a un sistema de fermentación en el que hay menos de aproximadamente cualquiera del 5, el 4, el 3, el 2, el 1, el 0,5, el 0,2 ó el 0,1% en peso de oxígeno. En algunos aspectos, el sistema de fermentación comprende menos de aproximadamente el 0,01% en peso de oxígeno. En algunos aspectos, el sistema de fermentación comprende menos de aproximadamente el 0,001% en peso de oxígeno.

En algunos aspectos, el sistema de fermentación comprende menos de aproximadamente 100 ppm de oxígeno. En algunos aspectos, el sistema de fermentación comprende menos de aproximadamente 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, 2 ó 1 ppm de oxígeno. En algunos aspectos, la cantidad de oxígeno en el sistema de fermentación está a un nivel suficientemente bajo de manera que un anaerobio obligado es capaz de reproducirse y/o producir isopreno. En algunos aspectos, la cantidad de oxígeno en el sistema de fermentación está a un nivel suficientemente bajo de manera que un anaerobio facultativo favorece la fermentación anaerobia con respecto a la respiración aerobia.

En algunos aspectos, se adoptan etapas para eliminar el oxígeno del medio de cultivo. El oxígeno puede eliminarse añadiendo un catalizador y opcionalmente añadiendo hidrógeno al medio de cultivo. En algunos aspectos, el catalizador es cobre.

#### *Materia prima*

Pueden usarse diversos tipos de materia prima para las células clostridiales recombinantes descritas en el presente documento. La materia prima puede ser una fuente de carbono o sintegas. A continuación se proporciona información sobre materias primas a modo de ejemplo.

#### *Fuente de carbono*

Puede usarse cualquier fuente de carbono para cultivar las células huésped. El término "fuente de carbono" se refiere a uno o más compuestos que contienen carbono capaces de metabolizarse por células clostridiales recombinantes descritas en el presente documento. Por ejemplo, el medio celular usado para cultivar las células clostridiales recombinantes descritas en el presente documento puede incluir cualquier fuente de carbono adecuada para mantener la viabilidad o el crecimiento de las células.

En algunas realizaciones, la fuente de carbono es un carbohidrato (tal como monosacárido, disacárido, oligosacárido o polisacáridos), azúcar invertido (por ejemplo, jarabe de sacarosa tratado enzimáticamente), glicerol, glicerina (por ejemplo, un subproducto de glicerina de un proceso de fabricación de biodiesel o jabón), dihidroxiacetona, fuente de un carbono, aceite (por ejemplo, un aceite de planta o vegetal tal como maíz, palma o aceite de soja), grasa animal, aceite animal, ácido graso (por ejemplo, un ácido graso saturado, ácido graso insaturado o ácido graso poliinsaturado), lípido, fosfolípido, glicerolípido, monoglicérido, diglicérido, triglicérido, polipéptido (por ejemplo, una proteína o péptido microbiano o vegetal), fuente de carbono renovable (por ejemplo una fuente de carbono de biomasa tal como una fuente de carbono de biomasa hidrolizada), extracto de levadura, componente de un extracto de levadura, polímero, ácido, alcohol, aldehído, cetona, aminoácido, succinato, lactato, acetato, etanol, o cualquier combinación de dos o más de los anteriores. En algunas realizaciones, la fuente de carbono es un producto de la fotosíntesis, que incluye, pero no se limita a, glucosa.

Los monosacáridos a modo de ejemplo incluyen glucosa y fructosa; los oligosacáridos a modo de ejemplo incluyen lactosa y sacarosa, y los polisacáridos a modo de ejemplo incluyen almidón y celulosa. Los hidratos de carbono a modo de ejemplo incluyen azúcares C6 (por ejemplo, fructosa, manosa, galactosa o glucosa) y azúcares C5 (por ejemplo, xilosa o arabinosa). En algunas realizaciones, el medio celular incluye un hidrato de carbono así como una fuente de carbono distinta de un hidrato de carbono (por ejemplo, glicerol, glicerina, dihidroxiacetona, fuente de un carbono, aceite, grasa animal, aceite animal, ácido graso, lípido, fosfolípido, glicerolípido, monoglicérido, diglicérido, triglicérido, fuente de carbono renovable o un componente de un extracto de levadura). En algunas realizaciones, el medio celular incluye un hidrato de carbono así como un polipéptido (por ejemplo, una proteína o péptido microbiano o vegetal). En algunas realizaciones, el polipéptido microbiano es un polipéptido de levadura o bacteria. En algunas realizaciones, el polipéptido vegetal es un polipéptido de soja, maíz, canola, jatrofa, palma, cacahuete, girasol, coco, mostaza, colza, semilla de algodón, palmiste, oliva, cártamo, sésamo o linaza.

En algunas realizaciones, las células se cultivan en condiciones limitadas de glucosa. Por "condiciones limitadas de glucosa" quiere decirse que la cantidad de glucosa que se añade es menor de o aproximadamente el 105% (tal como aproximadamente el 100%) de la cantidad de glucosa que consumen las células. En realizaciones particulares, la cantidad de glucosa que se añade al medio de cultivo es aproximadamente la misma que la cantidad de glucosa que consumen las células durante un período específico de tiempo. En algunas realizaciones, la velocidad de crecimiento celular se controla limitando la cantidad de glucosa añadida de modo que las células crezcan a la velocidad que puede soportar la cantidad de glucosa en el medio celular. En algunas realizaciones, la glucosa no se acumula durante el tiempo en que se cultivan las células. En diversas realizaciones, las células se cultivan en condiciones de glucosa limitadas durante más de o aproximadamente 1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60 ó 70 horas. En diversas realizaciones, las células se cultivan en condiciones de glucosa limitadas durante más de o aproximadamente el 5, el 10, el 15, el 20, el 25, el 30, el 35, el 40, el 50, el 60, el 70, el 80, el 90, el 95 o el 100% del tiempo total de cultivo de las células. Aunque no se pretende restringirse a ninguna teoría en particular, se cree que las condiciones limitadas de glucosa pueden permitir una regulación más favorable de las células.

En algunas realizaciones, las células se cultivan en presencia de un exceso de glucosa. En realizaciones particulares, la cantidad de glucosa que se añade es mayor de aproximadamente el 105% (tal como

aproximadamente o mayor del 110, el 120, el 150, el 175, el 200, el 250, el 300, el 400 o el 500%) o más de la cantidad de glucosa que consumen las células durante un período específico de tiempo. En algunas realizaciones, la glucosa se acumula durante el tiempo en que las células se cultivan. Lípidos a modo de ejemplo son cualquier sustancia que contenga uno o más ácidos grasos que sean ácidos grasos C4 y superiores que estén saturados, insaturados o ramificados.

Aceites a modo de ejemplo son lípidos que son líquidos a temperatura ambiente. En algunas realizaciones, el lípido contiene uno o más ácidos grasos C4 o superiores (por ejemplo, contiene uno o más ácidos grasos saturados, insaturados o ramificados con cuatro o más carbonos). En algunas realizaciones, el aceite se obtiene a partir de soja, maíz, canola, jatrofa, palma, cacahuete, girasol, coco, mostaza, colza, semilla de algodón, palmiste, oliva, cártamo, sésamo, linaza, células microbianas oleaginosas, sebo chino o cualquier combinación de dos o más de los anteriores.

Ácidos grasos a modo de ejemplo incluyen compuestos de fórmula RCOOH, donde "R" es un hidrocarburo. Los ácidos grasos insaturados a modo de ejemplo incluyen compuestos en los que "R" incluye al menos un doble enlace carbono-carbono. Los ácidos grasos insaturados a modo de ejemplo incluyen, pero no se limitan a, ácido oleico, ácido vaccénico, ácido linoleico, ácido palmitoelaídico y ácido araquidónico. Los ácidos grasos poliinsaturados a modo de ejemplo incluyen compuestos en los que "R" incluye una pluralidad de dobles enlaces carbono-carbono. Los ácidos grasos saturados a modo de ejemplo incluyen compuestos en los que "R" es un grupo alifático saturado. En algunas realizaciones, la fuente de carbono incluye uno o más ácidos grasos C<sub>12</sub>-C<sub>22</sub>, tales como un ácido graso saturado C<sub>12</sub>, un ácido graso saturado C<sub>14</sub>, un ácido graso saturado C<sub>16</sub>, un ácido graso saturado C<sub>18</sub>, un ácido graso saturado C<sub>20</sub> o un ácido graso saturado C<sub>22</sub>. En una realización a modo de ejemplo, el ácido graso es ácido palmítico. En algunas realizaciones, la fuente de carbono es una sal de un ácido graso (por ejemplo, un ácido graso insaturado), un derivado de un ácido graso (por ejemplo, un ácido graso insaturado) o una sal de un derivado de ácido graso (por ejemplo, un ácido graso insaturado). Las sales adecuadas incluyen, pero no se limitan a, sales de litio, sales de potasio, sales de sodio y similares. Di y triglicérols son ésteres de ácidos grasos de glicerol.

En algunas realizaciones, la concentración de lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido es de al menos o aproximadamente 1 gramo por litro de caldo (g/l, en donde el volumen de caldo incluye tanto el volumen de medio celular como el volumen de las células), tal como al menos o aproximadamente 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400 o más g/l. En algunas realizaciones, la concentración del lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido está entre aproximadamente 10 y aproximadamente 400 g/l, tal como entre aproximadamente 25 y aproximadamente 300 g/l, entre aproximadamente 60 y aproximadamente 180 g/l o entre aproximadamente 75 y aproximadamente 150 g/l. En algunas realizaciones, la concentración incluye la cantidad total de lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido que se añade antes y/o durante el cultivo de las células huésped. En algunas realizaciones, la fuente de carbono incluye tanto (i) un lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido como (ii) un hidrato de carbono, tal como glucosa. En algunas realizaciones, la razón de lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido con respecto al hidrato de carbono es de aproximadamente 1:1 en una base de carbono (es decir, un carbono en el lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido por carbono de hidratos de carbono). En realizaciones particulares, la cantidad del lípido, aceite, grasa, ácido graso, monoglicérido, diglicérido o triglicérido está entre aproximadamente 60 y 180 g/l, y la cantidad de hidrato de carbono está entre aproximadamente 120 y 360 g/l.

Los ejemplos de fuentes de carbono de polipéptidos microbianos incluyen uno o más polipéptidos de levadura o bacterias. Las fuentes de carbono de polipéptidos vegetales a modo de ejemplo incluyen uno o más polipéptidos de soja, maíz, canola, jatrofa, palma, cacahuete, girasol, coco, mostaza, colza, semilla de algodón, palmiste, oliva, cártamo, sésamo o linaza.

Las fuentes de carbono renovables a modo de ejemplo incluyen permeado de suero de queso, licor de maíz, melazas de remolacha azucarera, malta de cebada y componentes de cualquiera de los anteriores. Las fuentes de carbono renovables a modo de ejemplo también incluyen glucosa, hexosa, pentosa y xilosa presentes en la biomasa, tal como maíz, pasto varilla, caña de azúcar, desechos celulares de procesos de fermentación y subproductos de proteínas de la molienda de soja, maíz o trigo. En algunas realizaciones, la fuente de carbono de la biomasa es un material lignocelulósico, hemicelulósico o celulósico tal como, pero sin limitarse a, pasto, trigo, paja de trigo, bagazo, bagazo de caña de azúcar, pulpa de madera blanda, maíz, mazorca o cáscara de maíz, grano de maíz, fibra de granos de maíz, rastrojo de maíz, pasto varilla, producto de cáscara de arroz o un subproducto de la molienda de granos en húmedo o en seco (por ejemplo, granos de maíz, sorgo, centeno, triticale, cebada, trigo y/o granos de destilación). Los materiales celulósicos a modo de ejemplo incluyen madera, papel y residuos de pulpa, plantas herbáceas y pulpa de fruta. En algunas realizaciones, la fuente de carbono incluye cualquier parte de la planta, tal como tallos, granos, raíces o tubérculos. En algunas realizaciones, la totalidad o parte de cualquiera de las siguientes plantas se usan como fuente de carbono: maíz, trigo, centeno, sorgo, triticale, arroz, mijo, cebada, mandioca, legumbres, tales como alubias y guisantes, patatas, batatas, plátanos, caña de azúcar y/o tapioca. En algunas realizaciones, la fuente de carbono es un hidrolizado de biomasa, tal como un hidrolizado de biomasa que incluye tanto xilosa como glucosa o que incluye tanto sacarosa como glucosa.

En algunas realizaciones, la fuente de carbono renovable (tal como biomasa) se trata previamente antes de añadirse al medio de cultivo celular. En algunas realizaciones, el tratamiento previo incluye tratamiento previo enzimático, tratamiento previo químico o una combinación de tratamiento previo tanto enzimático como químico (véanse, por ejemplo, Farzaneh *et al.*, *Bioresource Technology* 96 (18): 2014-2018, 2005; patente estadounidense n.º 6.176.176; patente estadounidense n.º 6.106.888; a cada una de las cuales se hace referencia por el presente documento, particularmente con respecto al tratamiento previo de fuentes de carbono renovables). En algunas realizaciones, la fuente de carbono renovable se hidroliza parcial o completamente antes de añadirse al medio de cultivo celular.

En algunas realizaciones, la fuente de carbono renovable (tal como rastrojo de maíz) experimenta un tratamiento previo de expansión de fibras con amoníaco (AFEX) antes de añadirse al medio de cultivo celular (véase, por ejemplo, Farzaneh *et al.*, *Bioresource Technology* 96 (18): 2014-2018, 2005). Durante el tratamiento previo con AFEX, una fuente de carbono renovable se trata con amoníaco anhidro líquido a temperaturas moderadas (tal como de aproximadamente 60 a aproximadamente 100°C) y presión alta (tal como de aproximadamente 250 a aproximadamente 300 psi) durante aproximadamente 5 minutos. Entonces, la presión se libera rápidamente. En este proceso, los efectos químicos y físicos combinados de la solubilización de la lignina, la hidrólisis de la hemicelulosa, la descristalización de la celulosa y el aumento del área de superficie permiten la conversión enzimática casi completa de la celulosa y la hemicelulosa en azúcares fermentables. El tratamiento previo de AFEX tiene la ventaja de que casi todo el amoníaco puede recuperarse y reutilizarse, mientras que el resto sirve como fuente de nitrógeno para los microbios en procesos posteriores. Además, no se requiere una corriente de lavado para el tratamiento previo de AFEX. Por tanto, la recuperación de materia seca después del tratamiento de AFEX es esencialmente del 100%. AFEX es básicamente un proceso seco a seco. La fuente de carbono renovable tratada es estable durante largos períodos de tiempo y puede alimentarse a cargas sólidas muy altas en procesos de hidrólisis enzimática o fermentación. La celulosa y hemicelulosa se conservan bien en el proceso de AFEX, con poca o ninguna degradación. No hay necesidad de neutralización antes de la hidrólisis enzimática de una fuente de carbono renovable que se ha sometido a un tratamiento previo de AFEX. La hidrólisis enzimática de fuentes de carbono tratadas con AFEX produce corrientes de azúcar limpia para su uso en fermentación posterior.

En algunas realizaciones, la concentración de la fuente de carbono (por ejemplo, una fuente de carbono renovable) es equivalente a al menos o aproximadamente el 0,1, el 0,5, el 1, el 1,5, el 2, el 3, el 4, el 5, el 10, el 15, el 20, el 30, el 40 o el 50% de glucosa (p/v). La cantidad equivalente de glucosa puede determinarse usando métodos de HPLC convencionales con glucosa como referencia para medir la cantidad de glucosa generada a partir de la fuente de carbono. En algunas realizaciones, la concentración de la fuente de carbono (por ejemplo, una fuente de carbono renovable) es equivalente a entre aproximadamente el 0,1 y aproximadamente el 20% de glucosa, tal como entre aproximadamente el 0,1 y aproximadamente el 10% de glucosa, entre aproximadamente el 0,5 y aproximadamente el 10% de glucosa, entre aproximadamente el 1 y aproximadamente el 10% de glucosa, entre aproximadamente el 1 y aproximadamente el 5% de glucosa o entre aproximadamente el 1 y aproximadamente el 2% de glucosa.

En algunas realizaciones, la fuente de carbono incluye extracto de levadura o uno o más componentes de extracto de levadura. En algunas realizaciones, la concentración de extracto de levadura es de al menos 1 gramo de extracto de levadura por litro de caldo (g/l, en donde el volumen de caldo incluye tanto el volumen del medio celular como el volumen de las células), tal como al menos o aproximadamente 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300 o más g/l. En algunas realizaciones, la concentración de extracto de levadura es de entre aproximadamente 1 y aproximadamente 300 g/l, tal como entre aproximadamente 1 y aproximadamente 200 g/l, entre aproximadamente 5 y aproximadamente 200 g/l, entre aproximadamente 5 y aproximadamente 100 g/l o entre aproximadamente 5 y aproximadamente 60 g/l. En algunas realizaciones, la concentración incluye la cantidad total de extracto de levadura que se añade antes y/o durante el cultivo de las células huésped. En algunas realizaciones, la fuente de carbono incluye tanto extracto de levadura (o uno o más componentes del mismo) como otra fuente de carbono, tal como glucosa. En algunas realizaciones, la razón de extracto de levadura con respecto a la otra fuente de carbono es de aproximadamente 1:5, aproximadamente 1:10 o aproximadamente 1:20 (p/p).

Adicionalmente la fuente de carbono puede ser también sustratos de un carbono tales como dióxido de carbono o metanol. Se ha notificado la producción de glicerol a partir de fuentes de un solo carbono (por ejemplo, metanol, formaldehído o formiato) en levaduras metilotróficas (Yamada *et al.*, *Agric. Biol. Chem.*, 53(2) 541-543, 1989, al que se hace referencia por el presente documento, particularmente con respecto a las fuentes de carbono) y en bacterias (Hunter *et al.*, *Biochemistry*, 24, 4148-4155, 1985, al que se hace referencia por el presente documento, particularmente con respecto a las fuentes de carbono). Estos organismos pueden asimilar compuestos de un solo carbono, cuyo estado de oxidación oscila entre metano y formiato, y producen glicerol. La ruta de asimilación de carbono puede ser a través de monofosfato de ribulosa, a través de serina o a través de monofosfato de xilulosa (Gottschalk, *Bacterial Metabolism*, segunda edición, Springer-Verlag: Nueva York, 1986, al que se hace referencia por el presente documento, particularmente con respecto a las fuentes de carbono). La ruta de monofosfato de ribulosa implica la condensación de formiato con ribulosa-5-fosfato para formar un azúcar de seis carbonos que se convierte en fructosa y finalmente el producto de tres carbonos gliceraldehído-3-fosfato. Asimismo, la ruta de serina asimila el compuesto de un carbono en la ruta glicolítica por medio de metilentetrahidrofolato.

55 Sintegas

Puede usarse sintegas como fuente de energía y/o carbono para cualquiera de las células clostridiales recombinantes descritas en el presente documento. El sintegas puede incluir CO y H<sub>2</sub>. En algunos aspectos, el sintegas comprende CO, CO<sub>2</sub> y H<sub>2</sub>. En algunos aspectos, el sintegas comprende además H<sub>2</sub>O y/o N<sub>2</sub>. Por ejemplo, el sintegas puede comprender CO, H<sub>2</sub> y H<sub>2</sub>O (por ejemplo, CO, H<sub>2</sub>, H<sub>2</sub>O y N<sub>2</sub>). El sintegas puede comprender CO, H<sub>2</sub> y N<sub>2</sub>. El sintegas puede comprender CO, CO<sub>2</sub>, H<sub>2</sub> y H<sub>2</sub>O (por ejemplo, CO, CO<sub>2</sub>, H<sub>2</sub>, H<sub>2</sub>O y N<sub>2</sub>). El sintegas puede comprender CO, CO<sub>2</sub>, H<sub>2</sub> y N<sub>2</sub>. El CO y/o CO<sub>2</sub> en el gas de síntesis puede usarse como fuente de carbono para las células.

En algunos aspectos, la razón molar de hidrógeno con respecto a monóxido de carbono en el sintegas es aproximadamente cualquiera de 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2,0, 3,0, 4,0, 5,0 ó 10,0. En algunos aspectos, el sintegas comprende aproximadamente cualquiera del 10, el 20, el 30, el 40, el 50, el 60, el 70, el 80 o el 90% en volumen de monóxido de carbono. En algunos aspectos, el sintegas comprende aproximadamente cualquiera del 10, el 20, el 30, el 40, el 50, el 60, el 70, el 80 o el 90% en volumen de hidrógeno. En algunos aspectos, el sintegas comprende aproximadamente cualquiera del 10, el 20, el 30, el 40, el 50, el 60, el 70, el 80 o el 90% en volumen de dióxido de carbono. En algunos aspectos, el sintegas comprende aproximadamente cualquiera del 10, el 20, el 30, el 40, el 50, el 60, el 70, el 80 o el 90% en volumen de agua. En algunos aspectos, el sintegas comprende aproximadamente cualquiera del 10, el 20, el 30, el 40, el 50, el 60, el 70, el 80 o el 90% en volumen de nitrógeno.

El gas de síntesis usado según la presente invención puede derivarse de fuentes naturales o sintéticas. En algunos aspectos, el sintegas se deriva de biomasa (por ejemplo, madera, pasto varilla, desechos agrícolas, desechos municipales) o hidratos de carbono (por ejemplo, azúcares). En otros aspectos, el sintegas se deriva de carbón, petróleo, querógeno, arenas alquitranadas, esquisto bituminoso, gas natural, o una mezcla de los mismos. En otros aspectos, el sintegas se deriva de caucho, tal como de neumáticos de caucho. En algunos aspectos, el sintegas se deriva de una mezcla (por ejemplo, combinación) de biomasa y carbón. En algunos aspectos, la mezcla tiene aproximadamente o al menos aproximadamente cualquiera del 1%, el 2%, el 5%, el 10%, el 15%, el 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 90%, el 95% o el 99% de biomasa. En algunos aspectos, la mezcla tiene aproximadamente o al menos aproximadamente cualquiera del 1%, el 2%, el 5%, el 10%, el 15%, el 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 90%, el 95% o el 99% de carbón. En algunos aspectos, la razón de biomasa con respecto a carbón en la mezcla es de aproximadamente cualquiera de 5:95, 10:90, 15:85, 20:80, 25:75, 30:70, 35:65, 40:60, 45:55, 50:50, 55:45, 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10 ó 95:5.

El sintegas puede derivarse de una materia prima mediante una variedad de procedimientos, incluyendo reformado con metano, licuefacción de carbón, cocción conjunta, reacciones fermentativas, reacciones enzimáticas y gasificación de biomasa. La gasificación de biomasa se consigue sometiendo la biomasa a oxidación parcial en un reactor a temperaturas por encima de aproximadamente 700°C en presencia de menos que una cantidad estequiométrica de oxígeno. El oxígeno se introduce en el biorreactor en forma de aire, oxígeno puro o vapor. La gasificación puede producirse en tres etapas principales: 1) calentamiento inicial para secar cualquier humedad incrustada en la biomasa; 2) pirólisis, en la que la biomasa se calienta hasta 300-500°C en ausencia de agentes oxidantes para producir gas, alquitranes, aceites y residuos de carbón sólidos; y 3) gasificación de carbón sólido, alquitranes y gas para producir los componentes primarios del sintegas. La cocción conjunta se consigue mediante gasificación de una mezcla de carbón/biomasa. La composición del sintegas, tal como la identidad y razones molares de los componentes del sintegas, puede variar dependiendo de la materia prima de la que se deriva y el método mediante el que la materia prima se convierte en sintegas.

El gas de síntesis puede contener impurezas, cuya naturaleza y cantidad varían según tanto la materia prima como el proceso usado en la producción. Las fermentaciones pueden ser tolerantes con algunas impurezas, pero sigue habiendo la necesidad de eliminar del sintegas materiales tales como alquitranes y materiales particulados que podrían obstruir el fermentador y equipo asociado. También es aconsejable eliminar compuestos que podrían contaminar el producto de isopreno tales como compuestos orgánicos volátiles, gases ácidos, metano, benceno, tolueno, etilbenceno, xilenos, H<sub>2</sub>S, COS, CS<sub>2</sub>, HCl, O<sub>3</sub>, compuestos de organoazufre, amoniaco, óxidos de nitrógeno, compuestos orgánicos que contienen nitrógeno y vapores de metales pesados. La eliminación de las impurezas del sintegas puede lograrse mediante uno de varios medios, incluyendo lavado del gas, tratamiento con adsorbentes de fase sólida y purificación usando membranas permeables a los gases.

Se describen ejemplos de otros sistemas de fermentación y condiciones de cultivo que pueden usarse en las publicaciones de solicitud de patente internacional n.ºs WO2009/076676, WO2010/003007, WO2009/132220, WO2010/031062, WO2010/031068, WO2010/031076, WO2010/013077, WO2010/031079, WO2010/148150, WO2010/078457 y WO2010/148256, que se incorporan por el presente documento en su totalidad, particularmente con respecto a sistemas de fermentación y condiciones de cultivo para bacterias clostridiales.

En algunos aspectos, el medio de cultivo se prepara usando técnicas anóxicas. En algunos aspectos, el medio de cultivo comprende uno o más de NH<sub>4</sub>Cl, NaCl, KCl, KH<sub>2</sub>PO<sub>4</sub>, MgSO<sub>4</sub>·7H<sub>2</sub>O, CaCl<sub>2</sub>·2H<sub>2</sub>O, NaHCO<sub>3</sub>, extracto de levadura, clorhidrato de cisteína, Na<sub>2</sub>S·9H<sub>2</sub>O, oligoelementos y vitaminas. En algunos aspectos, el medio de cultivo contiene, por litro, aproximadamente 1,0 g de NH<sub>4</sub>Cl, aproximadamente 0,8 g de NaCl, aproximadamente 0,1 g de

KCl, aproximadamente 0,1 g de  $\text{KH}_2\text{PO}_4$ , aproximadamente 0,2 g de  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , aproximadamente 0,02 g de  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , aproximadamente 1,0 g de  $\text{NaHCO}_3$ , aproximadamente 1,0 g de extracto de levadura, aproximadamente 0,2 g de clorhidrato de cisteína, aproximadamente 0,2 g de  $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$ , aproximadamente 10 ml de disolución de oligoelementos y aproximadamente 10 ml de disolución de vitaminas. En algunos aspectos, la

condición de cultivo comprende mevalonato.

Las condiciones de cultivo, fuentes de carbono, fuentes de energía y medios de cultivo pueden ser según cualquiera de las condiciones de crecimiento, fuentes de carbono, fuentes de energía y medios de cultivo descritos en los ejemplos de la presente divulgación.

#### *Sistemas de expresión clostridial*

La presente divulgación proporciona sistemas de expresión de *Clostridium* para la producción de uno o más bioproductos industriales (por ejemplo, isopreno, butadieno o etanol). En algunas realizaciones de la divulgación el sistema puede incluir uno o más de: (a) una metiltransferasa (por ejemplo, un plásmido que comprende pDW268 o pMCS466), (b) un plásmido lanzadera (por ejemplo, pDW280, pMCS537, pMCS200, pMCS201, pMCS444 o pMCS445), (c) una célula bacteriana de *E. coli* capaz de interactuar con una célula bacteriana de *Clostridium* para permitir la transferencia de (a) y (b); y (d) una célula bacteriana de *Clostridium* capaz de interactuar con una célula bacteriana de *Escherichia* de manera que uno o más ácidos nucleicos se expresan en la célula bacteriana de *Clostridium*. En algunas realizaciones de la divulgación, la célula de bacterias de *E. coli* capaz de interactuar con una célula bacteriana de *Clostridium* es una célula de *E. coli* S17-1. En algunas realizaciones de la divulgación, la célula bacteriana de *Clostridium* capaz de interactuar con una célula bacteriana de *Escherichia* se selecciona del grupo de *Clostridium acetivum*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum* o *Clostridium acetobutylicum*. En algunas realizaciones, el sistema proporciona la expresión de uno o más ácidos nucleicos de interés (por ejemplo, ácidos nucleicos que codifican para isopreno sintasa o enzimas implicadas en la producción de etanol a partir de acetil-CoA). Tal como se describe en el presente documento, el sistema de restricción-modificación clostridial puede usarse para modificar por ingeniería genética células clostridiales de modo que el sistema de restricción-modificación pueda evitarse. Esta modificación por ingeniería genética permite el uso de las células clostridiales para producir diversos bioproductos industriales, incluyendo pero sin limitarse a, isopreno, butadieno, etanol, propanodiol (por ejemplo, 1,2-propanodiol, 1,3-propanodiol), hidrógeno, acetato, combustibles microbianos, alcoholes no fermentativos, alcoholes grasos, ésteres de ácidos grasos, alcoholes isoprenoides, alquenos, alcanos, terpenoides, isoprenoides, carotenoides u otros productos C5, C10, C15, C20, C25, C30, C35 o C40. La producción de estos bioproductos industriales se describe en detalle adicional a continuación y en el presente documento.

#### *Métodos de uso de bacterias clostridiales modificadas por ingeniería genética para la producción de bioproductos industriales*

Tal como se describe en el presente documento, el sistema de restricción-modificación clostridial puede usarse para modificar por ingeniería genética células clostridiales de modo que el sistema de restricción-modificación pueda evitarse. Esta modificación por ingeniería genética permite el uso de las células clostridiales para producir diversos bioproductos industriales, incluyendo pero sin limitarse a, isopreno, butadieno, etanol, propanodiol (por ejemplo, 1,2-propanodiol, 1,3-propanodiol), hidrógeno, acetato, combustibles microbianos, alcoholes no fermentativos, alcoholes grasos, ésteres de ácidos grasos, alcoholes isoprenoides, alquenos, alcanos, terpenoides, isoprenoides, carotenoides u otros productos C5, C10, C15, C20, C25, C30, C35 o C40. La producción de estos bioproductos industriales se describe en detalle adicional a continuación y en el presente documento.

#### *Producción de isopreno*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales que contienen una o más rutas para la producción de isopreno (por ejemplo, bacterias clostridiales que contienen las rutas ilustradas en la figura 41 a la figura 45) con uno o más polinucleótidos heterólogos que codifican para una o más enzimas de la ruta del isopreno expresadas en una cantidad suficiente para producir isopreno.

#### *Polipéptidos y ácidos nucleicos de isopreno sintasa a modo de ejemplo*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales con polinucleótidos que codifican para un polipéptido de isopreno sintasa. Los polipéptidos de isopreno sintasa convierten fosfato de dimetilalilo (DMAPP) en isopreno. Los polipéptidos de isopreno sintasa a modo de ejemplo incluyen polipéptidos, fragmentos de polipéptidos, péptidos y polipéptidos de fusión que tienen al menos una actividad de un polipéptido de isopreno sintasa. Pueden usarse métodos convencionales para determinar si un polipéptido tiene actividad de polipéptido de isopreno sintasa midiendo la capacidad del polipéptido para convertir DMAPP en isopreno *in vitro*, en un extracto celular, o *in vivo* (por ejemplo, tal como se describe en el ejemplo 1 del documento US 420360 B2, que se incorpora en el presente documento en su totalidad, particularmente con respecto a métodos para evaluar la actividad isopreno sintasa). La actividad del polipéptido de isopreno sintasa en extractos celulares puede medirse, por ejemplo, tal como se describe en Silver *et al.*, J. Biol. Chem. 270:13010-13016, 1995 y referencias en el mismo, a cada una de las cuales se hace referencia por el presente documento, particularmente

con respecto a ensayos para la actividad del polipéptido de isopreno sintasa.

En algunas realizaciones, el polipéptido o ácido nucleico de isopreno sintasa es de la familia *Fabaceae*, tal como la subfamilia *Faboideae*. En algunas realizaciones, el polipéptido o ácido nucleico de isopreno sintasa es un polipéptido o ácido nucleico que se produce de manera natural de *Pueraria montana* (kudzu) (Sharkey *et al.*, Plant Physiology 137: 700-712, 2005), *Pueraria lobata*, chopo (tal como *Populus alba x tremula* CAC35696) Miller *et al.*, Planta 213: 483-487, 2001), álamo (tal como *Populus tremuloides*) Silver *et al.*, JBC 270(22): 13010-1316, 1995) o roble inglés (*Quercus robur*) (Zimmer *et al.*, documento WO 98/02550), a cada uno de los cuales se hace referencia por el presente documento, particularmente con respecto a ácidos nucleicos de isopreno sintasa y la expresión de polipéptidos de isopreno sintasa. Las isopreno sintasas adecuadas incluyen, pero no se limitan a, las identificadas por los n.ºs de registro de Genbank AY341431, AY316691, AY279379, AJ457070 y AY1 82241, que se incorporan cada una por el presente documento como referencia en su totalidad, particularmente con respecto a secuencias de ácidos nucleicos y polipéptidos de isopreno sintasa. En algunas realizaciones, el polipéptido o ácido nucleico de isopreno sintasa no es un polipéptido o ácido nucleico que se produce de manera natural de *Quercus robur* (es decir, el polipéptido o ácido nucleico de isopreno sintasa es un polipéptido o ácido nucleico de isopreno sintasa distinto de un polipéptido o ácido nucleico que se produce de manera natural de *Quercus robur*). En algunas realizaciones, el ácido nucleico o polipéptido de isopreno sintasa no es un polipéptido o ácido nucleico que se produce de manera natural de chopo (tal como *Populus alba x tremula* CAC35696).

Los ácidos nucleicos de isopreno sintasa a modo de ejemplo incluyen ácidos nucleicos que codifican para un polipéptido, fragmento de un polipéptido, péptido o polipéptido de fusión que tiene al menos una actividad de un polipéptido de isopreno sintasa. Los polipéptidos y ácidos nucleicos de isopreno sintasa a modo de ejemplo incluyen polipéptidos y ácidos nucleicos que se producen de manera natural a partir de cualquiera de los organismos fuente descritos en el presente documento así como polipéptidos y ácidos nucleicos mutantes derivados de cualquiera de los organismos fuente descritos en el presente documento.

#### *Polipéptidos y ácidos nucleicos de DXS a modo de ejemplo*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales con polinucleótidos que codifican para polipéptidos de 1-desoxi-D-xilulosa-5-fosfato sintasa (DXS). Los polipéptidos de DSX convierten piruvato y D-gliceraldehído-3-fosfato en 1-desoxi-D-xilulosa-5-fosfato. Los polipéptidos de DXS a modo de ejemplo incluyen polipéptidos, fragmentos de polipéptidos, péptidos y polipéptidos de fusión que tienen al menos una actividad de un polipéptido de DXS. Pueden usarse métodos convencionales para determinar si un polipéptido tiene actividad de polipéptido de DXS midiendo la capacidad del polipéptido para convertir piruvato y D-gliceraldehído-3-fosfato en 1-desoxi-D-xilulosa-5-fosfato *in vitro*, en un extracto celular, o *in vivo* (véase, por ejemplo, el documento US 8420360 B2, que se incorpora por el presente documento en el presente documento en su totalidad, particularmente con respecto a métodos de evaluación de la actividad del polipéptido de DXS). Los ácidos nucleicos de DXS a modo de ejemplo incluyen ácidos nucleicos que codifican para un polipéptido, fragmento de un polipéptido, péptido o polipéptido de fusión que tiene al menos una actividad de un polipéptido de DXS. Los polipéptidos y ácidos nucleicos de DXS a modo de ejemplo incluyen polipéptidos y ácidos nucleicos que se producen de manera natural de cualquiera de los organismos fuente descritos en el presente documento así como polipéptidos y ácidos nucleicos mutantes derivados de cualquiera de los organismos fuente descritos en el presente documento.

#### *Polipéptidos y ácidos nucleicos de IDI a modo de ejemplo*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales con polinucleótidos que codifican para polipéptidos de isopentenil difosfato isomerasa (isopentenil-difosfato delta-isomerasa o IDI). IDI cataliza la interconversión de difosfato de isopentenilo (IPP) y difosfato de dimetilalilo (DMAPP) (por ejemplo, conversión de IPP en DMAPP y/o conversión de DMAPP en IPP). Los polipéptidos de IDI a modo de ejemplo incluyen polipéptidos, fragmentos de polipéptidos, péptidos y polipéptidos de fusión que tienen al menos una actividad de un polipéptido de IDI. Pueden usarse métodos convencionales para determinar si un polipéptido tiene actividad de polipéptido de IDI midiendo la capacidad del polipéptido para interconvertir IPP y DMAPP *in vitro*, en un extracto celular, o *in vivo* (véase, por ejemplo, el documento US 8420360 B2, al que se hace referencia por el presente documento, particularmente con respecto a ensayos para la actividad de IDI). Los ácidos nucleicos de IDI a modo de ejemplo incluyen ácidos nucleicos que codifican para un polipéptido, fragmento de un polipéptido, péptido o polipéptido de fusión que tiene al menos una actividad de un polipéptido de IDI. Los polipéptidos y ácidos nucleicos de IDI a modo de ejemplo incluyen polipéptidos y ácidos nucleicos que se producen de manera natural a partir de cualquiera de los organismos fuente descritos en el presente documento así como polipéptidos y ácidos nucleicos mutantes derivados de cualquiera de los organismos fuente descritos en el presente documento.

#### *Polipéptidos y ácidos nucleicos de la ruta de MVA a modo de ejemplo*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales con polinucleótidos que codifican para polipéptidos de la ruta de MVA. Los polipéptidos de la ruta de

MVA incluyen polipéptidos de acetil-CoA acetiltransferasa (AA-CoA tiolasa), polipéptidos de 3-hidroxi-3-metilglutaril-CoA sintasa (HMG-CoA sintasa), polipéptidos de 3-hidroxi-3-metilglutaril-CoA reductasa (HMG-CoA reductasa), polipéptidos de mevalonato cinasa (MVK), polipéptidos de fosfomevalonato cinasa (PMK), polipéptidos de difosfomevalonato descarboxilasa (MVD), polipéptidos de IDI y polipéptidos (por ejemplo, polipéptidos de fusión) que tienen una actividad de dos o más polipéptidos de la ruta de MVA. En particular, los polipéptidos de la ruta de MVA incluyen polipéptidos, fragmentos de polipéptidos, péptidos y polipéptidos de fusión que tienen al menos una actividad de un polipéptido de la ruta de MVA. Los ácidos nucleicos de la ruta de MVA a modo de ejemplo incluyen ácidos nucleicos que codifican para un polipéptido, fragmento de un polipéptido, péptido o polipéptido de fusión que tiene al menos una actividad de un polipéptido de la ruta de MVA. Los polipéptidos y ácidos nucleicos de la ruta de MVA a modo de ejemplo incluyen polipéptidos y ácidos nucleicos que se producen de manera natural a partir de cualquiera de los organismos fuente descritos en el presente documento así como polipéptidos y ácidos nucleicos mutantes derivados de cualquiera de los organismos fuente descritos en el presente documento.

En particular, polipéptidos de acetil-CoA acetiltransferasa (AA-CoA tiolasa o AACT) convierten dos moléculas de acetil-CoA en acetoacetil-CoA. Pueden usarse métodos convencionales (tales como los descritos en el presente documento) para determinar si un polipéptido tiene actividad de polipéptido de AA-CoA tiolasa midiendo la capacidad del polipéptido para convertir dos moléculas de acetil-CoA en acetoacetil-CoA *in vitro*, en un extracto celular, o *in vivo*.

Los polipéptidos de 3-hidroxi-3-metilglutaril-CoA sintasa (HMG-CoA sintasa o HMGS) convierten acetoacetil-CoA en S-hidroxi-S-metilglutaril-CoA. Pueden usarse métodos convencionales (tales como los descritos en el presente documento) para determinar si un polipéptido tiene actividad de polipéptido de HMG-CoA sintasa midiendo la capacidad del polipéptido para convertir acetoacetil-CoA en 3-hidroxi-3-metilglutaril-CoA *in vitro*, en un extracto celular, o *in vivo*.

Los polipéptidos de 3-hidroxi-3-metilglutaril-CoA reductasa (HMG-CoA reductasa o HMGR) convierten 3-hidroxi-3-metilglutaril-CoA en mevalonato. Pueden usarse métodos convencionales (tales como los descritos en el presente documento) para determinar si un polipéptido tiene actividad de polipéptido de HMG-CoA reductasa midiendo la capacidad del polipéptido para convertir 3-hidroxi-3-metilglutaril-CoA en mevalonato *in vitro*, en un extracto celular, o *in vivo*.

Los polipéptidos de mevalonato cinasa (MVK) fosforilan mevalonato para formar mevalonato-5-fosfato. Pueden usarse métodos convencionales (tales como los descritos en el presente documento) para determinar si un polipéptido tiene actividad de polipéptido de MVK midiendo la capacidad del polipéptido para convertir mevalonato en mevalonato-5-fosfato *in vitro*, en un extracto celular, o *in vivo*.

Los polipéptidos de fosfomevalonato cinasa (PMK) fosforilan mevalonato-5-fosfato para formar mevalonato-5-difosfato. Pueden usarse métodos convencionales (tal como los descritos en el presente documento) para determinar si un polipéptido tiene actividad de polipéptido de PMK midiendo la capacidad del polipéptido para convertir mevalonato-5-fosfato en mevalonato-5-difosfato *in vitro*, en un extracto celular, o *in vivo*.

Los polipéptidos de difosfomevalonato descarboxilasa (MVD o DPMD) convierten mevalonato-5-difosfato en polipéptidos de difosfato de isopentenilo (IPP). Pueden usarse métodos convencionales (tales como los descritos) para determinar si un polipéptido tiene actividad de polipéptido de MVD midiendo la capacidad del polipéptido para convertir mevalonato-5-difosfato en IPP *in vitro*, en un extracto celular, o *in vivo*.

Las composiciones y los métodos descritos en el presente documento pueden usarse para transformar bacterias clostridiales que se han modificado por ingeniería genética para producir isopreno a partir de sintegas y/o a partir de hidratos de carbono o mezclas de los mismos.

*Método de uso de células clostridiales modificadas por ingeniería genética para la producción de butadieno*

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales que contienen una o más rutas para la producción de butadieno (mostrado en la figura 46 a la figura 48) con uno o más polinucleótidos heterólogos que codifican para una o más enzimas de la ruta de butadieno expresadas en una cantidad suficiente para producir butadieno. La ruta de butadieno incluye una acetil-CoA:acetil-CoA aciltransferasa, una acetoacetil-CoA reductasa, una 3-hidroxi-3-metilglutaril-CoA reductasa (que forma aldehído), una crotonaldehído reductasa (que forma alcohol), una alcohol crotilico cinasa, una 2-butenil-4-fosfato cinasa, una butadieno sintasa, una crotonil-CoA hidrolasa, sintetasa o transferasa, una crotonato reductasa, una crotonil-CoA reductasa (que forma alcohol), una glutaconil-CoA descarboxilasa, una glutaril-CoA deshidrogenasa, una 3-aminobutil-CoA desaminasa, una 4-hidroxi-4-butil-CoA deshidratasa o una alcohol crotilico difosfocinasa. La producción de butadieno a partir de bacterias se describe en el documento WO 2011/140171 A2, al que se hace referencia por el presente documento, particularmente con respecto a las rutas para la producción de butadieno a partir de acetil-CoA (figura 46), a partir de eritrosa-4-fosfato (figura 47) y a partir de malonil-CoA más acetil-CoA (figura 48).

*Método de uso de células clostridiales modificadas por ingeniería genética para la producción de etanol*

Se sabe que varias bacterias del género *Clostridium* producen etanol a través de la ruta de acetil-CoA, que puede utilizar tanto monóxido de carbono como hidrógeno como fuentes de carbono y como fuentes de energía. La producción de etanol a partir de bacterias clostridiales se describe en Kopke *et al.*, 2011, Fermentative production of ethanol from carbon monoxide, Current Opinion in Biotechnology, vol. 22:320-323, y en Wilkins *et al.*, 2011, Microbial production of ethanol from carbon monoxide, Current Opinion in Biotechnology, vol. 22:326-330, ambos de los cuales se incorporan por el presente documento en su totalidad, particularmente con respecto a su discusión de la ruta para la producción de etanol a partir de acetil-CoA en bacterias clostridiales.

Las composiciones y los métodos divulgados en el presente documento pueden usarse para transformar bacterias clostridiales que contienen la ruta de etanol (incluyendo, pero sin limitarse a *Clostridium acetivum*, *Clostridium ljungdahli*, *Clostridium acetobutylicum* o *Clostridium autoethanogenum*) con uno o más polinucleótidos heterólogos que codifican para una o más enzimas de la ruta de etanol expresadas en cantidad suficiente para producir etanol. En bacterias clostridiales, la ruta para la producción de etanol a partir de acetil-CoA incluye la enzima aldehído deshidrogenasa y la enzima alcohol deshidrogenasa (véase, por ejemplo, la figura 41).

*Método de uso de células clostridiales modificadas por ingeniería genética para la producción de otros bioproductos industriales*

En algunos aspectos de la presente divulgación, cualquiera de los métodos descritos en el presente documento puede usarse para producir productos distintos de isopreno, butadieno y etanol. Tales productos pueden excretarse, secretarse o ser productos intracelulares. Uno cualquiera de los métodos descritos en el presente documento puede usarse para producir isopreno y/o uno o más de los otros productos. Los productos descritos en el presente documento pueden ser, por ejemplo, propanodiol (por ejemplo, 1,2-propanodiol, 1,3-propanodiol), hidrógeno, acetato o combustibles microbianos. Combustibles microbianos a modo de ejemplo son alcoholes fermentativos (por ejemplo, etanol o butanol), alcoholes no fermentativos (por ejemplo, isobutanol, metilbutanol, 1-propanol, 1-butanol, metilpentanol o 1-hexanol), alcoholes grasos, ésteres de ácidos grasos, alcoholes isoprenoides, alquenos y alcanos. Los productos descritos en el presente documento puede ser también un terpenoide, isoprenoide (por ejemplo, farneseno) o carotenoide u otro producto C5, C10, C15, C20, C25, C30, C35 o C40.

En algunos aspectos, los terpenoides se seleccionan del grupo que consiste en hemiterpenoides, monoterpenoides, sesquiterpenoides, diterpenoides, sesterpenoides, triterpenoides, tetraterpenoides y politerpenoides superiores. En algunos aspectos, el hemiterpenoide es prenol, isoprenol o ácido isovalérico. En algunos aspectos, el monoterpenoide es pirofosfato de geranilo, eucaliptol, limoneno o pineno. En algunos aspectos, el sesquiterpenoide es pirofosfato de farnesilo, artemisinina o bisabolol. En algunos aspectos, el diterpenoide es pirofosfato de geranylgeranilo, retinol, retinal, fitol, taxol, forskolina o afidicolina. En algunos aspectos, el triterpenoide es escualeno o lanosterol. En algunos aspectos, el tetraterpenoide es licopeno o caroteno. En algunos aspectos, los carotenoides se seleccionan del grupo que consiste en xantofilas y carotenos. En algunos aspectos, la xantofila es luteína o zeaxantina. En algunos aspectos, el caroteno es  $\alpha$ -caroteno,  $\beta$ -caroteno,  $\gamma$ -caroteno,  $\beta$ -criptoxantina o licopeno.

Los productos descritos en el presente documento pueden derivarse de acetil-CoA producido por medio de fermentación de sintegas o por medio de fermentación de otras fuentes de carbono tales como fructosa. En algunos aspectos, la célula se hace crecer en condiciones adecuadas para la producción del/de los producto(s) distinto(s) de isopreno.

Los productos descritos en el presente documento puede producirlos de manera natural la célula. En algunos aspectos, las células producen de manera natural uno o más productos incluyendo productos excretados, secretados o intracelulares. En algunos aspectos, las células producen de manera natural etanol, propanodiol, hidrógeno o acetato. En algunos aspectos, la producción de un producto que se produce de manera natural aumenta en relación con células de tipo natural. Puede usarse cualquier método conocido en la técnica para aumentar la producción de un producto celular metabólico para aumentar la producción de un producto que se produce de manera natural. En algunos aspectos, el ácido nucleico que codifica para la totalidad o una parte de la ruta para la producción de un producto descrito en el presente documento está operativamente unido a un promotor tal como un promotor fuerte. En algunos aspectos, el ácido nucleico que codifica para la totalidad o una parte de la ruta para la producción de un producto descrito en el presente documento está operativamente unido a un promotor constitutivo. En algunos aspectos, la célula se modifica por ingeniería genética para que comprenda copias adicionales de un ácido nucleico endógeno que codifica para un polipéptido para la producción de un producto descrito en el presente documento. En algunos aspectos, el producto descrito en el presente documento no lo produce de manera natural la célula. En algunos aspectos, la célula comprende uno o más ácidos nucleicos heterólogos que codifican para uno o más polipéptidos para la producción de un producto descrito en el presente documento.

En condiciones de crecimiento normales, los acetógenos producen acetato y etanol. El acetato se produce en una reacción de 2 etapas en la que el acetil-CoA se convierte en primer lugar en acetil-fosfato mediante la fosfotransacetilasa (pta), luego el acetil-fosfato se desfosforila mediante la acetato cinasa (ack) para formar acetato. El etanol se forma mediante un proceso de dos etapas en el que el acetil-CoA se convierte en acetaldehído y luego

en etanol mediante la enzima multifuncional alcohol deshidrogenasa (*adhE*). La producción de acetato y etanol puede no ser deseable en células que producen isopreno, ya que aleja el flujo de carbono del isopreno y en última instancia da como resultado un rendimiento disminuido de isopreno. Por tanto, algunos o todos los genes que codifican para fosfotransacetilasa (*pta*), acetato cinasa (*ack*) y alcohol deshidrogenasa (*adhE*) pueden alterarse o las expresiones de los mismos se reducen en células anaerobias para el fin de redirigir el flujo de carbono lejos del acetato y/o etanol y aumentar la producción de isopreno.

En algunos aspectos, las células son deficientes en al menos un polipéptido implicado en la producción de acetato, etanol, succinato y/o glicerol. En algunos aspectos, una o más rutas para la producción de un metabolito distinto de isopreno (por ejemplo, lactato, acetato, etanol (u otro(s) alcohol(es)), succinato o glicerol) se bloquean, por ejemplo, la producción de un metabolito distinto de isopreno puede reducirse en al menos aproximadamente cualquiera del 10%, el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 90% o el 95%. En algunos aspectos, una o más de las rutas para la producción de lactato, acetato, etanol, succinato o glicerol se bloquean, por ejemplo, la producción de lactato, acetato, etanol, succinato y/o glicerol se reduce en al menos aproximadamente cualquiera del 10%, el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 90% o el 95%. En algunos aspectos, las células son deficientes en al menos un polipéptido en ruta(s) de producción de acetato, etanol, succinato y/o glicerol. Los polipéptidos en ruta(s) de producción de acetato, etanol, succinato y/o glicerol pueden tener actividades reducidas o las expresiones de los mismos se reducen. Pueden alterarse ácidos nucleicos que codifican para los polipéptidos en ruta(s) de producción de acetato, etanol, succinato y/o glicerol. Un experto en la técnica conoce polipéptidos implicados en diversas rutas (por ejemplo, rutas para producir etanol y/o acetato), incluyendo, por ejemplo, los descritos en Misoph *et al.* 1996, *Journal of Bacteriology*, 178(11):3140-45, a cuyo contenido se hace referencia expresamente con respecto a los polipéptidos implicados en rutas de producción de succinato, acetato, lactato y/o etanol.

En algunos aspectos, las células son deficientes en *pta*. En algunos aspectos, las células son deficientes en *ack*. En algunos aspectos, las células son deficientes en *adhE*. En algunos aspectos, las células son deficientes en *pta*, *ack*, y/o *adhE*. En algunos aspectos, las expresiones de fosfotransacetilasa, acetato cinasa y/o alcohol deshidrogenasa se reducen. En algunos aspectos, las actividades de fosfotransacetilasa, acetato cinasa y/o alcohol deshidrogenasa se reducen. En algunos aspectos, las células son deficientes en polipéptido(s) que tiene(n) actividades similares a fosfotransacetilasa, acetato cinasa y/o alcohol deshidrogenasa. La expresión de *pta*, *ack*, *adhE* y/o polipéptido(s) que tiene(n) actividades similares a fosfotransacetilasa, acetato cinasa y/o alcohol deshidrogenasa puede reducirse mediante cualquiera de los métodos conocidos por un experto en la técnica, por ejemplo, la expresión puede reducirse mediante ARN antisentido (por ejemplo, ARN antisentido dirigido por cualquiera de los promotores descritos en el presente documento tal como cualquiera de los promotores inducibles). En algunos aspectos, el/los ARN antisentido está(n) operativamente unido(s) a un promotor adecuado tal como cualquiera de los promotores descritos en el presente documento incluyendo promotores inducibles.

En algunos aspectos, se recuperan tanto isopreno como producto(s) distinto(s) de isopreno a partir de la fase gaseosa. En algunos aspectos, se recupera isopreno a partir de la fase gaseosa (por ejemplo a partir de la fermentación de gas), y el/los otro(s) producto(s) se recupera(n) de la fase líquida (por ejemplo del caldo celular).

#### *Biorreactores*

Puede usarse una variedad de diferentes tipos de reactores para la producción de isopreno u otros bioproductos industriales. En algunas realizaciones, se usa un hidrato de carbono como fuente de energía y/o carbono. En algunas realizaciones, se usan un hidrato de carbono e hidrógeno como fuente de energía y/o carbono. En algunas realizaciones, se usa gas de síntesis como fuente de energía y/o carbono. Hay un gran número de diferentes tipos de procesos de fermentación que se usan comercialmente. Los biorreactores para su uso en la presente invención deben ser susceptibles a condiciones anaerobias. El biorreactor puede diseñarse para optimizar el tiempo de retención de las células, el tiempo de residencia del líquido y la velocidad de rociado de sintegas.

En diversos aspectos, las células se hacen crecer usando cualquier modo conocido de fermentación, tales como procesos discontinuos, de alimentación discontinua, continuos o continuos con recirculación. En algunos aspectos, se utiliza un método de fermentación discontinua. La fermentación discontinua clásica es un sistema cerrado donde la composición de los medios se establece al comienzo de la fermentación y no está sujeta a alteraciones artificiales durante la fermentación. Por tanto, al comienzo de la fermentación, el medio celular se inocula con las células huésped deseadas y se permite que se produzca la fermentación sin añadir nada al sistema. Normalmente, sin embargo, la fermentación "discontinua" es discontinua con respecto a la adición de la fuente de carbono y con frecuencia se hacen intentos de controlar factores tales como el pH y la concentración de oxígeno. En los sistemas discontinuos, las composiciones de metabolito y biomasa del sistema cambian constantemente hasta el momento en que se detiene la fermentación. Dentro de los cultivos discontinuos, las células se moderan a través de una fase de latencia estática a una fase logarítmica de alto crecimiento y finalmente a una fase estacionaria donde la tasa de crecimiento disminuye o se detiene. En algunos aspectos, las células en la fase logarítmica son responsables de la mayor parte de la producción de isopreno. En algunos aspectos, las células en fase estacionaria producen isopreno.

En algunos aspectos, se usa una variación en el sistema discontinuo convencional, como el sistema de alimentación

discontinua. Los procesos de fermentación de alimentación discontinua comprenden un sistema discontinuo típico con la excepción de que la fuente de carbono (por ejemplo, sintegas, glucosa, fructosa) se añade en incrementos a medida que avanza la fermentación. Los sistemas de alimentación discontinua son útiles cuando la represión por catabolitos es apta para inhibir el metabolismo de las células y cuando es deseable tener cantidades limitadas de fuente de carbono en el medio celular. Las fermentaciones de alimentación discontinua pueden realizarse con la fuente de carbono (por ejemplo, sintegas, glucosa, fructosa) en una cantidad limitada o en exceso. La medición de la concentración real de la fuente de carbono en los sistemas de alimentación discontinua es difícil y, por tanto, se estima basándose en los cambios de factores medibles tales como el pH, el oxígeno disuelto y la presión parcial de los gases de desecho, tales como CO<sub>2</sub>. Las fermentaciones discontinuas y de alimentación discontinua son comunes y se conocen bien en la técnica y pueden encontrarse ejemplos en Brock, *Biotechnology: A Textbook of Industrial Microbiology*, segunda edición (1989) Sinauer Associates, Inc..

En algunos aspectos, se usan métodos de fermentación continua. La fermentación continua es un sistema abierto en el que un medio de fermentación definido se añade continuamente a un biorreactor y se retira simultáneamente una cantidad igual de medio acondicionado para su procesamiento. La fermentación continua mantiene generalmente los cultivos en una densidad alta y constante donde las células se encuentran principalmente en el crecimiento en fase logarítmica.

La fermentación continua permite la modulación de un factor o cualquier número de factores que afectan el crecimiento celular o la producción de isopreno. Por ejemplo, un método mantiene un nutriente limitante, tal como la fuente de carbono o el nivel de nitrógeno a una tasa fija y permite que todos los demás parámetros se moderen. En otros sistemas, varios factores que afectan el crecimiento pueden alterarse continuamente mientras la concentración celular (por ejemplo, la concentración medida por la turbidez del medio) se mantiene constante. Los sistemas continuos se esfuerzan por mantener las condiciones de crecimiento en estado estacionario. Por tanto, la pérdida celular debida a la extracción de los medios se equilibra con la velocidad de crecimiento celular en la fermentación. Los métodos de modulación de nutrientes y factores de crecimiento para procesos de fermentación continua, así como las técnicas para maximizar la velocidad de formación de productos, se conocen bien en la técnica de microbiología industrial y se detallan una variedad de métodos en Brock, *Biotechnology: A Textbook of Industrial Microbiology*, segunda edición (1989) Sinauer Associates, Inc., al que se hace referencia por el presente documento, particularmente con respecto al cultivo celular y las condiciones de fermentación.

Una variación del método de fermentación continua es el método continuo con recirculación. Este sistema es similar al biorreactor continuo, con la diferencia de que las células retiradas con el contenido líquido se devuelven al biorreactor por medio de un dispositivo de separación de masa celular. Se usan unidades de filtración cruzada, centrifugadoras, tanques de sedimentación, astillas de madera, hidrogeles y/o fibras huecas para la separación o retención de la masa celular. Este proceso se usa normalmente para aumentar la productividad del sistema de biorreactor continuo, y puede ser particularmente útil para los anaerobios, que pueden crecer más lentamente y en concentraciones más bajas que los aerobios.

En un aspecto, puede usarse un biorreactor de membrana para el crecimiento y/o la fermentación de las células anaerobias descritas en el presente documento, en particular, si se espera que las células crezcan lentamente. Un filtro de membrana, como un filtro de flujo cruzado o un filtro de flujo tangencial, puede funcionar conjuntamente con un biorreactor de fermentación líquida que produce gas isopreno. Un biorreactor de membrana de este tipo puede mejorar la producción fermentativa de gas isopreno combinando la fermentación con la recirculación de componentes del caldo seleccionados que, de lo contrario, se descartarían. El MBR filtra el caldo de fermentación y devuelve el componente que no permea ("material retenido" del filtro) al reactor, aumentando eficazmente la concentración en el reactor de las células, los residuos celulares y otros sólidos del caldo, mientras mantiene la productividad específica de las células. Esto mejora sustancialmente el título, la producción total y la productividad volumétrica del isopreno, lo que conduce a menores costes de capital y operativos.

El filtrado líquido (o permeado) no se devuelve al reactor y, por tanto, proporciona una reducción beneficiosa en el volumen del reactor, similar a la recolección de una extracción de caldo. Sin embargo, a diferencia de una extracción de caldo, el permeado recogido es un líquido clarificado que puede esterilizarse fácilmente por filtración después de su almacenamiento en un recipiente ordinario. Por tanto, el permeado puede reutilizarse fácilmente como fuente de reciclaje de nutrientes y/o agua. Un permeado, que contiene medio agotado soluble, puede añadirse a la misma u otra fermentación para mejorar la producción de isopreno.

#### *Métodos de recuperación*

Cualquiera de los métodos descritos en el presente documento incluye además recuperar el bioproducto industrial (por ejemplo, isopreno, butadieno, etanol, etc.). Por ejemplo, el isopreno producido usando las composiciones y los métodos de la invención puede recuperarse usando técnicas convencionales, tales como separación de gas, separación potenciada por membrana, fraccionamiento, adsorción/desorción, evaporación, desorción térmica o a vacío de isopreno a partir de una fase sólida o extracción de isopreno inmovilizado o absorbido a una fase sólida con un disolvente (véanse, por ejemplo, las patentes estadounidenses n.ºs 4.703.007 y 4.570.029). En un aspecto, el isopreno se recupera mediante extracción por absorción (véase, por ejemplo, la solicitud de patente internacional n.º

PCT/US2010/060552 (WO 2011/075534)). En aspectos particulares, se usa destilación extractiva con un alcohol (tal como etanol, metanol, propanol, o una combinación de los mismos) para recuperar el isopreno. En algunos aspectos, la recuperación de isopreno implica el aislamiento de isopreno en una forma líquida (tal como una disolución pura de isopreno o una disolución de isopreno en un disolvente). La separación de gas implica la retirada de vapor de isopreno de la corriente de gas residual de fermentación de una manera continua. Tal eliminación puede lograrse de varios modos diferentes incluyendo, pero sin limitarse a, adsorción a una fase sólida, reparto en una fase líquida o condensación directa (tal como condensación debida a exposición a un serpentín de condensación o a un aumento en la presión). En algunos aspectos, enriquecimiento del vapor con membrana de una corriente de vapor de isopreno diluida por encima del punto de rocío del vapor dando como resultado la condensación de isopreno líquido. En algunos aspectos, el isopreno se comprime y se condensa.

La recuperación de isopreno puede implicar una etapa o múltiples etapas. En algunos aspectos, la retirada de vapor de isopreno del gas residual de fermentación y la conversión de isopreno en una fase líquida se realizan simultáneamente. Por ejemplo, el isopreno puede condensarse directamente a partir de la corriente de gas residual para formar un líquido. En algunos aspectos, la retirada del vapor de isopreno del gas residual de fermentación y la conversión de isopreno en una fase líquida se realizan secuencialmente. Por ejemplo, el isopreno puede adsorberse a una fase sólida y luego extraerse de la fase sólida con un disolvente.

En algunos aspectos, cualquiera de los métodos descritos en el presente documento incluye además una etapa de recuperación de los compuestos producidos. En algunos aspectos, cualquiera de los métodos descritos en el presente documento incluye además una etapa de recuperación del isopreno. En algunos aspectos, el isopreno se recupera mediante separación por absorción (véase, por ejemplo, la publicación estadounidense n.º 2011/0178261).

Las composiciones de isopreno recuperadas a partir de fermentaciones en organismos anaerobios pueden contener impurezas. Las identidades y los niveles de impurezas en una composición de isopreno pueden analizarse mediante métodos convencionales, tales como CG/EM, CG/FID y <sup>1</sup>H-RMN. Una impureza puede ser de origen microbiano, o puede ser un contaminante en la alimentación de gas de síntesis u otras materias primas de fermentación.

En algunos aspectos, la composición de isopreno recuperada de la fermentación en un organismo anaerobio comprende una o más de las siguientes impurezas: sulfuro de hidrógeno, sulfuro de carbonilo, disulfuro de carbono, etanol, acetona, metanol, acetaldehído, metacroleína, metil vinil cetona, 2-metil-2-viniloxirano, *cis*- y *trans*-3-metil-1,3-pentadieno, un alcohol prenilico C5 (tal como 3-metil-3-buten-1-ol o 3-metil-2-buten-1-ol), 2-heptanona, 6-metil-5-hepten-2-ona, 2,4,5-trimetilpiridina, 2,3,5-trimetilpirazina, citranelal, metanotiol, etanotiol, acetato de metilo, 1-propanol, diacetilo, 2-butanona, 2-metil-3-buten-2-ol, acetato de etilo, 2-metil-1-propanol, 3-metil-1-butanal, 3-metil-2-butanona, 1-butanol, 2-pentanona, 3-metil-1-butanol, isobutirato de etilo, 3-metil-2-butenal, acetato de butilo, acetato de 3-metilbutilo, acetato de 3-metil-3-buten-1-ilo, acetato de 3-metil-2-buten-1-ilo, (E)-3,7-dimetil-1,3,6-octatrieno, (Z)-3,7-dimetil-1,3,6-octatrieno, (E,E)-3,7,11-trimetil-1,3,6,10-dodecatetraeno y (E)-7,11-dimetil-3-metilen-1,6,10-dodecatrieno, 3-hexen-1-ol, acetato de 3-hexen-1-ilo, limoneno, geraniol (*trans*-3,7-dimetil-2,6-octadien-1-ol), citranelol (3,7-dimetil-6-octen-1-ol), (E)-3-metil-1,3-pentadieno, (Z)-3-metil-1,3-pentadieno, tiol(es), mono y disulfuro(s) o gas(es) tal(es) como CS<sub>2</sub> y COS. La composición de isopreno recuperada de la fermentación de sintegas en un organismo anaerobio puede comprender uno o más de los componentes descritos en Rimbault A *et al.* 1986, J of Chromatography, 375:11-25, cuyo contenido se incorpora expresamente en el presente documento como referencia en su totalidad con respecto a diversos componentes en gases de cultivos de *Clostridium*.

En algunos aspectos, cualquiera de los métodos descritos en el presente documento incluye además purificar el isopreno. Por ejemplo, el isopreno producido usando las composiciones y los métodos de la invención puede purificarse usando técnicas convencionales. La purificación se refiere a un procedimiento a través del cual se separa isopreno de uno o más componentes que están presentes cuando se produce el isopreno. En algunos aspectos, el isopreno se obtiene como un líquido sustancialmente puro. Los ejemplos de los métodos de purificación incluyen (i) destilación a partir de una disolución en un extractor líquido y (ii) cromatografía. Tal como se usa en el presente documento, "isopreno purificado" significa isopreno que se ha separado de uno o más componentes que están presentes cuando se produce el isopreno. En algunos aspectos, el isopreno está al menos aproximadamente al 20%, en peso, libre de otros componentes que están presentes cuando se produce el isopreno. En diversos aspectos, el isopreno es al menos o aproximadamente el 25%, el 30%, el 40%, el 50%, el 60%, el 70%, el 75%, el 80%, el 90%, el 95% o el 99%, en peso, puro. La pureza puede someterse a ensayo mediante cualquier método apropiado, por ejemplo, mediante cromatografía en columna, análisis de HPLC o análisis de CG-EM.

En algunos aspectos, al menos una porción de la fase gaseosa que queda tras una o más etapas de recuperación para la retirada de isopreno se recicla introduciendo la fase gaseosa en un sistema de cultivo celular (tal como un fermentador) para la producción de isopreno.

En algunas realizaciones, la recuperación de enzimas industriales puede usar cualquier método conocido por un experto en la técnica y/o cualquiera de los protocolos a modo de ejemplo que se dan a conocer en las publicaciones de solicitud estadounidense n.ºs 2009/0311764, 2009/0275080, 2009/0252828, 2009/0226569, 2007/0259397 y las patentes estadounidenses n.ºs 7.629.451; 7.604.974; 7.541.026; y 7.527.959 y para nutracéuticos (véase, por ejemplo, la patente estadounidense n.º 7.622.290), y para antimicrobianos (véase, por ejemplo, la publicación de

solicitud estadounidense n.º 2009/0275103).

Los siguientes ejemplos se han proporcionado para fines ilustrativos sólo y no pretenden limitar la invención.

## 5 Ejemplos

### Ejemplo 1: métodos y materiales

10 Las cepas bacterianas usadas en los ejemplos descritos en el presente documento se enumeran en la tabla 1 a continuación.

Tabla 1: Cepas bacterianas

Cepa	Descripción	Referencia/fuente
<i>Escherichia coli</i>		
TOP10	mcrA, ΔmcrBC, recA1, StrR	Life Technologies, Carlsbad CA
XL1-Blue	Δ (mcrA) 183 Δ(mcrCB-hsdSMR-mrr) 171 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac (F'proAB lacI <sup>q</sup> ZΔM15 Tn10 (tetR))	Stratagene, La Jolla CA
S17-1	Tp <sup>R</sup> , Sm <sup>R</sup> , recA <sup>-</sup> , thi, pro, hsdR <sup>-</sup> , hsdM <sup>+</sup>	Colección americana de cultivos tipo, cepa 47055
Anaerobios	Número de registro de la ATCC	
<i>Clostridium acetobutylicum</i>	ATCC 824	Colección americana de cultivos tipo
<i>Clostridium ljungdahlii</i>	ATCC 55383	Colección americana de cultivos tipo
<i>Clostridium aceticum</i>	ATCC 35044	Colección americana de cultivos tipo
<i>Clostridium autoethanogenum</i>	DSM 10061	Colección americana de cultivos tipo

15 Todos los plásmidos se construyeron en células de *E. coli* TOP10 y se enumeran en la tabla 2.

Tabla 2: Plásmidos

Identificador de plásmido	Características	Descrito en
pMCS244	ermB, ColE1 RNA II, lacZ alfa, repL (tiene 4 sitios CCWGG que se producen de manera natural)	Figuras 10A-B y figura 11
pDW265	ermB, ColE1 RNA II, lacZ alfa, repL (los 4 sitios CCWGG que se producen de manera natural mutados)	Figura 14 y figura 15
pDW268	Promotor P <sub>BAD</sub> , RBS de pBAD, RYBO02455 metiltransferasa, terminador rrnB, araC	Figura 23 y figuras 24A-C
pBAD33	Promotor P <sub>BAD</sub> , terminadores 5s y rrnB T <sub>1</sub> T <sub>2</sub> , bla truncado, Cm <sup>R</sup> , pACYC184 ori, araC	Guzman <i>et al.</i> , 1995, Journal of Bacteriology, vol. 177, n.º 14: 4121-4130.
pCA1	repB, transposasa de la familia IS605 OrfB,	Figura 6 y figura 7A-B
pMCS203	ori pIM13, catP, TraJ, lacZ alfa, ARN II de ColE1	Figura 8 y figura 9
pDW263	Marcador de resistencia a cloranfenicol repB, ARN II de ColE1, TraJ, transposasa de la familia IS605 OrfB,	Figura 19 y figuras 20A-C
pDW264	repB, marcador de resistencia a cloranfenicol, ARN II de ColE1, TraJ, transposasa de la familia IS605 OrfB	Figura 21 y figuras 22A-C
pDW280	repB, ermB, ARN II de ColE1, TraJ, transposasa de la familia IS605 OrfB	Figura 23 y figuras 24A-C

pMCS537	repB, ermB, ARN II de ColE1, TraJ	Figura 30 y figuras 31A-B
pMCS444	repA, ermB, ARN II de ColE1, TraJ	Figura 49 y figura 50
pMCS445	repH, ermB, ARN II de ColE1, TraJ	Figura 51 y figura 52
pMCS200	repA, ermB, ARN II de ColE1, TraJ	Figura 32 y figuras 33A-B
pMCS201	repH, ermB, ARN II de ColE1, TraJ	Figura 34 y figuras 35A-B
pMCIjs	ORF de metiltransferasa de <i>Clostridium lungdahlii</i>	Figura 53 y figura 54
pMCS94	ori de pIM13, ori de pB322, EmR, gentR, ApR,	Figura 55 y figura 56
pMCS466	ApR, casete Carb de pMCS94, ORF de metiltransferasa de <i>Clostridium lungdahlii</i>	Figura 57 y figura 58

### Ejemplo 2: identificación de endonucleasa en *Clostridium*

5 Para identificar una endonucleasa de restricción activa en *Clostridium acetivum*, se recogieron cultivos durante la noche de bacterias de tipo natural hechas crecer en medio líquido AcM (tabla 3) mediante centrifugación y se resuspendieron en una disolución que contenía lisozima, penicilina G y sacarosa 0,6 M para inducir la formación de protoplastos. Después de varias horas, se sometieron los protoplastos suspendidos a lisis hipotónica mediante centrifugación y resuspensión en tampón que contenía Tris 100 mM pH 7,4, NaCl 50 mM y PMSF 1 mM. Las células lisadas se retiraron mediante centrifugación, y el sobrenadante se usó en todos los experimentos posteriores para examinar e identificar la actividad endonucleasa. Todas las técnicas y métodos usados siguieron las prácticas convencionales de microbiología y biología molecular.

Tabla 3: receta de AcM

Componente	Cantidad en 1 l de AcM
NH <sub>4</sub> Cl	10 ml
KH <sub>2</sub> PO <sub>4</sub>	3,3 ml
K <sub>2</sub> HPO <sub>4</sub>	4,5 ml
MgSO <sub>4</sub> · 7H <sub>2</sub> O	1 ml
Cisteína HCl	10 ml
Disolución mineral de Wolfe	20 ml
Disolución de vitaminas de Wolfe	20 ml
Resazurina (disolución al 0,1%)	1 ml
NaHCO <sub>3</sub>	10 g
Extracto de levadura	2 g
pH	7,4
H <sub>2</sub> O	Hasta 1 l

### Ejemplo 3: identificación de la secuencia de reconocimiento de ADN para la endonucleasa de restricción

15 El plásmido pMCS244, un vector erm<sup>R</sup> usado para transformar *E. coli* para conferir resistencia a eritromicina, se incubó con 1 µl de lisado de *C. acetivum* en tampón NEB 2 en un volumen final de 20 µl durante 30 minutos a 30°C, y se observó el patrón de digestión por restricción mediante electroforesis en gel (E-gel, Life Technologies). Se observó un patrón de restricción diferenciado (figura 12A), y la endonucleasa de *Clostridium acetivum* no identificada se denominó "Cacl" según la nomenclatura convencional para enzimas de restricción. También se mapearon los sitios de escisión de Cacl en relación con los sitios de escisión de las enzimas de restricción HindIII y ApaLI en pMCS244 (figura 12A, carriles 4 y 5, respectivamente). HindIII y ApaLI son endonucleasas de restricción disponibles comercialmente con secuencias de reconocimiento de ADN bien establecidas de AAGCTT y GTGCAC, respectivamente. El mapa de restricción se refinó adicionalmente generando un producto de PCR lineal, utilizando los cebadores M13R y oMCS25 (tabla 4), sometiéndolo a digestión mediante el lisado de *C. acetivum* y determinando la proximidad de cualquier sitio de escisión en relación con HindIII. La figura 12B muestra los patrones de digestión por restricción del producto de PCR. Usando esta información de secuencia, se identificó la secuencia de reconocimiento CCWGG (W = T o A) como el sitio de reconocimiento de la enzima Cacl que está presente en el

## ES 2 728 307 T3

lisado de *C. acetikum*. La secuencia CCTGG (SEQ ID NO: 10) que es proximal a la secuencia de reconocimiento de *HindIII* de AAGCTT se muestra en la figura 13.

5 Por tanto, este ejemplo ilustra la identificación de la secuencia de reconocimiento de ADN, CCWGG (W = T o A), para la endonucleasa de restricción, *Cacl*, presente en lisado de *C. acetikum*.

Tabla 4. Nombres y secuencias de los cebadores.

Cacl M1 Dir	gaaaaccctgacgttaccactta
Cacl M1 Inv	tgggtaacgtcagggtttccca
Cacl M2 Dir	gaaacgctgntatctttagtct
Cacl M2 Inv	acaggactataaagatancagcgt
Cacl M3 Dir	acggttcctgacctttgctggct
Cacl M3 Inv	ggccagcaaaagtcaggaaccgta
Cacl M2 Inv 2	ataaagatancagcgtttcccctngaagctccctcgtgcgct
Cacl M2 Inv 3	ataaagatancagcgtttcccnnggaagctccctcgtgcgct
Cacl M2 Inv 4	ataaagataacagcgtttcccctagaagctccctcgtgcgct
Cacl M2 Inv 5	ataaagataacagcgtttcccnntggaagctccctcgtgcgctcctcgt
Cacl M2 Dir 2	gaaacgctgttatctttagtct
M13R	caggaaacagctatgacc
oMCS25	ctcattagtagttcagggttaaca
Bad33 2455 frag 1 directo	taccgggggaggaataataaatggccgtactccgcaatattgat
Bad33 2455 frag 2 inverso	ttattattctccccgggtaccgagctcgaattcgcta
Bad33 2455 frag 2 directo	caaagatcgttgaggctgtttggcggatgagagaagat
Bad33 2455 frag 1 inverso	aacagcctcaacgatctttgcgcagcagcagatgtgctcgttct
O105	agggacagctagtcttagagtcggtgaacgctctcc
O106	ccaacttttaaatcaatctaagtatatatgagtaaactggctgcac
O107	gatttaaaaagttggcccagggctccccg
O108	gaactagctgtcccgtatggctgcctctac
oMCS158	cagcacttaacattaaccatataatcacgaac
oMCS159	cagctatagcagctactctttggtattattatcaaaatg
oMCS418	ggtagaccctaattatcgtgaacgc
oMCS419	tgattattattatgaaccgattgtaaatgatttttag
oMCS420	ttggatgagaagataactaaagatgaaggg
oMCS421	ttcagagtataattttctaaatacgtaaatatttttc
oMCS422	atgaacaaaaataaaaatttctcaaaacttttaac
oMCS423	ttattcctcccgttaaataatagataactatta
oMCS426	ctataaatattagcgttggactttttctcccttaaatc
oMCS427	tccaacgctaataattatagatcagtttaaaactgaaactgcaac
GA CA1_1 Plásmido Dir	ccgcgccgcccattatagcataaagaggct
GA CA1_1 Plásmido Inv	agattgacctttattttagaggtatattttct
GA CA1_1 203 Dir	tgaataataaaggtcaatctatgaaatcgca
GA CA1_1 203 Inv	tgctataatggcggcccggtcatagctgtt
GA CA1_2 Plásmido Dir	ccgcgccgcccagctatagcagctactctt
GA CA1_2 Plásmido Inv	agattgacctcagcacttaacattaacct

GA CA1_2 203 Dir	ttaagtgtgagggtcaatctatgaaatgcga
GA CA1_2 203 Inv	gctatagctggcgccgcggtcatagctgtt

Ejemplo 4: identificación de la metiltransferasa de *Clostridium acetikum* (M.Cacl, RYBO02455) y caracterización de su actividad

5 El marco de lectura abierto de *Clostridium acetikum* RYBO02455 (SEQ ID NO: 2) codifica para una enzima con homología a M.Mval, una metiltransferasa de *Micrococcus varians* que transfiere un grupo metilo sobre el resto 4-amino del segundo residuo de citosina de la secuencia de reconocimiento CCWGG (W = T o A) (Butkus *et al.*, 1985, Nucl. Acids Res., vol. 13, n.º 16: 5727-5746). Para determinar si el producto proteico de RYBO02455 metila CCWGG, y por tanto protege esta secuencia de reconocimiento de su escisión por la actividad de Cacl en el lisado de *C. acetikum*, la empresa DNA2.0 optimizó los codones de la secuencia codificante de RYBO02455 para la expresión en *E. coli* y la clonó mediante la clonación sin costuras GeneArt (Life Technologies) en el vector inducible por arabinosa pBAD33 para crear el plásmido pDW268. Los cebadores usados se proporcionan en la tabla 4, y el mapa de plásmido para pDW268, así como su secuencia de ADN, se muestran en la figura 23 y la figura 24A-C (SEQ ID NO. 14), respectivamente.

15 Entonces se cotransformó pDW268 con pMCS244 en células químicamente competentes de *E. coli* Top10 (Life Technologies). Se hicieron crecer las células durante la noche en LB con antibióticos apropiados, se diluyeron de nuevo el día siguiente en medio nuevo en una razón 1:1 y se indujeron con arabinosa (120 µl de una disolución al 15% p/v en 5 ml de LB) durante 3 horas. Entonces se purificaron los plásmidos (Qiagen) y se sometieron a escisión mediante el lisado de *C. acetikum*. La figura 25 muestra que el ADN metilado por RYBO02455 era resistente a la escisión por Cacl, porque pMCS244 pudo volver a transformarse en *E. coli* tras la incubación en lisado de *C. acetikum*. A la inversa, la figura 26 muestra que no era posible transformar células de *E. coli* con pMCS244 no metilado tras la incubación en lisado de *C. acetikum*, debido a la digestión completa por la actividad endonucleasa de Cacl. La enzima codificada por RYBO02455 se denominó "M.Cacl", siguiendo los sistemas de nomenclatura convencionales para metiltransferasas en sistemas de restricción-modificación.

Ejemplo 5: identificación del marco de lectura abierto (ORF) que codifica para la endonucleasa de restricción Cacl

30 RYBO02454 es un ORF que es directamente adyacente a, y transcrito en la dirección opuesta de, RYBO02455 (M.Cacl). RYBO02454 codifica para una enzima con baja identidad de secuencia con M.Mval, una endonucleasa de restricción de *Micrococcus varians* que escinde CCWGG. Debido a su proximidad a M.Cacl, su homología con una enzima que se sabe que escinde la secuencia de reconocimiento CCWGG, y la tendencia de las parejas de enzimas de restricción/metilación a localizarse conjuntamente en genomas bacterianos, se consideró RYBO02454 un candidato a codificar para Cacl, una enzima de restricción en el lisado de *C. acetikum*.

Ejemplo 6: creación de un plásmido resistente a Cacl, pDW265

40 Para determinar si Cacl, que selecciona como diana CCWGG, era la actividad endonucleasa de restricción predominante en el lisado de *C. acetikum*, se ensambló el plásmido pDW265, en el que los 4 sitios de reconocimiento CCWGG identificados se mutaron, usando tanto el kit de clonación sin costuras GeneArt (Life Technologies) como la mutagénesis por PCR QuikChange (Stratagene) según los protocolos recomendados por el fabricante (véase la tabla 4 para los cebadores). El mapa de plásmido para pDW265 se proporciona en la figura 14, y la secuencia de ADN para pDW265 se proporciona en la figura 15A (SEQ ID NO. 11).

45 La figura 16 muestra los resultados de un ensayo de endonucleasas de restricción usando un plásmido de control, pMCS244, o el plásmido pDW265 (que tenía los cuatros sitios de reconocimiento de ADN de Cacl CCWGG mutados) tratado con lisado de *Clostridium acetikum*, la endonucleasa de restricción HindIII o ambos. Carril 1: marcador de peso molecular de ADN X de Roche; carril 2: plásmido de control pMCS244; carril 3: plásmido pDW265 no tratado; carril 4: control de pMCS244 tratado con lisado de *C. acetikum*; carril 5: plásmido pDW265 tratado con lisado de *C. acetikum*; carril 6: pMCS244 tratado con HindIII; carril 7: pDW265 tratado con HindIII; carril 8: plásmido pMCS244 tratado con tanto lisado de *C. acetikum* como HindIII; carril 9: pDW265 con tanto lisado de *C. acetikum* como HindIII.

55 Los carriles 5, 7 y 9 de la figura 16 muestran que pDW265 resiste la escisión cuando se incuba con lisado de *Clostridium acetikum* (carril 5), cuando se incuba con HindIII (carril 7) o cuando se incuba con ambos (carril 9). A la inversa, la figura 16 también muestra que el plásmido pMCS244, que es idéntico a pDW265 excepto porque todavía contiene los 4 sitios de reconocimiento CCWGG identificados, no resiste la escisión cuando se incuba con lisado de *Clostridium acetikum* (compárese pMCS244 no tratado en el carril 2 con pMCS244 tratado con *C. acetikum* en el carril 4), HindIII (carril 6) o ambos (carril 8).

60 Entonces se incubaron pDW265 y pMCS244 con el lisado de *C. acetikum* tal como se describió anteriormente y se transformaron en células de *E. coli* químicamente competentes Top10 (Life Technologies) según el protocolo recomendado por el fabricante. El día siguiente, la presencia de colonias resistentes a eritromicina transformadas

con pDW265 (figura 17), y la ausencia completa de colonias resistentes transformadas con pMCS244 (figura 18), confirmó que pDW265 estaba protegido de la escisión por el lisado de *C. acetivum* que contiene *Cacl*, que reconoce específicamente CCWGG.

#### 5 Ejemplo 7: creación de un plásmido lanzadera de *E. coli* - *C. acetivum* conjugativo, pDW280

Para transformar satisfactoriamente *C. acetivum* con ADN heterólogo, se construyeron en primer lugar vectores lanzadera para su propagación en *E. coli*. La construcción de una serie de vectores lanzadera modulares entre *E. coli* y diversas especies bacterianas clostridiales (conocida como "la serie pMTL80000") se describe en Heap *et al.*, 2009 (Journal of Microbiological Methods, vol. 78: 79-85). Estos vectores pMTL80000 portan uno de cuatro replicones Gram-positivos, un origen de replicación de p15A o ColE1 en *E. coli*, un sitio de clonación múltiple con terminadores de la transcripción flanqueantes y un marcador resistente a antibióticos, *catP*, *ermB*, *aad9* o *tetA*. Algunos de los vectores también portan un promotor de ferredoxina (*Pfdx*) de *C. sporogenes* y un sitio de unión al ribosoma (RBS) o un promotor de tiolasa de *C. acetobutylicum* y RBS para la expresión génica.

Para crear el vector lanzadera pDW280, se amplificó la estructura principal de plásmido de pMCS203 (pMTL85151) mediante PCR (PfuUltra II, Agilent Technologies) usando los pares de cebadores indicados en la tabla 4 (por ejemplo, GA CA1\_1 203 Dir y GA CA1\_1 203 Inv). El mapa de plásmido y la secuencia de ADN para pMCS203 se proporcionan en la figura 8 y la figura 9A-B, respectivamente. Se amplificó el plásmido pCA1 usando los pares de cebadores indicados (por ejemplo, GA CA1\_1 Plásmido Dir y GA CA1\_1 Plásmido Inv, tal como se enumera en la tabla 4). El mapa de plásmido y la secuencia de ADN para pCA1 se proporcionan en la figura 6 y la figura 7A-B, respectivamente. Los productos de PCR del peso molecular apropiado se purificaron mediante electroforesis en gel (Qiagen) y se combinaron usando el kit de clonación sin costuras GeneArt (Life Technologies). Entonces se transformaron estos productos de PCR en células de *E. coli* TOP10 químicamente competentes (Life Technologies) según el protocolo recomendado por el fabricante. Se recuperaron las células y se sembraron en placa sobre medio selectivo, y se seleccionaron transformantes resistentes a cloranfenicol para su análisis adicional. Se hicieron crecer varias colonias individuales durante la noche en medio LB selectivo, y el día siguiente se purificaron los plásmidos (Qiagen) y se compararon los pesos moleculares con el del plásmido pCA1 original mediante electroforesis en gel. Esto dio como resultado el plásmido pDW264. Tal como se indica en el mapa de plásmido pDW264 mostrado en la figura 20, el vector lanzadera pDW264 contiene el plásmido pCA1 de *Clostridium acetivum* nativo y casetes de ADN que permiten la replicación en *E. coli*, la transferencia conjugativa y la resistencia al antibiótico cloranfenicol. La secuencia de ADN para pDW264 se muestra en la figura 22A-C.

A continuación, se cortó pDW264 con las enzimas de restricción *FseI* y *PmeI* (New England Biolabs), siguiendo el protocolo recomendado por el fabricante, para eliminar el casete de resistencia a cloranfenicol. Entonces se ligaron estos vectores (ligasa de T4, NEB) a un casete de resistencia a eritromicina que se había aislado del molde pDW265 mediante digestión por restricción con *FseI*, *PmeI* y *AscI*, y se transformó en células de *E. coli* químicamente competentes Top10 (Life Technologies), usando técnicas de biología molecular convencionales. El plásmido lanzadera conjugativo resultante, pDW280, contenía toda la secuencia nativa de pCA1 de *Clostridium acetivum*, un origen de transferencia, un origen de replicación en *E. coli* y el casete de resistencia a eritromicina. El mapa de plásmido y la secuencia de pDW280 se proporcionan en la figura 27 y la figura 28A-C, respectivamente.

#### 45 Ejemplo 8: creación de los plásmidos lanzadera de *E. coli* - *C. acetivum* conjugativos pMCS537, pMCS244 y pMCS245

El plásmido lanzadera conjugativo pDW280 (mostrado en su forma no modificada en la figura 27) se modificó mediante la eliminación de sus cuatro proteínas hipotéticas así como la eliminación del marco de lectura abierto de transposasa de *repB* para crear el plásmido lanzadera de *E. coli* - *C. acetivum* conjugativo más pequeño pMCS537. En resumen, se amplificó el plásmido pDW280 mediante PCR usando los cebadores oMCS426 y oMCS427 (tabla 4) y luego se purificó y se sometió a autoligación usando el kit de clonación sin costuras GeneArt de Invitrogen.

Se crearon pMCS444 y pMCS445 reemplazando el casete de *catP* por el casete de *emR* en los plásmidos pMCS200 y pMCS201. Se realizó esto usando métodos de digestión-ligación descritos para la colección de plásmidos modulares (clostron.com) o tal como se indica en Heap *et al.*, 2009. El mapa de plásmido para pMCS444 se muestra en la figura 49, y su secuencia de ADN se proporciona en la figura 50, mientras que el mapa de plásmido para pMCS445 se muestra en la figura 51, y su secuencia de ADN se proporciona en la figura 52.

#### 60 Ejemplo 9: creación de los plásmidos lanzadera de *E. coli* - *C. ljungdahlii* conjugativos pMCS200 y pMCS201

Para transformar satisfactoriamente *C. ljungdahlii* con ADN heterólogo, se construyeron en primer lugar vectores lanzadera para su propagación en *E. coli*. La construcción de una serie de vectores lanzadera modulares entre *E. coli* y diversas especies bacterianas clostridiales (conocida como "la serie pMTL80000") se describe en Heap *et al.*, 2009 (Journal of Microbiological Methods, vol. 78: 79-85). Estos vectores pMTL80000 portan uno de cuatro replicones Gram-positivos, un origen de replicación de p15A o ColE1 en *E. coli*, un sitio de clonación múltiple con terminadores transcripcionales flanqueantes y un marcador resistente a antibióticos, *catP*, *ermB*, *aad9* o *tetA*. Algunos de los vectores también portan un promotor de ferredoxina de *C. sporogenes* (*Pfdx*) y un sitio de unión al

ribosoma (RBS) o un promotor de tiolasa de *C. acetobutylicum* y RBS para la expresión génica.

El vector pMTL82151, renombrado pMCS200, porta el origen de replicación Gram-positivo de pCB102, el marcador de resistencia a cloranfenicol *catP* y el origen de replicación de *E. coli* ColE1. El mapa de plásmido para pMCS201/pMTL83151 se proporciona en la figura 32 y la secuencia de ADN se proporciona en la figura 33A-B, y SEQ ID NO: 17.

El vector pMTL83151, renombrado pMCS201, porta el origen de replicación Gram-positivo de pCB102, el marcador de resistencia a cloranfenicol *catP* y el origen de replicación de *E. coli* ColE1. El mapa de plásmido para pMCS201/pMTL83151 se proporciona en la figura 34, y la secuencia de ADN se proporciona en la figura 35A-B y SEQ ID NO: 18.

Ejemplo 10: transformación de *Clostridium aceticum* mediante transferencia conjugativa (pDW268 con pDW280 o con pMCS537)

La transferencia conjugativa implica la transferencia de ADN de una célula bacteriana a otra a través de contacto directo célula a célula. La cepa donadora que se moviliza usada en los ejemplos de la presente solicitud es la cepa de *E. coli* S17-1, que contiene un derivado del plásmido RP4 integrado en su ADN cromosómico y carece de la enzima de restricción de ADN específica de *E. coli* K12, permitiendo por tanto la captación eficaz de ADN clonado foráneo (McFarlane *et al.*, 1987, Journal of Microbiological Methods, vol. 6: 301-305). El sitio *oriT* de RP4 es el origen de la transferencia conjugativa, correspondiente al sitio en el que se mella el dúplex de ADN en la preparación para la transferencia de una única hebra del donador al receptor (William *et al.*, 1990, Journal of General Microbiology, vol. 136: 819-826; Burkhardt *et al.*, 1979, Journal of General Microbiology, vol. 114:341-348). La cepa de *E. coli* S17-1 también contiene una inserción del transposón T7n, lo que da como resultado la resistencia a trimetoprima y bajo nivel de estreptomina de esta cepa.

Para generar una cepa de *E. coli* S17-1 capaz de tanto metilación en como conjugación a partir de *E. coli* en *C. aceticum*, se cotransformaron células de *E. coli* S17-1 (usando técnicas convencionales) con pDW268, un plásmido que codifica para M.Cacli inducible por arabinosa, y o bien el plásmido pDW280 o bien pMCS537. En resumen, se hicieron crecer las cepas S17-1 con tanto el plásmido de metilación pDW268 como cualquiera del plásmido lanzadera pDW280 o pMCS537 durante la noche en medio LB líquido que contenía los antibióticos apropiados, y se diluyó el día siguiente en medio nuevo. Durante la fase exponencial media, a una DO600 de aproximadamente 0,6, se recogieron 5 ml de células mediante centrifugación, se lavaron tres veces en medio LB líquido sin antibióticos y se resuspendieron en 250 µl de LB con 12 µl de una disolución de arabinosa al 15% antes de la conjugación. Simultáneamente, se recogió un cultivo de *C. aceticum* en medio AcM líquido mediante centrifugación y se resuspendió en 100 µl de AcM líquido. Entonces se llevaron las células de *E. coli* a la cámara anaerobia, y se mezclaron suspensiones celulares (100 µl de cada una) y se sembraron en placa junto con una placa de medio sólido AcM. El día siguiente, se rasparon las células de la superficie de la placa de conjugación, y se sembraron en placa sobre placas de AcM nuevo que contenían ácido nalidíxico (10 µg/ml) y eritromicina (5 µg/ml) para seleccionar transformantes positivos. Se realizaron pases sucesivamente de colonias resistentes a eritromicina y ácido nalidíxico para verificar la transformación. La figura 29 muestra células de *C. aceticum* sometidas a múltiples pases que se hacen crecer sobre placas con eritromicina y ácido nalidíxico.

Las cepas de *C. aceticum* transformadas se validaron adicionalmente mediante siembra en estrías sobre LB y pruebas para determinar el crecimiento aerobio (*C. aceticum* no crecerá de manera aerobia), purificación de plásmidos (Qiagen) a partir de la cepa de *C. aceticum* transformada, retransformación en células químicamente competentes de *E. coli* Top10, purificación de plásmidos a partir de la *E. coli* retransformada y confirmación mediante secuenciación completa (Quintara BioSciences). Para su confirmación adicional, los productos de PCR amplificados a partir de plásmidos pDW280 aislados a partir de una cepa de *C. aceticum* transformada, usando los cebadores oMCS418 a oMCS423 (enumerados en la tabla 4), confirmaron la presencia de toda la secuencia heteróloga, el origen de replicación de *C. aceticum* y el casete de resistencia a eritromicina, respectivamente (figura 18).

Ejemplo 11: transformación de *Clostridium aceticum* mediante transferencia conjugativa (pDW268 con pMCS444 o con pMCS445)

Para generar una cepa de *E. coli* S17-1 capaz de tanto metilación en como conjugación a partir de *E. coli* en *C. aceticum*, se cotransformaron células de *E. coli* S17-1 (usando técnicas convencionales) con pDW268, un plásmido que codifica para M.Cacli inducible por arabinosa, y o bien el plásmido pMCS444 o bien pMCS445. Se hicieron crecer las cepas S17-1 con tanto el plásmido de metilación pDW268 como cualquiera del plásmido lanzadera pMCS444 o pMCS445 durante la noche en medio LB líquido que contenía los antibióticos apropiados, y se diluyó el día siguiente en medio nuevo. Durante la fase exponencial media, a una DO600 de aproximadamente 0,6, se recogieron 5 ml de células mediante centrifugación, se lavaron tres veces en medio LB líquido sin antibióticos y se resuspendieron en 250 µl de LB con 12 µl de una disolución de arabinosa al 15% antes de la conjugación. Simultáneamente, se recogió un cultivo de *C. aceticum* en medio AcM líquido mediante centrifugación y se

resuspendió en 100 µl de AcM líquido. Entonces se llevaron las células de *E. coli* a la cámara anaerobia, y se mezclaron suspensiones celulares (100 µl de cada una) y se sembraron en placa junto con una placa de medio sólido AcM. El día siguiente, se rasparon las células de la superficie de la placa de conjugación, y se sembraron en placa sobre placas de AcM nuevo que contenían ácido nalidíxico (10 µg/ml) y eritromicina (5 µg/ml) para seleccionar transformantes positivos. Se realizaron pases sucesivamente de colonias resistentes a eritromicina y ácido nalidíxico para verificar la transformación.

Las cepas de *C. acetikum* transformadas se validaron adicionalmente mediante siembra en estrías sobre LB y pruebas para determinar el crecimiento aerobio (*C. acetikum* no crecerá de manera aerobia), purificación de plásmidos (Qiagen) a partir de la cepa de *C. acetikum* transformada, retransformación en células químicamente competentes de *E. coli* Top10, purificación de plásmidos a partir de la *E. coli* retransformada y confirmación mediante secuenciación completa (Quintara BioSciences).

Conjuntamente, los ejemplos 10 y 11 demuestran la transformación satisfactoria de *Clostridium acetikum* con cuatro plásmidos (pDW280, pMCS537, pMCS444 y pMCS445) que tienen un total de tres orígenes de replicación distintos en *Clostridium acetikum*.

Ejemplo 12: comparación de métodos de transformación para *Clostridium acetikum*

Se generaron protoplastos de *Clostridium acetikum* y se recuperaron según el método de Allock *et al.*, 1982, "*Clostridium acetobutylicum* protoplast formation and regeneration", Applied Environmental Microbiology, vol. 43, n.º 3: 719-721.

Tal como se indica en la tabla 5, los solicitantes sometieron a prueba múltiples métodos para transformar *Clostridium acetikum*, incluyendo: (1) electroporación de protoplastos (según el método descrito en Romero *et al.* para la transformación de protoplastos de *Bacillus subtilis*); (2) transformación mediada por polietilenglicol (PEG), según el método descrito en Chang y Cohen para la transformación de protoplastos de *Bacillus subtilis*; (3) transformación mediada por liposomas (usando DOTAP), según el método de Metcalf *et al.* para la transformación de *Metanosarcina acetivorans*; y (4) la transferencia conjugativa de plásmidos pDW268 y o bien pDW280 o bien pMCS537 tal como se describe en el ejemplo 9 de la presente solicitud.

Tabla 5 - Resultados de los intentos de transformar *Clostridium acetikum* usando diversos métodos

Método de transformación	Resultado de la prueba	Método adaptado de
Electroporación de protoplastos	Lisis celular	Romer <i>et al.</i> , 2006. "Transformation of undomesticated strains of <i>Bacillus subtilis</i> by protoplast electroporation". Journal of Microbiological Methods, vol. 66: 556-559.
Protoplastos + PEG	Falsos positivos	Chang y Cohen, 1979. "High frequency transformation of <i>Bacillus subtilis</i> protoplasts by plasmid DNA". Molecular Genes and Genetics, vol. 168 (1): 111-115.
Protoplastos + DOTAP	Falsos positivos	Metcalf <i>et al.</i> , 1997. "A genetic system for Archaea of the genus <i>Methanosarcina</i> : liposome-mediated transformation and construction of shuttle vectors". Proceedings of the National Academy of Sciences vol. 94: 2626-2631.
Conjugación a partir de <i>E. coli</i> usando pDW268 y (pDW280 o pMCS537)	Verdadero positivo	Presente solicitud.

Sólo la conjugación a partir de *E. coli* que alberga el plásmido inducible por arabinosa pDW268 y el plásmido lanzadera conjugativo pDW280 (o el plásmido lanzadera conjugativo más pequeño pMCS537 o pMCS444), tal como describieron los solicitantes en los ejemplos 9 y 10 de la presente solicitud, dio como resultado la transformación de *Clostridium acetikum*. No pudieron obtenerse transformantes satisfactorios de *Clostridium acetikum* usando electroporación de protoplastos, transformación de protoplastos mediada por PEG o transformación mediada por liposomas. Adicionalmente, no pudieron obtenerse transformantes satisfactorios de *Clostridium acetikum* usando electroporación de células vegetativas.

Los ejemplos 9 y 10 demuestran la transformación satisfactoria de cuatro plásmidos (pDW280, pMCS537, pMCS444 y pMCS445) en *Clostridium acetikum*, tres de los cuales albergan orígenes de replicación distintos (pDW280 y pMCS537 tienen el origen de replicación de repB, mientras que pMCS444 tiene un origen de replicación de repA y pMCS445 tiene un origen de replicación de repH).

Ejemplo 13: transformación de *Clostridium ljungdahlii* mediante transferencia conjugativa (pMCS466 con pMCS200 o

con pMCS201)

Para generar una cepa de *E. coli* S17-1 capaz de tanto metilación en como conjugación a partir de *E. coli* en *C. ljungdahlii*, se cotransformaron células de *E. coli* S17-1 (usando técnicas convencionales) con pMCS466 y o bien el plásmido pMCS200 o bien pMCS201. El plásmido pMCS466 codifica para la metiltransferasa de *C. ljungdahlii* que protege el ADN de la degradación por el sistema de restricción-modificación de *C. ljungdahlii* endógeno. Para crear el plásmido pMCS466, se amplificó el plásmido pMCljS mediante PCR con los cebadores o107 y o108 (tabla 4). Se amplificó el casete de resistencia a carbenicilina a partir del plásmido pMCS94 con los cebadores o105 y o106 (tabla 4). Se aparearon los dos productos de PCR usando los métodos de clonación sin costuras (Invitrogen) para crear el plásmido pMCS466, un derivado del plásmido pMCljS en donde el marcador de resistencia se ha cambiado de espectinomomicina a carbenicilina.

Se hicieron crecer las cepas S17-1 con tanto el plásmido de metilación pMCS466 como cualquiera del plásmido lanzadera pMCS200 o pMCS201 durante la noche en medio LB líquido que contenía los antibióticos apropiados, y se diluyó el día siguiente en medio nuevo. Durante la fase exponencial media, a una  $DO_{600}$  de aproximadamente 0,6, se recogieron 5 ml de células mediante centrifugación, se lavaron tres veces en medio LB líquido sin antibióticos y se resuspendieron en 250  $\mu$ l de LB con 12  $\mu$ l de una disolución de arabinosa al 15% antes de la conjugación. Simultáneamente, se recogió un cultivo de *C. ljungdahlii* en medio MES-F líquido mediante centrifugación y se resuspendió en 100  $\mu$ l de MES-F líquido. Entonces se llevaron las células de *E. coli* a la cámara anaerobia, y se mezclaron suspensiones celulares (100  $\mu$ l de cada una) y se sembraron en placa junto con una placa de medio MES-F sólido. El día siguiente, se rasparon las células de la superficie de la placa de conjugación, y se sembraron en placa sobre placas de MES-F nuevo que contenían ácido nalidíxico (10  $\mu$ g/ml) y el antibiótico apropiado para seleccionar transformantes positivos. Se realizaron pases sucesivamente de colonias resistentes a antibiótico y ácido nalidíxico para verificar la transformación.

Se validaron adicionalmente cepas de *C. ljungdahlii* transformadas mediante purificación de plásmidos (Qiagen) a partir de la cepa de *C. ljungdahlii* transformada, retransformación en células químicamente competentes de *E. coli* Top10, purificación de plásmidos a partir de la *E. coli* retransformada y confirmación mediante electroforesis en gel.

Este ejemplo demuestra la transformación satisfactoria de *Clostridium ljungdahlii* con dos plásmidos que albergan orígenes de replicación distintos: (1) pMCS200, con un origen de replicación de repA (también denominado pBP1), y (2) pMCS201, con un origen de replicación de repH (también denominado pCB102).

Ejemplo 14: determinación de las concentraciones de antibiótico inhibitorias mínimas (MIC) para bacterias clostridiales

La concentración inhibitoria mínima (MIC) es la concentración más baja de antibiótico que se determina que tiene un efecto inhibitorio sobre el crecimiento de un organismo. Las concentraciones inhibitorias mínimas de tiamfenicol y eritromicina para *Clostridium acetivum* (en medios AcM líquidos) y *Clostridium ljungdahlii* en medios MES-F o MES-X líquidos (tabla 6) se determinaron empíricamente diluyendo en serie los medios específicos de cada cepa con adiciones conocidas con antibiótico. La concentración de partida fue de 30  $\mu$ g/ml. Se añadió un inóculo de volumen 1:20 de un cultivo durante la noche de *Clostridium acetivum* a cada dilución en serie y se permitió que creciera durante la noche. Se midió la  $DO_{600}$  de cada muestra y la MIC que se determinó que era la concentración más baja de antibiótico a la que no se había observado crecimiento durante la noche. Los resultados para *Clostridium acetivum* se muestran en la figura 36A. Los resultados para *Clostridium ljungdahlii* se muestran en la figura 38A.

Para determinar la MIC de o bien tiamfenicol o bien eritromicina en placas de agar solidificado, se hicieron diluciones en serie del antibiótico en medios de agar fundido a partir de una concentración de partida de 30  $\mu$ g/ml. Se vertieron los medios en placas de Petri y se dejaron solidificar, luego se transfirieron a la cámara anaerobia y se permitió que se equilibraran durante 48 horas. Se extendió una muestra de 10  $\mu$ l de un cultivo durante la noche sobre cada placa de agar y se dejó que creciera durante 48 horas. La MIC era la concentración más baja de antibiótico a la que no se observó crecimiento. Los resultados para *Clostridium acetivum* (hecha crecer sobre medios AcM) se muestran en la figura 36B y los resultados para *Clostridium ljungdahlii* hecha crecer sobre medios MES-F (descritos en las tablas 6 y 7) se muestran en la figura 38B.

Tabla 6: receta de MES-Fructosa (MES-F) o MES-Xilosa (MES-X)

Componente de los medios	P.f.	Dis. madre, g/l	Molaridad de dis. madre (M)	Vol. de dis. madre/litro	1x MES F final (mM)
NH <sub>4</sub> Cl	53,4 g	100	1,87	10 ml	18,7
KH <sub>2</sub> PO <sub>4</sub>	136,09	100	0,73	2 ml	1,46
MgSO <sub>4</sub> .7H <sub>2</sub> O	246,47	100	0,406	2 ml	0,811
KCl	74,55	100	1,34	1 ml	1,34

CaCl <sub>2</sub> ·2H <sub>2</sub> O	147,01	20	0,136	1 ml	0,136
Acetato de sodio	136,08	166	1,22	2,5 ml	3,05
Cisteína HCl	175,6	879 mg			5,01
Disolución de vitaminas de Wolfe				10 ml	
Mezcla de oligoelementos de <i>Ljungdahlii</i>				10 ml	
Resazurina	229,19	1	0,00436	1 ml	4,36
Extracto de levadura				2 g	
MES	195,2	20		20 g	102,45
Fructosa*	180,16	10		10 g	55,5

\* Para crear medios MES-X, sustituir con 10 gramos de xilosa los 10 gramos de fructosa.

Tabla 7: Mezcla de oligoelementos de *Ljungdahlii* para su uso en la receta MES-F

5

Componente	Cantidad
Acido nitrilotriacético	2,0 g
MnSO <sub>4</sub> · H <sub>2</sub> O	1,0 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	20,0 mg
NiCl <sub>2</sub> · 6H <sub>2</sub> O	20,0 mg
Na <sub>2</sub> MoO <sub>4</sub> · 2H <sub>2</sub> O	20,0 mg
Na <sub>2</sub> SeO <sub>4</sub>	20,0 mg
Na <sub>2</sub> WO <sub>4</sub>	20,0 mg
Agua destilada	Llevar hasta 1,0 l

Ejemplo 15: transferencia conjugativa del plásmido lanzadera que contiene *ispS* pMCS537-*IspS* en *C. aceticum*

10 Se modifica un vector lanzadera pMCS537 para incluir una copia truncada, de codones optimizados del gen *ispS* (isopreno sintasa) de *Populus alba* para crear el plásmido lanzadera pMCS537-*IspS*, y se transforma en *Clostridium aceticum* mediante transferencia conjugativa.

15 La cepa de transferencia conjugativa de *E. coli* S17-1 se transforma conjuntamente con pDW268, un plásmido que codifica para M.CaCl inducible por arabinosa, y el plásmido pMCS537-*IspS*, para generar una cepa de *E. coli* S17-1 capaz de tanto metilación en como conjugación a partir de *E. coli* en *C. aceticum*.

20 Se hacen crecer las cepas S17-1 con tanto el plásmido de metilación pDW268 como el plásmido lanzadera pMCS537-*IspS* durante la noche en medio LB líquido que contiene los antibióticos apropiados, y se diluyen el día siguiente en medio nuevo. Durante la fase exponencial media, a una DO600 de aproximadamente 0,6, se recogen mediante centrifugación, se lavan tres veces en medio LB líquido sin antibióticos y se resuspenden en 250 µl de LB con 12 µl de una disolución de arabinosa al 15% antes de la conjugación. Simultáneamente, se recoge un cultivo de *C. aceticum* en medio AcM líquido mediante centrifugación y se resuspende en 100 µl de AcM líquido. Las células de *E. coli* se llevan entonces a una cámara anaerobia, y se mezclan suspensiones celulares (100 µl de cada una) y se siembran en placa juntas en una placa de medio sólido AcM. El día siguiente, se raspan las células de la superficie de la placa de conjugación, y se siembran en placa sobre placas nuevas que contienen ácido nalidixico (10 µg/ml) y el antibiótico apropiado para seleccionar transformantes positivos. Se realizan pases sucesivamente de colonias resistentes al antibiótico apropiado y ácido nalidixico para verificar la transformación. Se validan adicionalmente cepas de *C. aceticum* transformadas mediante siembra en estrías sobre LB y pruebas para determinar el crecimiento aerobio (*C. aceticum* no crecerá de manera aerobia), purificación de plásmidos (Qiagen) a partir de la cepa de *C. aceticum* transformada, retransformación en células químicamente competentes *E. coli* Top10, purificación de plásmidos a partir de la *E. coli* retransformada y confirmación mediante secuenciación completa (Quintara BioSciences). Para su confirmación adicional, los productos de PCR se amplifican a partir de plásmidos aislados de

25

30

una cepa de *C. acetikum* transformada para confirmar la presencia de toda la secuencia heteróloga, el origen de replicación de *C. acetikum*, el gen *ispS* de *Populus alba* y el casete de resistencia a eritromicina, respectivamente.

5 Ejemplo 16: producción de isopreno mediante *Clostridium acetikum* transformada con pMCS537-*IspS* y hecha crecer sobre fructosa

Se hace crecer *Clostridium acetikum* que alberga el plásmido lanzadera pMCS537-*IspS* para la producción de isopreno en medio DSZM 135 complementado con fructosa. Tras el crecimiento, se toman muestras del espacio superior mediante microextracción en fase sólida (SPME) y se usa software conocido en la técnica para extraer el ión m/z 67 que es característico de isopreno. Se usa un patrón de isopreno autenticado para confirmar el espectro y el tiempo de retención, y un pico al tiempo de elución del isopreno esperado (demostrado mediante el patrón de isopreno) demostraría que la *C. acetikum* transformada produce niveles detectables de isopreno cuando se hace crecer sobre fructosa.

15 Ejemplo 17: transformación de *Clostridium ljungdahlii* mediante transferencia conjugativa (pDW268 con pMCS200-A1)

Para mejorar los niveles de producción de etanol de *Clostridium ljungdahlii* de tipo natural, se modifica el vector lanzadera pMCS200 (por ejemplo, usando cualquiera de las técnicas divulgadas en el presente documento) para incluir genes de aldehído deshidrogenasa y alcohol deshidrogenasa heterólogos, creando así el vector lanzadera pMCS200-A1. Los genes heterólogos son de otro organismo clostridial, o de cualquier organismo que se sepa que posee estos dos genes. Para generar una cepa de *E. coli* S17-1 capaz de tanto metilación en como conjugación a partir de *E. coli* en *C. ljungdahlii*, se transforman conjuntamente células de *E. coli* S17-1 con pDW268, un plásmido que codifica para M.CaCl inducible por arabinosa, y pMCS200-A1. En resumen, se hacen crecer las cepas S17-1 con tanto el plásmido de metilación pDW268 como el plásmido lanzadera pMCS200-A1 durante la noche en medio LB líquido que contiene los antibióticos apropiados, y se diluye el día siguiente en medio nuevo. Durante la fase exponencial media, se recogen células mediante centrifugación, se lavan tres veces en medio LB líquido sin antibióticos y se resuspenden en 250 µl de LB con 12 µl de una disolución de arabinosa al 15% antes de la conjugación. Simultáneamente, se recoge un cultivo de *C. ljungdahlii* en medio MES-F líquido (tablas 6 y 7) mediante centrifugación y se resuspende en 100 µl de MES-F líquido. Se llevan las células de *E. coli* a la cámara anaerobia, y se mezclan suspensiones celulares y se siembran en placa junto con una placa de medio MES-F sólido. El día siguiente, se raspan las células de la superficie de la placa de conjugación, y se siembran en placa sobre placas de MES-F nuevo que contenían ácido nalidíxico (10 µg/ml) y el antibiótico apropiado para seleccionar transformantes positivos. Se realizan pases sucesivamente de colonias resistentes al antibiótico apropiado y ácido nalidíxico para verificar la transformación.

Las cepas de *C. ljungdahlii* transformadas se validan adicionalmente mediante purificación de plásmidos (Qiagen) a partir de la cepa de *C. ljungdahlii* transformada, retransformación en células químicamente competentes de *E. coli* Top10, purificación de plásmidos a partir de la *E. coli* retransformada y electroforesis en gel posterior.

40 Ejemplo 18: producción de etanol por *Clostridium ljungdahlii* transformada con pMCS200-A1 y hecha crecer sobre fructosa

Se hace crecer *Clostridium ljungdahlii* que alberga el plásmido lanzadera pMCS200-A1 para la producción de etanol en medios MES-F (tablas 6 y 7). Tras el crecimiento, se analiza una muestra mediante microextracción en fase sólida (SPME) y se usa software conocido en la técnica para extraer el ión m/z característico de etanol. Se usa un patrón de etanol autenticado para confirmar el espectro y tiempo de retención, y un pico al tiempo de elución del etanol esperado (demostrado por el patrón) demuestra que la *C. ljungdahlii* transformada produce niveles detectables de etanol cuando se hace crecer sobre fructosa. Se espera que *Clostridium acetikum* transformada con pMCS200-A1 y hecha crecer sobre fructosa produzca más etanol que *Clostridium acetikum* de tipo natural que no se ha transformado con pMCS200-A1.

## REIVINDICACIONES

1. Polinucleótido aislado que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 1, en el que el polinucleótido codifica para un polipéptido con actividad metiltransferasa.
2. Polinucleótido según la reivindicación 1,
  - (a) en el que el polinucleótido es SEQ ID NO: 2; o
  - (b) en el que el polipéptido codificado metila un polinucleótido en una secuencia que comprende CCWGG, en el que opcionalmente la secuencia que comprende CCWGG se selecciona del grupo que consiste en CCAGG (SEQ ID NO: 9) y/o CCTGG (SEQ ID NO: 10); o
  - (c) en el que el polipéptido codificado metila un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10.
3. Plásmido que comprende el polinucleótido según una cualquiera de las reivindicaciones 1-2, operativamente unido a una o más secuencias de control de manera que el polipéptido codificado puede expresarse en un huésped de expresión.
4. Plásmido según la reivindicación 3,
  - (a) en el que el huésped de expresión es *E. coli*; y/o
  - (b) en el que dicho plásmido comprende además SEQ ID NO: 14.
5. Plásmido según una cualquiera de las reivindicaciones 3-4, en el que dicho plásmido se transforma dentro de una célula de *E. coli* S17-1.
6. Célula huésped recombinante que comprende
  - (a) uno cualquiera de los polinucleótidos según una cualquiera de las reivindicaciones 1-2 como ácido nucleico heterólogo; o
  - (b) uno cualquiera de los plásmidos según una cualquiera de las reivindicaciones 3-4.
7. Polipéptido aislado que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos el 90% con SEQ ID NO: 3, en el que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende CCWGG.
8. Polipéptido según la reivindicación 7,
  - (a) en el que dicho polipéptido es capaz de metilar un polinucleótido en una secuencia que comprende SEQ ID NO: 9 y/o SEQ ID NO: 10; o
  - (b) en el que dicho polipéptido es capaz de metilar un polinucleótido en SEQ ID NO: 9 y/o SEQ ID NO: 10; o
  - (c) en el que dicho polipéptido es SEQ ID NO 3.
9. Polipéptido aislado producido mediante uno cualquiera de los polinucleótidos según una cualquiera de las reivindicaciones 1-2, en el que el polipéptido tiene actividad metiltransferasa.
10. Método de producción de una ADN metiltransferasa, que comprende: (a) cultivar una célula huésped recombinante que comprende el polinucleótido según una cualquiera de las reivindicaciones 1-2 en condiciones adecuadas para la producción de la ADN metiltransferasa codificada, y (b) recuperar la ADN metiltransferasa.
11. Método de producción de un transformante bacteriano de *Clostridium* recombinante, que comprende:
  - a) introducir el polinucleótido según una cualquiera de las reivindicaciones 1-2 o el plásmido según una cualquiera de las reivindicaciones 3-4 que codifican para una ADN metiltransferasa en una célula huésped bacteriana de *Escherichia*,
  - b) cultivar la célula huésped bacteriana de *Escherichia* en condiciones adecuadas para la expresión de la ADN metiltransferasa, y
  - c) transferir un polinucleótido metilado desde la célula huésped bacteriana de *Escherichia* hasta una célula

huésped bacteriana de *Clostridium*, en el que las bacterias transformadas usando este método se seleccionan del grupo que consiste en *Clostridium aceticum*, *Clostridium ljungdahlii*, *Clostridium acetobutylicum* y *Clostridium autoethanogenum*.

- 5 12. Método según la reivindicación 11, en el que el polinucleótido metilado codifica para una enzima isopreno sintasa.
- 10 13. Método según una cualquiera de las reivindicaciones 11-12, en el que el polinucleótido metilado comprende uno o más polinucleótidos heterólogos que codifican para una o más enzimas de la ruta del isopreno expresados en una cantidad suficiente para producir isopreno, y en el que las bacterias clostridiales contienen una o más rutas para la producción de isopreno.
- 15 14. Método según la reivindicación 12, en el que la célula huésped bacteriana de *Clostridium* se ha modificado por ingeniería genética para producir isopreno a partir de sintegas y/o a partir de hidratos de carbono o mezclas de los mismos.

**FIG. 1**

atggccgactccgcaatattgatgagcaactgaccgaggaatttaagaaactgccgatcgactattgggactttgaggg  
 tgaggacacgaaagaactgacgcacggcctgcacaactatccggcgggatgggttatccgatctaccgtaacattatc  
 gacatcgtgaagcgtcacggtgaggtcgaaaccttctggacccgttatgggtagcggtagggcctggggaaggca  
 agctggcgggttcaacaaagtgtacgggtacggatctgaatcctctggcagtgctgctgagcaaggtaagaccaccgtc  
 ttgaaagaggatagcgtggatattcaggacaagctgctgctgagagaatattgagcaggcgttcgtgtccagcaaacag  
 ctgctggataacattgacaattacattgcgggagaagggcctggacgtcagcgccaagacggctggggctctgatgctg  
 catgtcattttgcgagatctggatacctacaacagcggctgaaaaatcccagactttaagaataatgggtattgggtcaa  
 accgctgcttattctggagctgcaactgattaaggatatcattctgcagatcgagaatgaggacttccgtaacttcttctgg  
 ctgcttctctgaaactgccgctacgtgagcaacaccgtaatggtagtcaagctgtccgatcaagaaagaaaaag  
 tggcagattcaatccggacgtaagatcgagtttacaataatctggatcgtaacatcgaaaagattaagactttgacaa  
 acgttgtaacaacgattgcgaagttagcgttgctttgaagatacccgcttctggactcgggtccggacaatagcatcgat  
 ctgatgattaccagcccaccgtacggcgtatgcaaaaactacgggtggcgtacgggtcaatttagccgtccgtcttctggtggt  
 tggatctggaattgatggacatcgaagagctgaatcaagttgacaacaatctgctgggtggaagaagggtggacaaag  
 acttcgagtgtaactgagctcccgtacctggagaagggcattaaagaaatcaaagaaaaggacctggaccgctgca  
 cgtgacgttatagcttctacgaggatttggataaggctatggagtccattacgaaaaagatgctgcataacagctaccag  
 ttctgggtgctggttaaccgtaccgttaagaagtcaaactgctgaccaacgaaatcattagcgaactgggagagaat  
 atgggttgggtgaggtttacgatatcccgctaacatcccgaataaggctatgccgagccgtaattccccgaccaatgaaa  
 ccggcaagacggctcagcaccatgacgaacgagcacatcgtcgtgctgcaagatcgt

**FIG. 2**

atggctgtattgagaaatattgatgaacaattaacagaagaattcaaaaaactaccaatagattattgggatttgaagggtg  
 aagatacaaaaagaattaacgcatggacttcacaattaccctgctgttatggtatatacctataatagaaatataatagatatt  
 gtcaaaaggcatggtgaggtagaaacttttttagatcccttcatgggttctggtacaggactgttagagggaaaattggcag  
 gctttaataaagttatgggacagatttaaaccttttagcggcttattaagtaagggttaaacaactgtattaaaagaagatt  
 ctgtagatattcaagataaattacttagagagaatattgaacaagcattttagcagcaacaactacttgataatattgat  
 aattacattgcagaaaaaggtttagatgtatctgctaaagatggatgggggttcagatgcacatggtattctgagagaataact  
 agatacatataactcagggttaaaaattccagacttcaaaaatatggggactggttaaacacgtgtgatattagagcttc  
 aactattaaggatataatactacaaatagaaaacgaagattttagaaaatttctttagtatgttttagtgaactgcaagat  
 atgtagtaatacaagaaatggtgagtttaaaactattagaattaaaaaggaaaaagtagcagattcaatcctgatgttaa  
 aatcgagttctataagtatttagatagaacatcgaaaaataaaagactttgataaaagatgaataacgactgcgaag  
 ttagtgttcattgaggatactaggattttagatagtgtagcactgacaatagcatagatttaataactagtagccaccatagg  
 tgattctaaaactactgtagcatatggacagttcagtagaccctcttatggtggttagatctagagcttatggacatagaag  
 aattaaatcaagtagataacaacctactaggcggtaagaaagttgacaaggattttgaatgtgaattatcaagtagaactt  
 tagaaaaagcaataaaaagagattaaggaaaaagacctgatagagcaagagatggttatagttctatgaggacttaga  
 taaagcaatggaatcaataactaagaaaatgagacataatagttatcaattctgggtgtgggaacagaaacagtaaaa  
 gaagttaagctattaactaatgaaattttcagaattagggtgaaaagtagcggtttagtgggaagtatatgatatacctagaaa  
 tataccaataaagttatgccaaagcaggaattcaccaactaatgaaacaggaaaaactgtaagtacaatgacaaatga  
 acatatagtagtattaagaaaagatagggaa

**FIG. 3**

MAVLRNIDEQLTEEFKKLPIDYWDFEGEDTKELTHGLHNYPAVMVYPIYRNIIDIVKRHG  
EVETFLDPFMGSGTGLVEGKLAGFNKVYGTDLNPLAVLLSKVKTTVLKEDSVDIQDKLL  
RENIEQAFVSSKQLLDNIDNYIAEKGLDVSAKDGWGSDAHVILREYLDTYNSGLKIPDF  
KNMGYWFKPRVILELQLIKDIILQIENEDFRNFFLVCFSETARYVSNTRNGEFKLFRIKKE  
KVADFNPDVKIEFYKYLDRNIEKIKDFDKRCNNDCEVSVAFEDTRILDSVPDNSIDLMIT  
SPPYGDSKTTVAYGQFSRPSLWWLDLELMDIEELNQVDNLLGGKKVDKDFECESLSS  
RTLEKAIKEIKEKDLDRARDVYSFYEDLDKAMESITKKMRHNSYQFWVVGNRRTVKEVK  
LLTNEIISELGEKYGLVEVYDIPRNIPNKVMPSRNSPTNETGKTVSTMTNEHIVVLRKDR

## FIG. 4A

atgtataccctagagagattaaaaattaggtaagagaaaataaatcaaatgggatatgtagaactcacaggagtggctc  
tactggaataggtaaaactctgaagatttattaggaattgcagagaataatattgctggagcagatctgaccatctggcg  
agttaaatcatgtagaaacgggcaaattagcatggttacattgtttacaaaaagtcctagccctccacgagtaaacact  
gcacttctagaatcctatggctatgttgaccctacaagaggcggacgaaaaatacttcacacaactttaaatgggttaact  
acaatactgtaaacggaacccctatggattcaaagtcgaagttagaggaagtagggtatattactttctaattccctacgc  
aagttaatgcttattgggaaagagaagattacgttatgctttgaaagtaaactccacgtctaataattgtaaagcaaattc  
acgaggctgctggaagaaatgaagaatttcattttgtagaagcctatcatctgaaggctttagttttgaacaattgaagatt  
actagaacaaggaattataaaaatcgacattcgtataggacaatatccagatggacgaacctatgacctggtacagc  
ttttagaattatgaatgacagaatagatgacttattgaaaataaaataagattatta

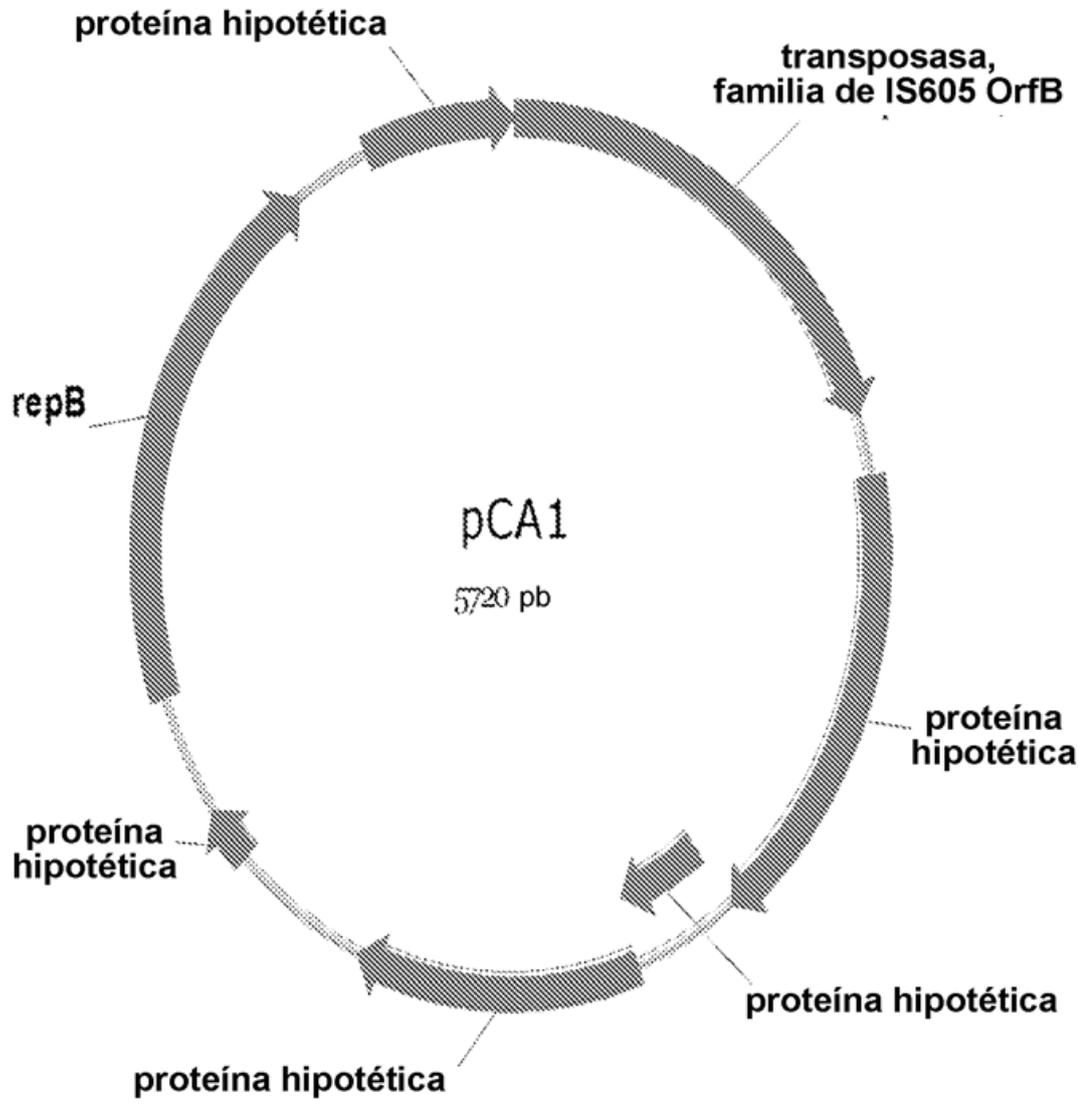




**FIG. 5**

MYTLERLKIRLREINQMGYVRTHRSRGPTGIGKTLEDLLGIAENNIAGADLDHLGELKS  
CRNGQISMVTLFTKSPSPRVNTALLESYGYVDPTRGGRKILHTTLNGVNYNTVNGT  
PYGFKVEVRGSRLYLLSNFPTQVNAYWEREDLRYAFESKLPRLIFVKANSRGAGR  
EEFHFVEAYHLEGFSFEQFEDLLEQGIKIDIRIGQYPDGRTHDHGTAFRIMNDRIDDL  
FENKIRL

**FIG. 6**



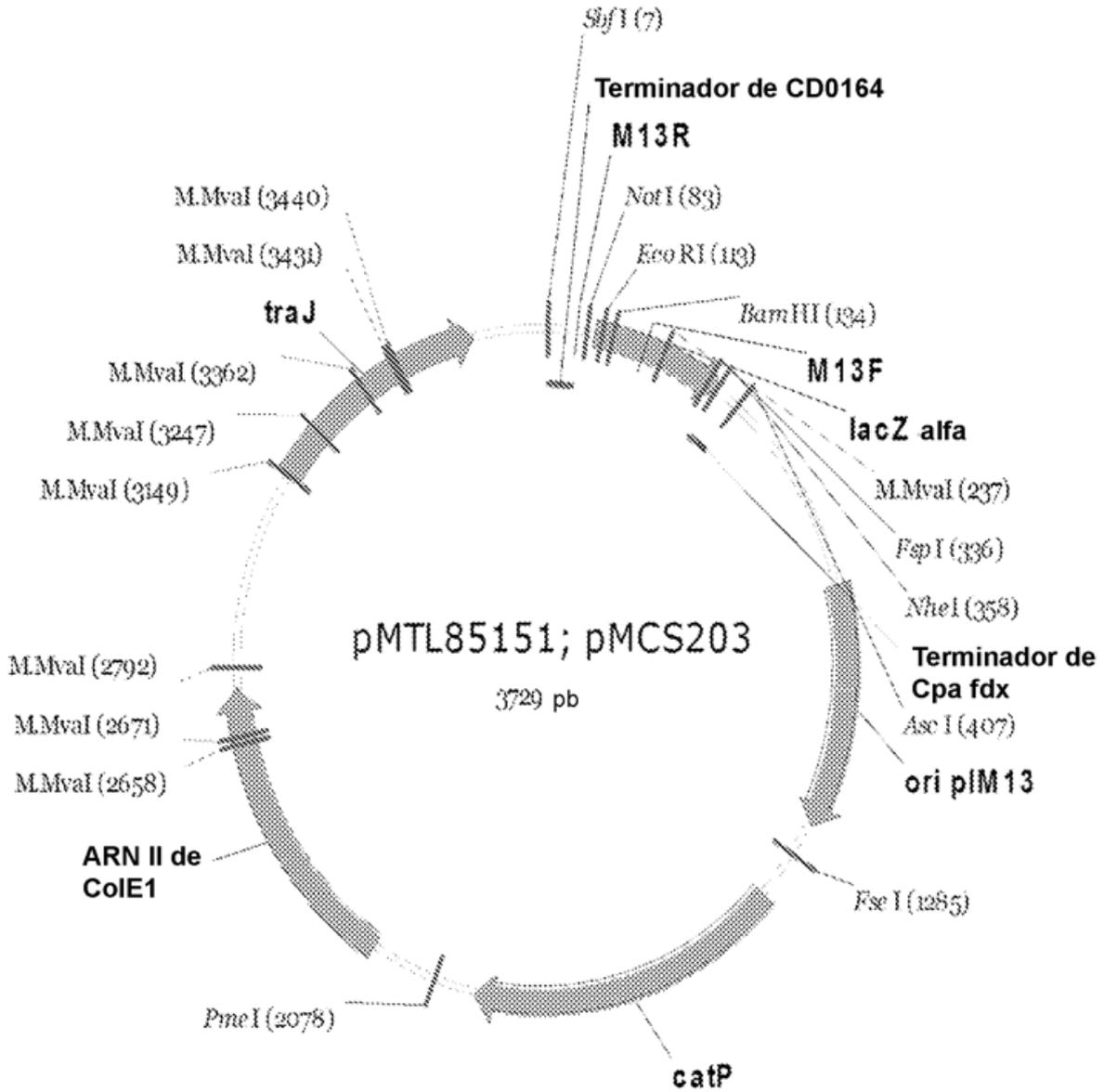
**FIG. 7A**

ttgaagaacaaaaacaagggggtgaacaatgcagataacagtaaaatataatattttgacaaaagaacaagta  
 caactaatagaatctatatcaaaagaatataccatactgttaatagccttgttcatctacgctccaatcagaagaagagt  
 aaagctatcatctaaagatgttttgcaaatatgccaaagtgacagtaaaaaatcaatctattagagatgccaaaagtatctgt  
 actaagtacaagaagctatcaaggctaattccaaactgcctactgataaacaacaaagtaaatcaatgtagctacccttaa  
 aaaacctgtctgtatgtgaataatcaaaattattcactaaagacggatcttctagtttcccgttattatagatgggaaatcg  
 cagcgtattcaaaactagaactatcatgacagactatcagctaaaacaactagaaggcattgggagcattgctgataact  
 aagaaaagcaataaataatcgctcaataagtggtgaaaagtatctcatatagttaaaggatgttgaatgggtgtg  
 acttaggcctaaaagttcctgtctgtagctgaaccgattcaggaaaaacgttttttggaaacggtaggcaaaaataaata  
 cgtaaacgtaataataaagcgaacgtaaaaaacttggaaaagccaagaagctaaagtcattaanaagctgatg  
 ataaagaacaacgttggatgacagaccaagaccacaaagtaagtagagaaataaataatgttcagtaataataatgt  
 ttctgatattcggcttgaanaataacgaatatcagaacacggcaagaacaagccgtaaaaaacgaaaaaaatctaca  
 tacatggctattctatcgtctagctcaattcatagagtataaggcactattgaaggggataaagggtgaatgttgatcctaa  
 atacacttctcaaatatgcctgaatgtaagaaactaaataaagcaagagatagaaaaataaagctcctgtggttttaa  
 aacacatagggatagagtaggtgctataaataaataatgcacctgtagtagatggtaaaagtctactagcctagggtag  
 tataatgactgctctaggaggggtaatggcataccctaagcttgaggctactcctgatagcagaatgtactcggttaat  
 cactcaagaatcccactgctttagctgtgggagtgcaaatgaagcatgatggctatttctgtaactagtgaaaggaagat  
 tgtattatgctgtagtcaaaaaattagtttaatagtggtatacctttaaatacaggagatggagttgttggtaagaaatt  
 caagatttaattcaacttctgatgtttatccgatgttactttaaaggatgaaattgcaaatcaaatatccaaataaattttg  
 aatatgatggaaaagaaccgattagcaatccgttttgggattatgaaaactacatacaggtagtagaagtattgatatag  
 gtgcaaatccagatttatcagctctagtagggaaacatatgaagatgttattagtgaaaatccaagtcaacaaaatccta  
 tgggtgctccgataccatttctgatctggttggcaaatggaaagatatagttaacgatagtggaacatggcaagggg  
 aaggcatagatggaagtactggaactgcaatagatagctcctcattagatattcctggaacgtggcaaggcaaatggtct  
 tggacagcagacgggcaattagtttctgatggttcttttcaggttctgacggaacaacatggcaaggaacatatacgcata  
 caggaatagggtgtcagaatcctgtactaaatccaccactaaccgggatttaacaggaataacaggttgggtatcatctat  
 aagttcatggtaactagtttggcttccaactgatttttagttgaatttagaccggtgaaaaatctacctatagcaacaa  
 aatttcttctgtttgccatttgattaaaaaataagcattgaatcattgcaatctctgtcgttctccagttttacgactactgg  
 aatttaccctttatcaaggagatatagagattaatttagcagctatggaacgatttgcaaaataaacacgttggggaacgtt  
 aattgtatttaactctggttaataactgttacaaggaaggtgtatcatgatatggcaagcactagcatcttttataatctactat  
 taaagcattaggaacggtttaggggcaattatcggattattaccitcaagctcctttcaaaactttcaaaatcagcagtaaca  
 gaatatttaggcatgttgaattggtttatccgtagatgccatgataactatattaacttactggactactgcaattataagtta  
 ctatgtaatatcaactgcatgagatggggaaaaacaattgaatagggggataaatgataagttttatagtggtactcca  
 ggaagtggaaaaagcttaatatagctagatacatatggattaaagttcgacatgctaaacaaaatataaacttgttaata  
 tgacagttaatagagagtattattacatcaaaactgaagcaactgttaataaaattagattgaaattaaaacttaaacct  
 attaataactaagttaaaagactatggcaaaaatctattctataagactcgatcagctgaacacaaaatttctagaagattatg  
 ctatgaaattcacaatgggtggcattgaaggacaatcaaaaataataatagatgaggcacaactgatttggccccaacg  
 gtgatgaaaaataaaaagcaggtagaccctaattatcgtgaacgctggatagagtttatgacactccatagacacttagg  
 tttgacatgataattataagtaatttgatagggtgatagatgcacaaatcgttctatttgaatacaatcatattcatcgga  
 aagtcaataactttgtatagggtattggctaaacctattcaaaataaaagtatttgcagaagtgaatattggtatggagtta  
 gagcaaggattggagtttaatt

**FIG. 7B**

cttcgctattactccatggacttcaaaacactataggaaaattataacgcacataaaagggtctcagattaaagggaa  
 agaaaaaagtagcgttagcgttggactttttctccctttaaatacaagaaatataatgttcgtaaaaaaatgaatcctgatgt  
 catggatcacgtggcagcagtcfaatatttagatctaaaaattgaataatatccaaacaaataggagggtgtgtaaaataa  
 atgttcgtgattataggttaagttaagtgcgcagctatagcagctactcttggattattatcaaaatgcttaataaaatag  
 attcaaaaagtgctatatacatgatagatatatttaatataggggggtgtatagattgtttacaaggaaccagaa  
 actaaaaataagcttttagttcttagaatgacagaaacgcaaaagaagatacttgagattatggctaagagagagggtt  
 atcacaatcagaattaattatgatattattggagaatgaattcaagaagcctgtattagaataaaagcagcaagattaa  
 ctgcccgcctggatagcggagcaacggtttatccaagcggtaaacaaatattctaaacagcgggtttaaaattatcaac  
 tagaagtgtattaatggctgcggaagaaatattaaaccagctactatcaaatcgcaccttaaaagtaagggttttaattg  
 ttaattttggcagcgaactgtctcttcttgatataattacaacaagtcggctaaaattgaaatttaacgttatcctgaaagg  
 ggggcaaaaattggatgagaagatacttaagatgtaagggttctaaaaatcatttacaatcgggtcataataataatca  
 gtataataagttgattgtaggttattacaatcaatacatagaagattctagacctgaaagaagaaaaagactattttggat  
 tatactagatttactatgaagattattttgtgaaaaattagaacataaaagagataagttagctaattgtaataagaaatg  
 ggaagttgaagttatgaaaaacttaaagtaaaagattatgtgctactttattatgtaatgataagttttgtagtaattgaa  
 aaagtaaaagcaagctcaaggatggcgaaaaatagcctttgcttgaacagtataaagataaattatcaaatggtttt  
 aactacaccaaataattgtagatcacaggggaagaattgaaaaaagagattaaaaagcaatttaaagcattaactta  
 ttaacagaatatttaaaaggtaaaaaacaagtaaaagggtttagattttgatattggatacttaggtgcaataaggctgtg  
 gaggtaacttatagcgggtgactattatcatccgcatttgcatttgatattgtagtaattgataatcaaaatgaattataacagat  
 aaaaaaataataaataactattcttatgattattataaaaaaagaccaactagatttttcagattttgaaatattgttacg  
 aaatcttggtatctttatataatggggaaagattgactaaggaaaatagataaactggaaaaagggtatagttgcatg  
 atggataaggcaaaaagaagatgatttttagaagtttttaatacatggtgaagaatgatccggcagaggagaatgtaa  
 aaggtagtaacaaaatgacttataaaaattttagagatttagaataatgcattgcatagataagacagatacaagggtatg  
 gagtttttataatattaaagatatattaatggctgaagaagtaaatgaaatgatgaatggataagagagattttaatcaa  
 aatgaaggagaagctcctgcatatcgtgtgagaagatacagaagctctagatgatactgagataactcttataatcaa  
 ggaaaaaataattacgtatttaagaaaaataactctgaataataacattatagcataaagagggttaattgctctcttt  
 ttaatttctttaaagcttcatttgggtgatgtttaatagattacagtaaaattcgctgaaagcccacgggttcaatcgtgggat  
 gaaaggcgttcttttaactcttctgtgcagtttcagtttaaaactgatactataaatatgggacaagattatagaagaac  
 acaacaacagtatctttaaataaactatcattttgtttctgccaagggtacagacgtaaaagtctagttggagaagttgaaa  
 taaaatttaaacagcttctcaatgagattgtaaaagacattgaaatagaaatttggcaatagaatgtgataaagaccact  
 gccatcttttgcattcacttctcatttaagtccagcagacataatggcaaaagtgaaaggagtgacttctcgattatta  
 aggcaggaatttaacatctgacgacatttgccaagctttggacaagaagctattttgtatctaccgcaggaaatgtatca  
 agtgaactataaaacgatatg

**FIG. 8**



**FIG. 9**

cctgcaggataaaaaaattgtagataaatttataaaaatagttttatctacaatttttatcaggaaacagctatgaccgcccgcctgt  
 atccatgatgacatgattacgaattcgagctcggtacccggggatcctctagagtcgacgtcacgctccatggagatctcaggcctgc  
 agacatgcaagcttgccactggccctgctttacaacgtcgtgactgggaaaacccctggcgttacccaactaatgccttgagcacat  
 cccctttcgcagctggcgtaatagcgaagaggcccgaccgatcgccctccaacagttgcgcagcctgaatggcgaatggcgcta  
 gcataaaaaataagaagcctgcatttgaggcttcttattttatggcgcgccgattcactcttttctataataatgagcgaagcgaat  
 aagcgtcggaaaaagcagcaaaaagttccttttctgttggagcatgggggtcaggggtgagatctgacgtcaatgcccagcga  
 aagcagcgaagggtagcattacgttagataaccccctgatgctccgacgtttatagaaaagaagattcaactaggtaaaat  
 ctaatataggtgagatgataaggttataaggaattgtttgttctaattttcactcattttgttctaatttctttaacaaatgttctttt  
 tttgaaacagttatgatagttagaatgtttaaaaaaggagtgagaaaaagatgaaagaaagatggaaacagctataaaggct  
 ctgagggtcatagacgaagaagtggaagtcagaggttagacaagttataccgtaaaaaacgtctggtaacttcgtaaaaggc  
 atatatagtgcaattaataagtagtttagatgattggcggaaaaaaactaaaatcgtaactatacctagataatgtccacttaagt  
 aacaatacaatgatagctacaacaagagaatagcaaaagctacaggaacaagctacaacagtaataacaacacttaaaatctta  
 gaagaaggaaatattataaaaagaaaaactggagtttaaatgttaaacctgaactactaatgagaggcgcagcaccaaaaacaaaa  
 tacctctactcgaatttgggaactttgagcaagaggcaaatgaaatagattgacctccaataacaccacgtagttattgggaggtcaa  
 tctatgaaatgagattaaggccggccagtgggcaagttgaaaaatcacaaaaatgtgtgataaatctttgttcattagagcgataaa  
 ctgaaattgagagggaacttagatggtattgaaaaaattgataaaaaatggtgaaacagaaaagagattttgaccactactttgcaa  
 gtgtacctgtacctacagcatgaccgttaaagtggatcacacaaaaaaaggaaaagggaatgaaactatacctgcaatgctttatt  
 atattgcaatgattgtaaaccgcatcagagtttaggacggcaatcaatcaagatggtgaattggggatataatgatgagatgatacaa  
 gctatacaatattcacaatgatactgaaacattttccagcctttggactgagtgtaagtctgactttaaatcatttttagcagattatgaa  
 gtgatacgaacggatggaacaaatcatagaatggaaggaaagcacaatgctccgaaaacatttttaagtatctatgatacctgg  
 tcaacctcgatggcttaactgaaattgagaaagatgattatttattctatttttactatggggaaatattataaagaagataa  
 caaaattatactcctttggcaattcaagttcatcacgcagtagtgacggatttcacattgcccgtttgtaaacgaattgaggaattg  
 aaatagtttaactcaggtttgctgtaactaaaaacaagatttaagcaaaaacatcgtagaatacgggtttttgttacctaagttta  
 aactccttttgataatctcatgacaaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccgtagaaaagatcaaaggatc  
 ttctgagatcctttttctgcgcgtaactctgctgctgcaaaaaaaaaccaccgctaccagcgggtggtttgttgcggatcaagagc  
 taccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagttagccgtagtttagccaccactc  
 aagaactctgtagaccgctacatacctcgtctgctaactcctgttaccagtggtctgctccagtgccgataagtcgtgcttaccgggt  
 ggactcaagacgatagttaccggataaggcgcagcggctcggctgaacggggggtcgtgcacacagcccagctggagcgaacgac  
 ctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacgcttcccgaaggagaaaggcggacaggtatccggtaa  
 cggcagggctggaacaggagagcgcagcagggagcttccaggggaaacgcctggtatctttatagtcctgtcgggtttccacctc  
 gactgagcgtcgtttttgtgctcgtcagggggcggagcctatggaaaaacccagcaacgcggccttttacggttctgctgctt  
 ttgctggcctttgctacatgttcttctcgttatccctgattctgtggataaccgtattaccgctttgagtgagctgataaccgctgcc  
 gcagccgaacgaccgagcgcagcagtgagtgagcaggaagcggaaagagcggcaatacgcagggccccctgcttgggggtcatta  
 tagcgatttttctgtatatacctccttttgcacgatatacaggattttccaaagggtcgtgtagacttcttgggtatccaacggc  
 tcagccgggaggataggtgaagtaggcccaccgagcgggtgttcttctcactgtcccttattcgacactggcgggtgctcaacgg  
 gaatcctgctcgcagggctggccggctaccgcccgtaacagatgagggaagcggatggctgatgaaaccaagccaaccaggaa  
 gggcagcccacctaaggtgtactgcttccagacgaacgaagagcgttaggaaaggcggcggcggcggcggcggcggcggcggc  
 ggctactcctgctggcgtcggccagggtacaaaatcacgggctcgtggactatgagcacgtccgcgagctggcccgatcaatggcg  
 acctgggcccctggcggcgtgctgaaactctggctaccgacgaccgcgcagggcgggttgggtgatccacgatctcgcctg  
 ctggcgaagatcgaagagaagcaggacgagctggcaaggtcatgtagggcgtggtccgcccaggggcagagccatgactttttagc  
 cgctaaaacggccggggggtgctgctgattgcaagcagctccccatcgctccatcaagaagagcacttcgggagctggtgaagt  
 acatcaccgacgagcaaggcaagaccgatcgggccc

**FIG. 10A**

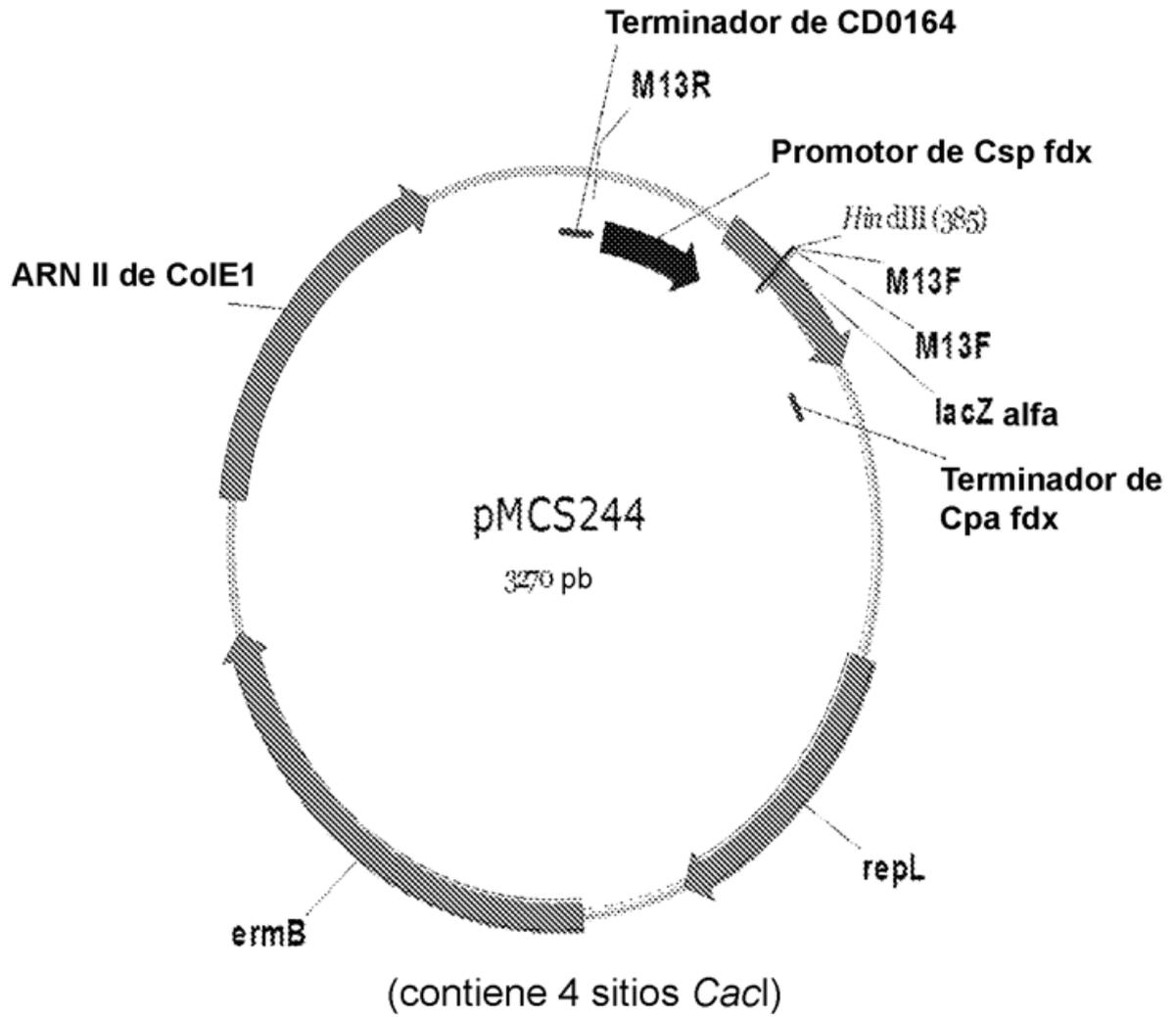
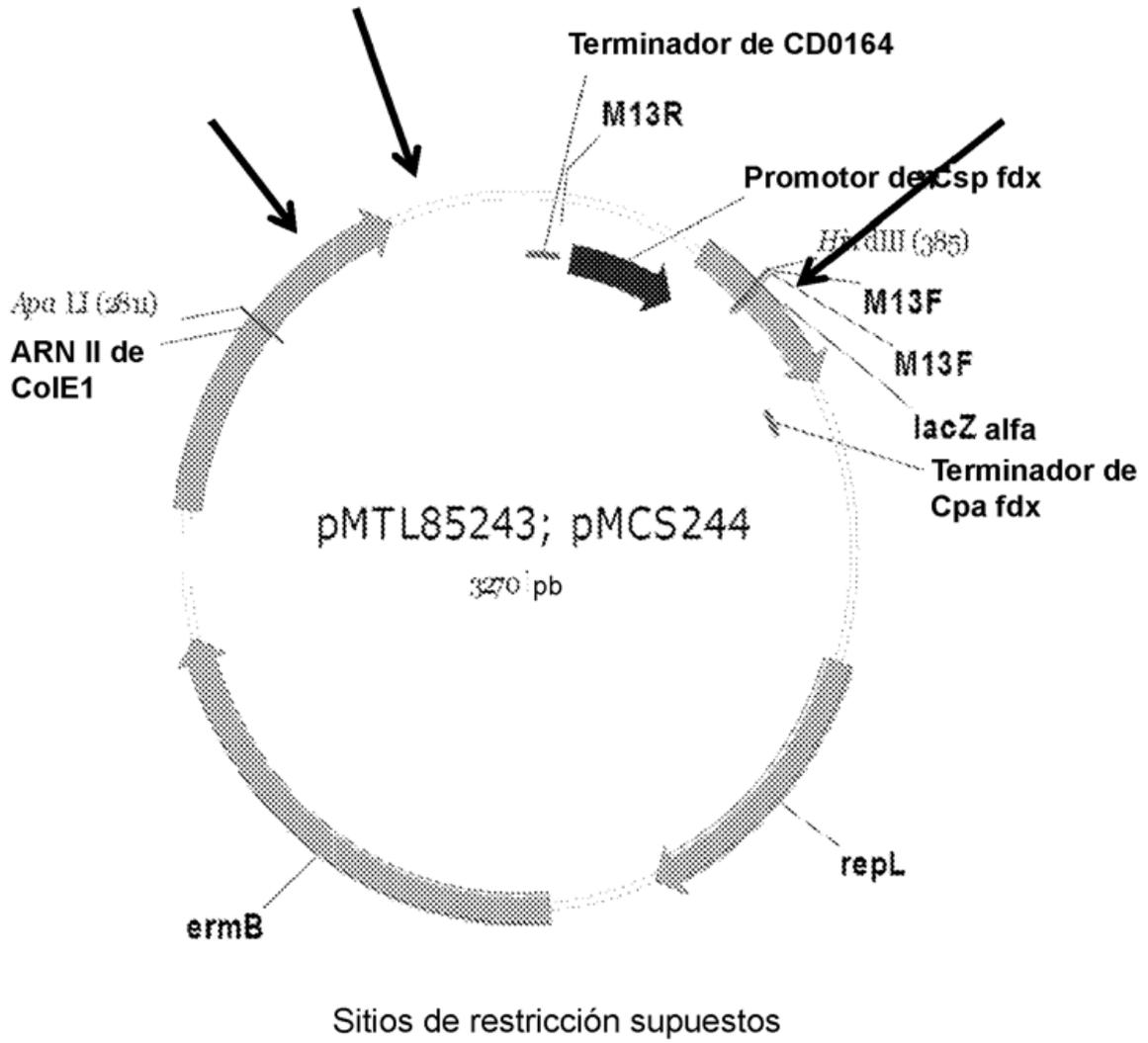


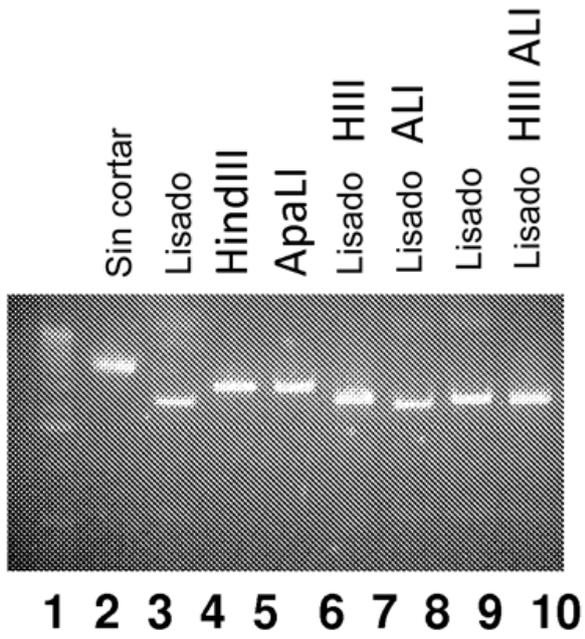
FIG. 10B



**FIG. 11**

cctgcaggataaaaaaattgtagataaaatttataaaaatagttttatctacaatttttatcaggaaacagctatgaccgcg  
ccgctgtagtagcctgtgaaataagtaaggaaaaaaaagaagtaagtggtatataatgatgattttttagatgtagata  
ggataatagaatccatagaaaatataggttatacagttataaaaaaactttaaataaaaaacatggtaaaaat  
aaatcgtataaagtgtagtaattttaaggaggtgttaccatagaccatgattacgaattcgagctcggtagccgggac  
ctctagagtcgacgtcacgctccatggagatctcgaggcctgcagacatgcaagctggcactggccgctgtttaca  
cgtcgtgactgggaaaaccctggcgttaccctaacctaatcgcctgcagcacatcccccttcgccagctggcgtatagc  
gaagaggcccgaccgatcgccctccaacagttgcgcagcctgaatggcgaatggcgtagcataaaaaataagaa  
gcctgcattgtaggcttctatttttatggcgcgccgcallcacttctttctataaaatagagcgaagcgaataagcgtc  
gaaaagcagcaaaaagttccttttctgttggagcatgggggttcaggggtgcagtagctgacgtcaatgccgagcg  
aaagcgagccgaagggtagcattacgttagataacccctgatatgctccgacgctttatagaaaaagaagattcaac  
taggtaaaatctaatataggtgagatgataaggttataaggaattgttttctaaatctcactcatttcttctaa  
caatgttctttttttagaacagttatgatagttagaatagtttaaaaataaggagtgagaaaaagatgaaagaaagat  
atggaacagictataaaggctctcagaggctcatagacgaagaaagtggaagtcataaggttagacaagttatacc  
gtaaacaaacgctcgttaactcgtaaaggcatatagtgcaattaataagtagttagatgattggcggaaaaaac  
ttaaactcgttaactatcctagataatgtccacttaagtaacaatacaatgatagctacaacaagagaaatagcaaaa  
gctacaggaacaagictacaacagtaataacaacacttaaaatcttagaagaaggaaatataaaaaagaaaaaac  
tggagtattaatgttaaaccctgaactactaatgagaggcgacgaccaaaaaacaaaataaccttactcgaattggga  
actttgagcaagaggcaaatgaaatagattgacctccaataacaccacgtagttattgggaggcaatctatgaaatgc  
gattaagggccggccgaagcaaaccttaagagtggtgatagtgacagtagcttaaaatgtgataataggaattgaagta  
aattagatgtaaaaaattgtaattaagaaggagtgattacatgaacaaaaataaaaatatttcaaaaacttttaacgagt  
gaaaaagtaactcaacaaataataaaacaattgaatttaaagaaccgataccgittacgaaattggaacaggtaaa  
gggcatttaacgacgaaactggctaaaataagtaaacaggtaacgtctattgaattagacagctcatctattcaactatcgt  
cagaaaaataaaactgaatactcgtgtcactttaattccaagatattctacagttcaattccctaacaacagaggtat  
aaaattgtgggagtagtcttaccatttaagcacacaaattataaaaaagtggttttgaaagccatgctgtagatctat  
ctgattgtgaagaaggattctacaagcgtacctggatattcaccgaacactaggggtgctctgcacactcaagtctgat  
tcagcaattgcttaagctgccagcggaaatgctttcatcctaaacaaaagtaaacagtgcttaataaaacttaccgcat  
accacagatgttccagataaatattggaagctatatacgtactttgttcaaaatgggtcaatcgagaatatcgtcaactgtt  
actaaaaatcagttcatcaagcaatgaaacacgccaagtaacaatttaagtaccgttactatgagcaagtagttgtct  
attttaatagttatctattttaacgggaggaaataattctatgagtcgctttgtaaattggaaagtacaggtactaaagg  
gaatgtgttaaacctttttgataatctcatgacaaaaatcccttaacgtgagtttctgtccactgagcgtcagacccgta  
gaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaatctgctgcttgcacaaaaaaaccaccgctaccag  
cgggtggtttgttggcggatcaagagctaccaactctttccgaaggtaactggcctcagcagagcgcagataccaaata  
ctgttctctagtagccgtagttaggaccacttcaagaactctgtagcaccgctacatacctcgtctgtaatcctgtt  
accagtggtcgtcagtgaggcagataagtcgtgtcttaccgggtggactcaagacgatagttaccggataaggcgcag  
cggctcgggctgaacgggggttcgtgcacacagcccagctggagcgaacgacctacaccgaactgagatacttaca  
gctgagctatgagaaagcgcacgcttcccgaagggagaaaggcggacaggtatccggttaagcggcagggctcg  
aacaggagagcgcacgagggagcttccagggggaacgccttggtatctttatagtcctgctcgggttcgccacctctga  
cttgagcgtcattttgtgatgctcgtcagggggcggagcctatggaaaaacgccagcaacgcggccttttacgggtc  
ctggccttttgcggcctttgtctacatgttcttctcgttatccccgatctgtggataaccgattaccgctttgagtgagc  
tgataccgctcggcgagccgaacgaccgagcgcagcagtagtagcagaggaagcggaaagagcggccaatac  
gcagggcc

**FIG. 12A**



**FIG. 12B**

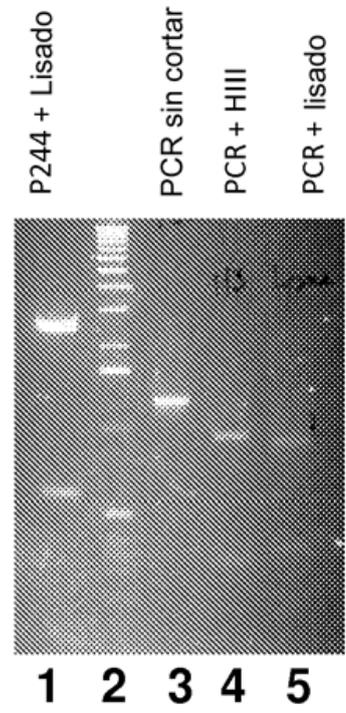


FIG. 13

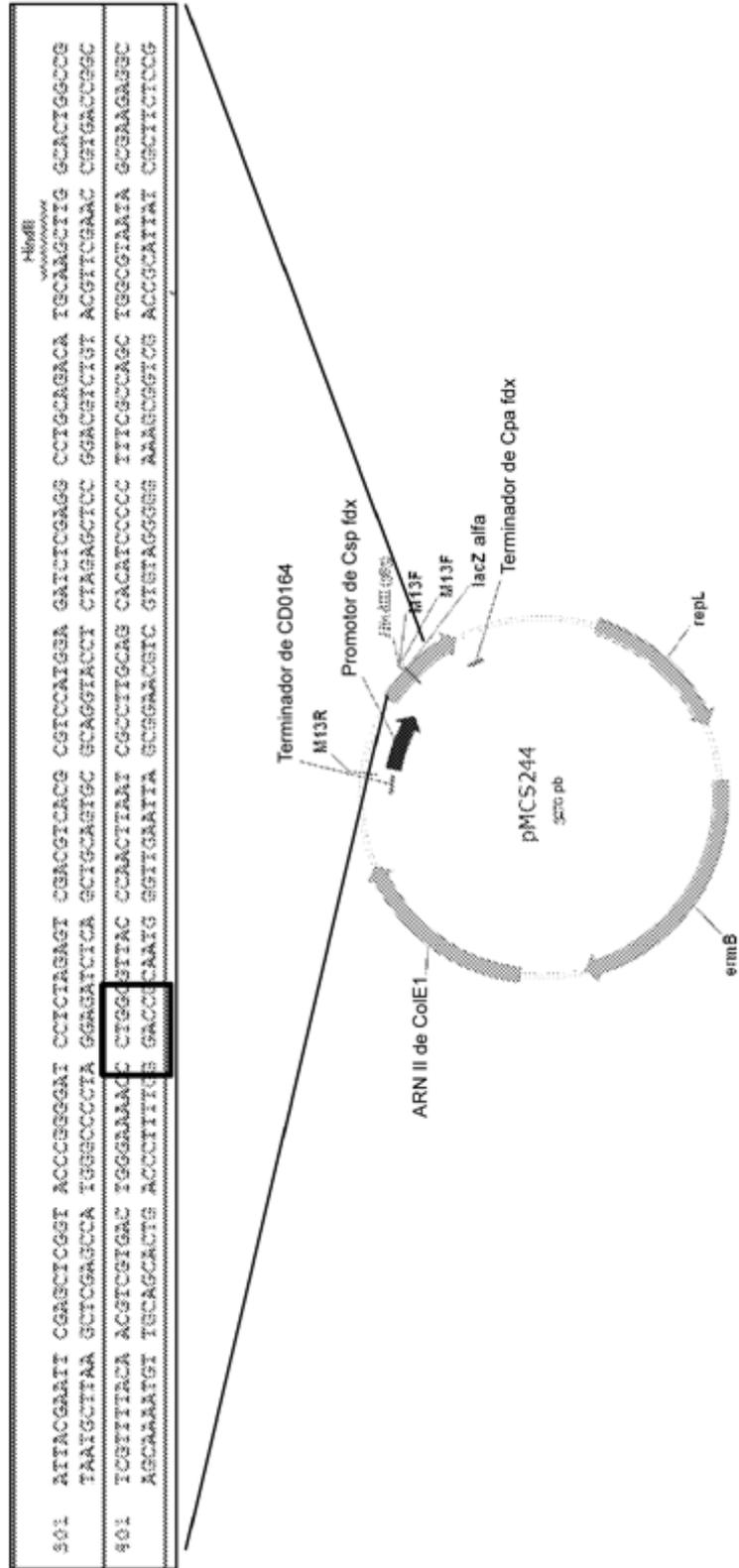
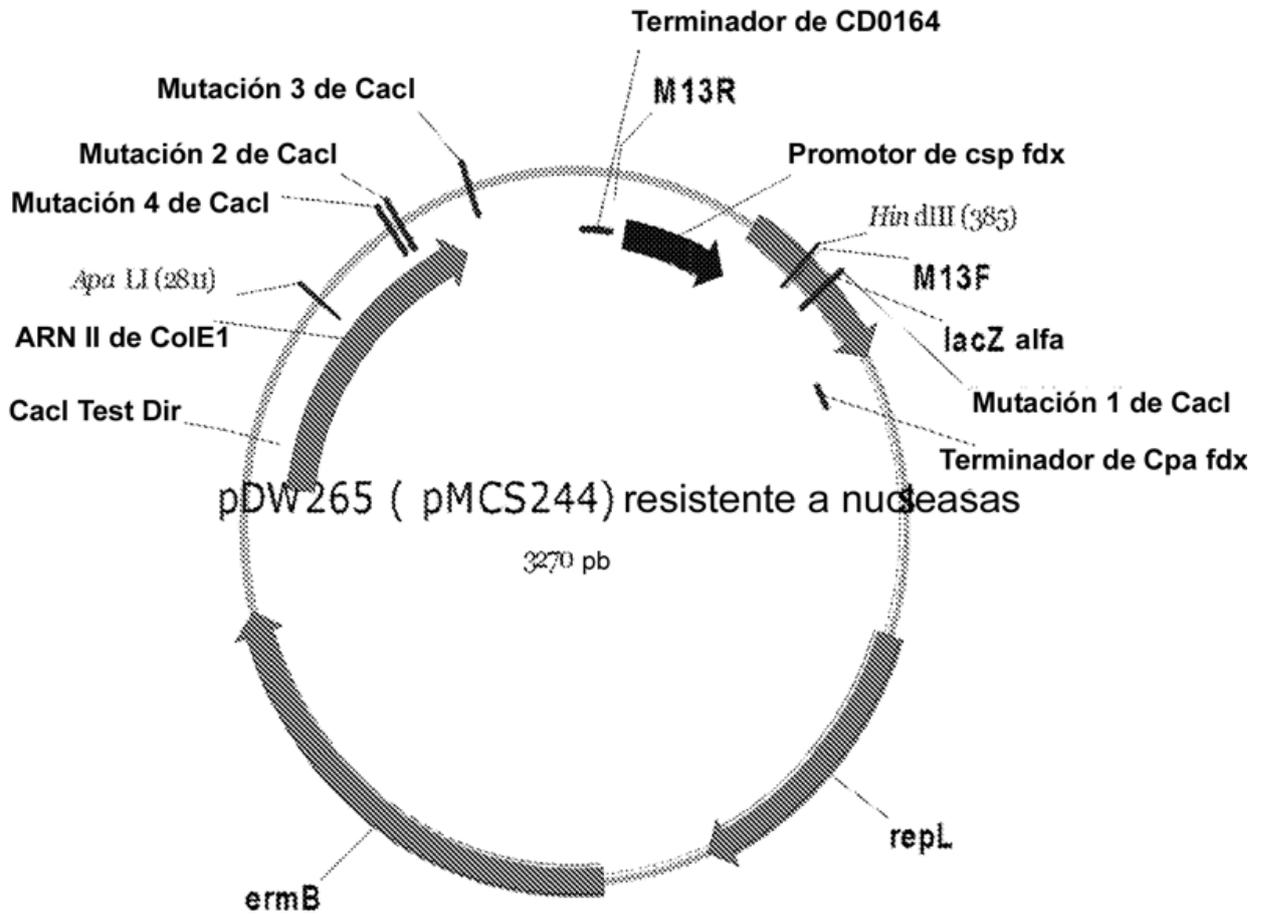


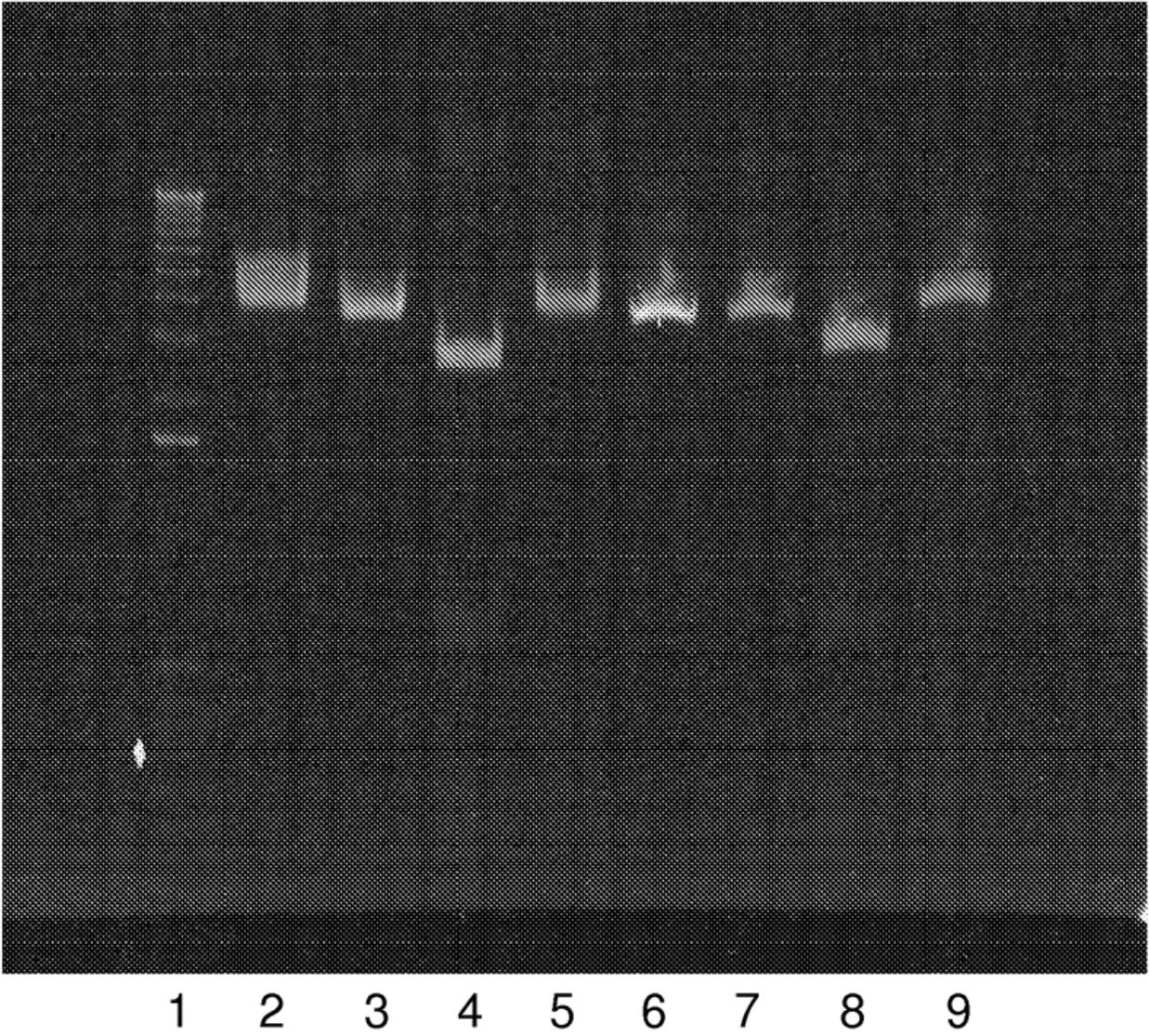
FIG. 14



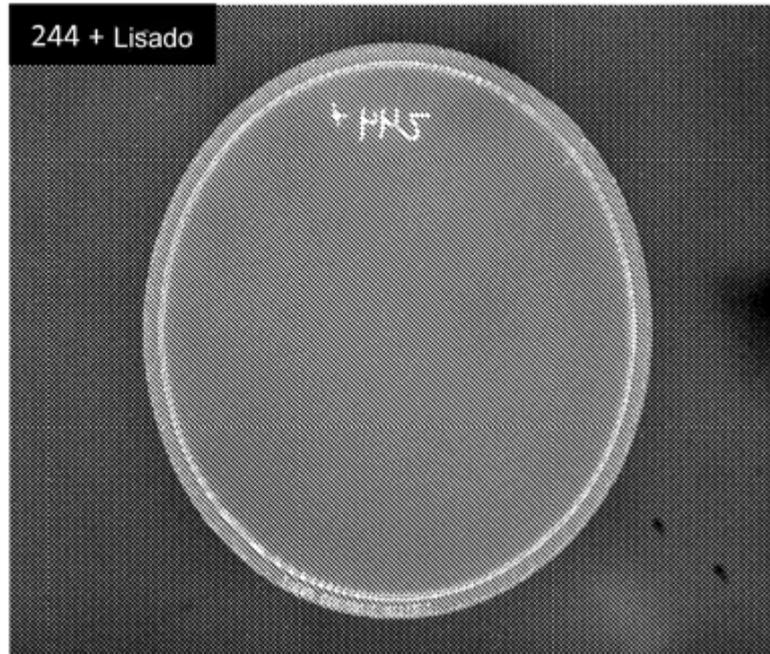
**FIG. 15**

cctgcaggataaaaaaattgtagataaattttataaaaatagttttatctacaatTTTTTatcaggaaacagctatgaccgcgcc  
gcgtgtagtagcctgtgaaataagtaaggaaaaaaaagaagtaagtggtatataatgatgattTTTTgtagatgtagataggat  
aatagaatccatagaaaaataggttatacagttataaaaaattactttaaaaattaataaaaacatggtaaaatataaatcg  
tataaagttgtgtaattttaaggagggtgtttacatatgaccatgattacgaattcgagctcggtagccggggatcctctagagt  
cgacgtcacgcgtccatggagatctcgaggcctgcagacatgcaagcttggcactggccgctgTTTTacaacgtcgtgactg  
ggaaaacctgacgttacccaacttaatcgccttgcagcacatcccccttccgagctggcgtaatagcgaagaggccccg  
caccgatcgccttcccaacagttgcgagcctgaatggcgaatggcgtagcataaaaaataagaagcctgcatttgcagg  
cttctatTTTTatggcgcgccgacttacttctttctataataatagagcgaagcgaataagcgtcggaaaagcagcaaaaa  
gttccTTTTgtggtggagcatgggggtcaggggtgcagtatctgacgtcaatgccgagcgaagcgcgagccgaagggtg  
gcattacgttagataacccccgatatgctccgacgcttataatagaaaagaagattcaactaggtaaaatcttaatataggtt  
gagatgataagggttataaggaattgttcttaattttcactcatttcttaatttcttaacaaatgttctTTTTtagaacagtt  
atgatatagttagaatagtttaaaaaggagtgagaaaaagatgaaagaaagatggaacagctataaaaggctctag  
aggctcatagacgaagaaagtgagaagtcataagaggtagacaagttataccgtaaacaaacgctggtaactcgtaaa  
ggcatatagtgcaattaataagtagttagatatgattggcggaaaaaaacttaaaatcgtaactatacctagataatgct  
cacttaagtaacaatacatgatagctacaacaagagaaatagcaaaagctacaggaacaagctacaaacagtaataa  
caacacttaaaatcttagaagaaggaaatattataaaaagaaaaactggagttaatgttaaacctgaactactaatgag  
aggcgcgacccaaaaacaaaaatacctctactcgaattgggaactttgagcaagaggcaaatgaaatagattgacctcc  
caataacaccacgtagttattgggagggtcaatctatgaaatgcgattaagggccggccgaagcaaacctaaagagtggttga  
tagtgcagtatcttaaaattttagataataggaattgaagttaaattagatgctaaaaattgtaattaagaaggagtgattacatg  
aacaanaataaaaatattctcaaaacttttaacgagtgaaaaagactcaaccaataataaaacaattgaatttaaaag  
aaaccgataccgTTTtacgaaattggaacaggtaaagggtcatttaacgcgaaactggctaaaaataagtaaacaggtaacg  
tctattgaattagacagtcactatctcaactatcgtcagaaaaattaaaactgaatactcgtgctacttaaitcaccagatattc  
tacagttcaattccctaacaacagagggtataaaattgttgggagttacccttaccatttaagcacacaaattattaaaaaagt  
ggttttgaaagccatgcgtctgacatctatctgattgtgaagaaggattctacaagcgtaccttggatattcaccgaacactag  
ggtgctctgacactcaagctctgattcagcaattgcttaagctgccagcggaaatgcttcatctaaacaaaagtaaaaca  
gtgtcttaataaaacttaccgccataccacagatgtccagataaatattggaagctataacgtacttggttcaaaatgggtc  
aatcgagaataatcgtaactgtttactaaaaatcagttcatcaagcaatgaaacacgccaaagtaaaacaatttaagtaccgt  
tacttatgagcaagtagtctatttttaatagttatctattttaacgggaggaaataattctatgagtcgcttttgtaaattggaaa  
gttacagttactaaagggaatgtgtttaaactccttttgataatctatgacccaaaatccctaacgtgagtttctgctcactga  
gcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctTTTTctgcgcgtaactctgctgctgcaacaaaaaaa  
ccaccgctaccagcgggtggttgttggcggatcaagagctaccaactctttccgaaggtaactggctcagcagagcgcga  
gataccaaaactgttctctagtgtagccgtagtaggcccaccactcaagaactctgtagcaccgcctacatacctcgtctg  
ctaactctgttaccagtggtgctgctccagtgccgataaagtcgtgcttaccgggtggactcaagacgatagttaccggataag  
gcgcagcggctgggctgaacgggggggtcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatac  
ctacagcgtgagctatgagaaagcgcacgcttcccgaaggagaaaggcggacaggtatccggtaagcggcaggggtc  
ggaacaggagagcgcacgaggagcttctagggggaaacgcctgatatcttatagtcctgctgggttccgacctctgac  
ttgagcgtcgattttgtgatgctcagggggcggagccatggaaaaacgccagcaacgcggccttttacggttcctga  
cctttgtggtcctttgctcactgttcttctgcgttatccccctgattctgtggataaccgtattaccgcctttgagtgagctgatac  
cgctcggcgcagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcggccaataacgcagggc  
cc

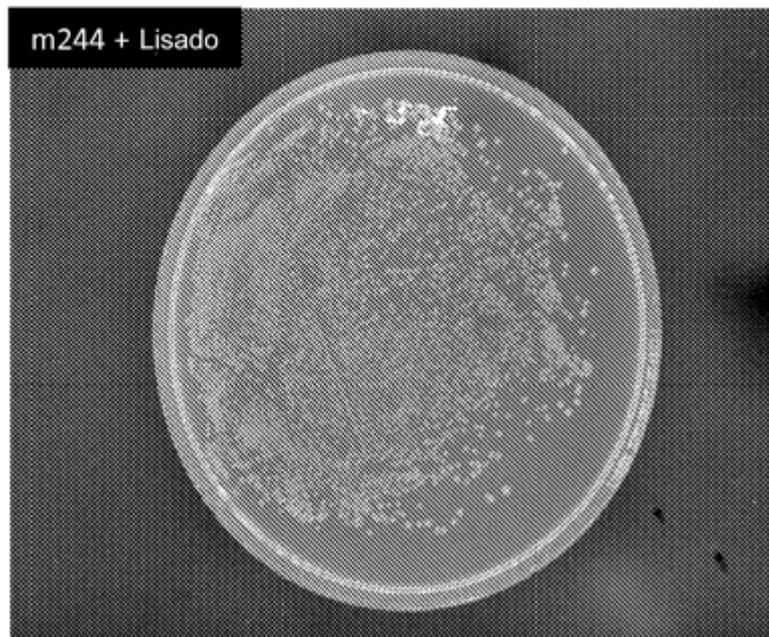
**FIG. 16**



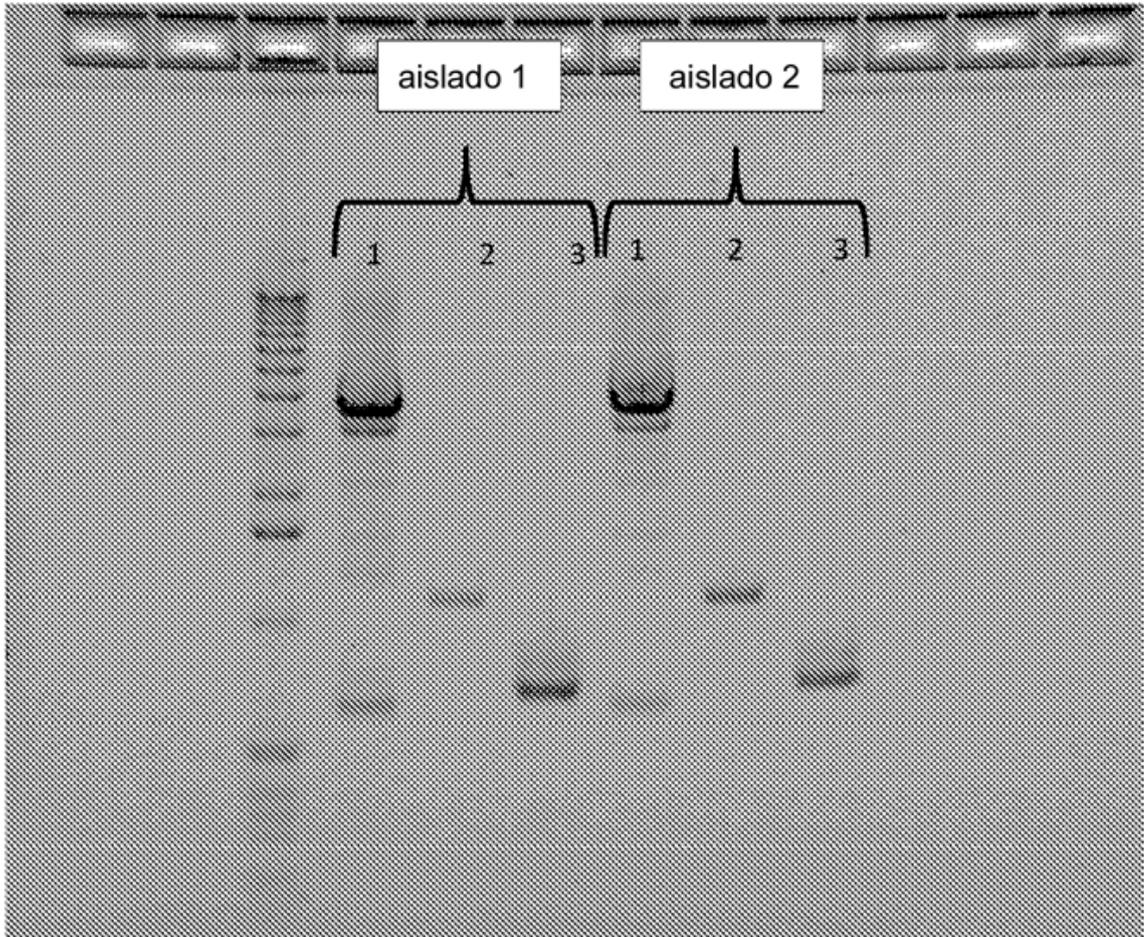
**FIG. 17A**



**FIG. 17B**

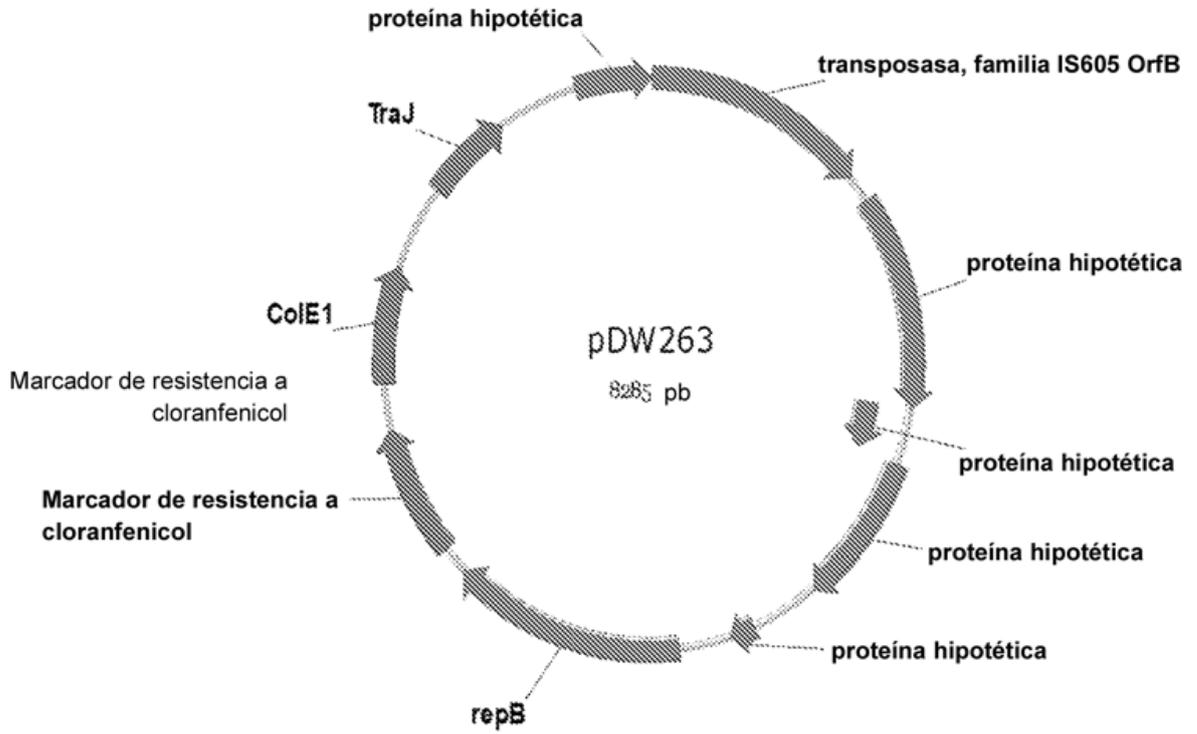


**FIG. 18**



1 = Cebadores 418 y 419  
2 = Cebadores 420 y 421  
3 = Cebadores 422 y 423

**FIG. 19**



**FIG. 20A**

ttgaagaacaaaaacaagggggtgaacaatgcagataacagtaaaatattaatattttgacaaaagaacaagta  
 caactaatagaatctatatcaaaagaatataccatactgtaatagcctgtttcatctacgctccaatcagaagaaagagt  
 aaagctatcatctaaagatgttttgcaaatatgccaaagtcagtgaaaaatcaatctattagagatgccaaaagtatctgt  
 actaagtacaagaaagctatcaaggctaattccaaactgcctactgataaacaanaagtaaatcaatgtagctacccttaa  
 aaaacctgtctgtatatggaataatcaaaattacttaagacgggtattcttagtttcccgttattatagatgggaaatcg  
 cagcgtattcaaactagaactatcatgacagactatcagctaaaacaactagaaggcatttggggagcattgcgtataac  
 taagaaaagcaataaataatcgcctcaataagtggtgaaaaagtatctcatatagttaaagggtgatgtgtaatgggtgttg  
 acttaggcctaaaagttcctgctgtagctgaaccgattcaggaaaaacgtttttttggaacggtagggcaaaaataaata  
 cgtaaacgtaaatataaagcgaacgtaaaaaacttggaagccaagaagcctaaagtcattaaaaagcttgatg  
 ataaagaacaacggtggatgacagaccaagaccacaaagtaagtagagaaataattaatgtcagtaataataatgt  
 ttctgatattcggcttgaaaaattaacgaatcagaaacacggcaagaacaagccgtaaaaaacgaaaaaatctaca  
 tacatggctattctatcgtctagctcaattcatagagtataaggcactattgaaggggataaagggtgaatgttgatcctaa  
 atacacttcaaaatagccctgaatgtaagaaactaaataaagcaagagatagaaaaataaatgctcctgtggttttaa  
 aacacatagggatagagtaggtgctataaataaataatgcacctgtagtagatggtaaaagcttactagcctagggtgta  
 ctatatgtactgctctaggaggggtaatggcataccctaagcttgaggctatactccgatagcagaaatgtactcggtttaa  
 tcaactcaagaatcccactgcttagctgtgggagtgtaaatgaagcatgatggctattatctgtaactagtgaaggaaga  
 ttgtattatgctggtagtcaaaaaattagtttaatagtggtatacctttaaatacaggagatggagttgtgtttggaatgaaatt  
 caagatttaattcaactctgatgtttattccgatgttactttaaacggatgaaattgcaaattcaaatatccaaatataaatttg  
 aatatgatggaaaagaaccgattagcaatccgttttgggattatgaaaacttacatacaggtagtagaagtattgatatag  
 gtgcaaatccagatttatcagctctagtagggaaaacatatgaagatgttattagtgaaaatccaagtcaacaaaaatccta  
 tgggtcctccgataccatttctgattcatggtttggcaaatggaaagatatagttaacgatagtggaacatggcaagggg  
 aaggcatagatggaagtagtggaaactgcaatagatagctccalltagatattcctggaacgtaggcaaggcaaatggct  
 tggacagcagacggcaattagtttctgatggttcttttcagggtctgacggaacaacatggcaaggaacatatacgcata  
 caggaatagggtgtcagaatcctgtactaaatccaccactaaccgggatttaacaggaataacaggttggttatcatctat  
 aagttcatggttaactagttgtttgcttccaactgatttttagttgaaatttagaccggtgaaaaatctacctatagcaaca  
 aatttctttctgtttgcaattgttttaaaaaatagcattgaatcattgcaatctcctgtcgtgtccagttttacgactactgg  
 aatttaccctttatcaaggagatatagagattaatttagcagctatggaacgatttgcaaaaataacagttggggaacgtt  
 aattgtatttaacttggtttaacttgttacaaggaagggttatcatgatatggcaagcactagcatcttttataatctactat  
 taaagcattaggaacgggttttaggggcaattatcggattattacctcaagtcctttcaaaactattcaaaatcagcagtaac  
 agaataatttaggcatgtgaaatgggttataatccgtagatgccatgataactatattaacttactggactactgcaattataagtt  
 actatgtaatatcaactgcatgagatggggaaaaacaattgaatagggggataatgataagttttatagtggtactcc  
 aggaagtggaaaaagcttaatatagctagatacatatggattaagttcgcacatgctaaacaaaataataacttgttaat  
 atgacagttaatagagagtatcttattacatcaaaactgaagcaactgttaataaaaattagattgaaattaaacttaaac  
 ctattaactaagttaaaagactatggcaaaatctattctataagactcgcacagctgaacacaaaatttctagaagattat  
 gctatgaaattcacatggtgggcatggaaggacaatcaaaaataataatagatgaggcacaactgatttggccccaac  
 ggtgatgaaaaataaaaagcaggtagaccctaattatcgtgaacgctggatagagttatgacactccatagacacttag  
 gtttgacatgataaattataagtcatttgataggtgatagatgcacaaatagcttctatttgaatacaatcatattcatcg  
 aaagtcaataactttgtataggttattggctaaacctattcaaaaataaaaagatttgcagaagtgcaaatatgggatggagtt  
 agagcaaggattggagtttaatttctcgtctattactccatggacttcaaaacactataggaaaattataacgcacat

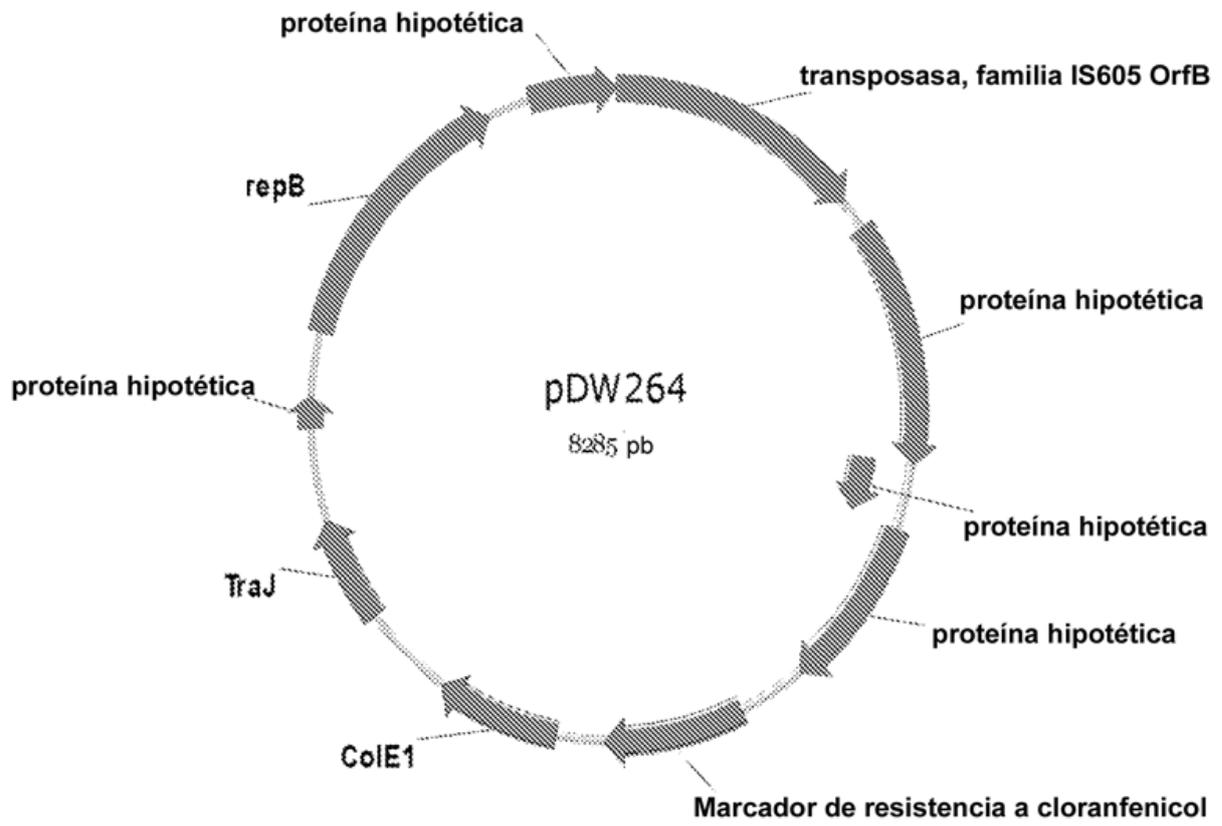
**FIG. 20B**

aaaagggtctcagatttaagggaaagaaaaagtagcgttagcgttgactttttctcccttaaatcaagaaatataat  
 gttcgtaaaaaaatgaatcctgatgcatggatcacgtggcagcagcaatatttagatctaaaaattgaataatccaaa  
 caaataggagggtgtgtaaaataaatgttcgtgattataggtaatgtaagtgctgcagctatagcagctactctttggtatta  
 ttatcaaaatgcttaataaaatagatttcaaaaagtgctatacatgatagatatatttaaatgatatataggggggtgatag  
 attgtttacaaggaaaccagaaactaaaaataagtccttagtcttagaatgacagaaacgcaaaagaagatacttgag  
 attatggctaagtagagaggttatcacaatcagaattaatgatattatggagaatgaattcaagaagcctgtattaga  
 aataaagcagcaagattaaactgccgcctggatagcggagcaacggtttatccaagcggtaaacatattctaaac  
 agcgggtgttaaaattcaactagaagtgattatggcgcggaaagaaatataaaccagctactatcacaattcgcac  
 cttaaagtaagggttttaattgtaattttggcacggaactgctcttcttgatataattacaacaagtcggctaaaattgaaat  
 ttaacgttatcctgaaaggggggcaaaattggatgagaagatactaaagatgtaaggggttctaaaaatcatttacaat  
 cggttcataataataatcagtataataagttgattgtaggttattacaatcaatcatagaagattctagacctgtaaagaag  
 aaaaagactatttggattatactagatttacttatgaagatttttggtaaaaaattagaacataaaagagataagttagct  
 aattgtaataagaaatgggaagtgaaagttatgaaaaactaaagtaaaagattatgtgctactttattatgtaatgataag  
 tttgtagtaattgtaagaaagtaagcaagctcaaggatggcgaaaaatagccttggctgaacagataaaagataaat  
 tatacaaatggtttaactacaccaaataattgtagatcatacaggggaagaattgaaaaagagattaaaaagcaattta  
 aagcattaactatttaacagaatatttaaaaggtaaaaaacaagtaaaaggggttagattttgatattggatacttaggtgca  
 ataaggctgtggaggtaacttatagcggtgactattatcatccgcatttgatattagtagttggataatcaaatgaat  
 ttataacagataaaaaaaatataaataactattcttatgattattataaaaaagaccaactagatttttcagattttgaaat  
 attgttacagaaatctggtaicctttatataatggggaaagattgactaaggaaaatatagataaactggaaaagggttat  
 agttgcatgatggataaggcaaaaagaagatgatttttagaagtttttaatacatgggtaagaatgatccggcagagga  
 gaatgtaaaaggtagtaacaaaatgacttataaaaatttttagagattagaatatgcattgcatagataagacagataca  
 aggtatggagtttttataatattaagatatattaatggctgaagaagtaaatgaaatgatgaatggataagagagatttt  
 aatcaaaaatgaaggagaagctcctgcataatcgtgtgagaagatacagaagcttctagatgatactgagatactcttat  
 atcaaggaaaaaaatattacgtatttaagaaaaatatactctgaataataaaggtaaatctatgaaatgagtaagggc  
 cggccagtgggcaagtgaaaaattcacaaaaatgtggtataatactttgttcattagagcgataaactgaattgagag  
 ggaacttagatggtattgaaaaatgataaaaatagttggaacagaaaagagattttgaccactactttgcaagtgac  
 ctgtacctacagcatgaccgttaaagtgatatacacacaataaaggaaaagggatgaaactatctcgtcaatgctt  
 tattatattgcaatgattgtaaaccgccattcagagtttaggacggcaatcaatcaagatgggtaattggggatatatgatg  
 agatgataccaagctatacaatatttcacaatgatactgaaacattttccagcctttggactgagtgtaagctgactttaa  
 catttttagcagattatgaaagtgatacgcacggatggaaacaatcatagaatggaaggaaagccaaatgctccgga  
 aaacatttttaatgtatctatgataccgtggtcaacctcgtatggcttaactcgaattgcagaaaggatgattatttgattcc  
 tatttttactatggggaaatattataaagaagataacaaaattatacttctttggcaattcaagttcatcacgcagatgtgac  
 ggatttcacatttgccgtttgtaaacgaattgcaggaattgataaatagtttaactcagggttctgtaactaaaacaagta  
 ttaagcaaaaacatcgtagaaatcgggtgtttttgtaccctaagtttaaaactccttttgataatctcatgacaaaatccctt  
 aacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttctgagatcctttttctgctgaa  
 tctgctgctgcaacaaaaaaaccaccgctaccagcgggtggtttgtttgcccggatcaagagctaccaactctttccgaa  
 ggtaactggcttcagcagagcgcagataccaaatactgttctctagtgtagccgtagttaggccaccactcaagaactct  
 gtagcaccgctacatacctcgtctgtaactcctgttaccagtggtcgtgcccagtgggcagataagtcgtgcttaccgggtt  
 ggactcaagacgatagttac

**FIG. 20C**

cggataaggcgcagcggctcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaac  
 tgagatacctacagcgtgagctatgagaaagcggccacgcttcccgaagggagaaaggcggacaggtatccggttaagcg  
 gcagggctcggaacaggagagcgcacgagggagcttccagggggaaacgcctggatctttatagtcctgtcgggttcgcc  
 acctctgactgagcgtcgatttttgatgctcgtcagggggcgaggcctatggaaaaacgccagcaacgcggcctttta  
 cggttctcggccttttgctggcctttgctcacatgttcttctcgttatcccctgattctgtggataaccgtattaccgccttgagt  
 agctgataccgctcggcgcagccgaacgaccgagcgcagcagtcagtgagcaggaagcgggaagagcgcccaata  
 cgagggccccctgctcggggtcattatagcgatttttcgggtatataccatcttttcgcacgatatacaggatttgccaaagg  
 gttcgttagactttccttggtgatccaacggcgtcagccgggcaggataggtgaagtagggccaccgcgagcgggtgtt  
 ccttctcactgtcccttaltcgcacctggcggtgtcaacgggaatctgctcgcgaggctggccgggtaccgccggcgtaa  
 cagatgagggaagcggatggctgatgaaaccaagccaaccaggaagggcagcccacctatcaagggtgactgccttc  
 cagacgaacgaagagcgattgaggaaaaggcggcgccggccggcatgagcctgtcggcctacctgtggccgtcggcc  
 agggctacaaaatcacgggcgtcgtggactatgagcacgtccgcgagctggcccgcataatggcgacctgggcccgcct  
 gggcggcctgctgaaactctggctcaccgacgaccgcgcagggcgcggttcgggtgatgccacgatcctcggcctgctgg  
 cgaagatcgaagagaagcaggacgagctggcaaggtcatgatggcggtggcccccaggggcagagccatgactttt  
 ttagccgctaaaacggccgggggtgctcgtgatgccaagcacgtccccatgctccatcaagaagagcgacttcgcg  
 gagctggtgaagtacatcacgacgagcaaggcaagaccgatcgggccccctgcaggataaaaaaattgtagataaatt  
 ttataaaatagttttatctacaattttttatcaggaaacagctatgaccgcggccgattatagcataaagagggttaattgct  
 ctcttttttaattcttttaagcttcatttgggtgatgttaatagattacagtaaatcgcctgaaagcccacggttcaatcgtggg  
 atgaaaggcgttcttttaactctctgttcagttcagtttaaaactgatactataaataatgaggacaagattatagaagaacac  
 aaacaacagtatcttaataaactatcatttgtttctgtccaaggtacagacgtaaagttctagttggagaagttgaaataaaa  
 ttaaacagcttciaatgagattgtaaagacattgaaatagaaatgtggcaatagaatgtgataaagaccactgccatctttt  
 gtcaatgcacttctcatttaagtcagcagacataatggcaaaagtgaaggagtgacttctcattattaaggcaggaattt  
 aaacatctcgcacatttgccaagcttttgacaagaagctatttgtatctaccgcaggaaatgtatcaagtgaaactataaaa  
 cgatatg

**FIG. 21**



**FIG. 22A**

ttgaagaacaaaaacaagggggtgaacaatgcagataacagtaaaatattaatattttgacaaaagaacaagta  
 caactaatagaatctatatcaaaagaatataccatactgttaatagccttgtttcatctacgctccaatcagaagaagagat  
 aaagctatcatctaaagatgttttgcaaatatgccaagtcagtgaaaaatcaatctattagagatgccaaaagatctgt  
 actaagtacaagaagctatcaaggctaattccaaactgcctactgataaacaagtaaatcaatgtagctacccttaa  
 aaaacctgtctgtatatggaataatcaaaattattcacttaagacgggtattcttagtttcccgttattatagatgggaaatcg  
 cagcgtattcaactagaactatcatgacagactatcagctaaaacaactagaaggctattgggagcattgcgtataac  
 taagaaaagcaataaataatcgcctcaataagtggtgaaaaagtatctcatatagttaaagggtgatgtgtaatgggtgttg  
 actaggcctaaaagttcctgtctgtagctgaaccgattcaggaaaaacgttttttggaaacggtagggcaaaaataaata  
 cgtaaacgtaaatataaagcgaacgtaaaaaacttggaaaagccaagaagcttaaaagtcattaaaagcttgatg  
 ataaagaacaacggttgatgacagaccaagaccacaaagtaagtagagaaataaataatggcagtaataataatgt  
 ttctgatattcggcttgaaaaattaacgaatatcagaacacggcaagaacaagccgtaaaaacgaaaaaaatctaca  
 tacatggctcattctatcgtctagctcaattcatagagtataaggcactattgaaggggataaagggtgaaatgtgtatcctaa  
 atacacttctcaaatatgccctgaatgtaagaaactaaataaagcaagagatagaaaatataaatgctcctgtggttttaa  
 aacacatagggatagagttaggtgctataaataaataatgcacctgtagtagatggtaaaagcttactagcctagggtta  
 ctatatgtactgtcttaggggggtaatggcataccctaagcttgaggctatactccgatagcagaaatgtacttcggtttaa  
 tcaactcaagaatcccactgcttagctgtgggagtgcaaatgaagcatgatggctattatctgtaactagtgaaggaaga  
 ttgtattatgctggtagtcaaaaaattagtttaatagtggtataccttaataacaggagatggaggtgtgttggaaatgaaatt  
 caagattaaattcaactctgtatgtttattccgatgttactttaacggatgaaattgcaaaattcaaatatccaaatataaatttg  
 aatatgatggaaaagaaccgattagcaatccgttttgggattatgaaaacttacatacaggtagtagaagtattgatatag  
 gtgcaaatccagatttatcagctctagtagggaaaacatatgaagatgttattagtgaaaatccaagtcaacaaaaatccta  
 tgggtgcctccgataccatttctgattcatggtttggcaaatggaaagatatagttaacgatagtggaacatggcaagggg  
 aaggcatagatggaagtagtggaaactgcaatagatagctccattagatattcctggaacgtggcaaggcaaatggct  
 tggacagcagacggcaattagtttctgatggttcttttcagggtctgacggaacaacatggcaaggaacatatacgcata  
 caggaataggtgtcagaatcctgtactaaatccaccactaaccocggattaacaggaataacaggttggttatcatctat  
 aagttcatggttaactagttgttgcgtttccaactgattttagtttgaatttagaccggtgaaaaatctacctatagcaacaa  
 aatttcttctgtttgccatttgattaaaaaatagcattgaatcattgcaatctcctgtcgttgcctcagttttacgactactgg  
 aatttaccctttatcaaggagatatagagattaatttagcagctatggaacgatttgcacaaataacacggtggggaacgtt  
 aattgtatttaacttggtttaacttgttacaaggaagggtgtatcatgatatggcaagcactagcatcttttataatctacttat  
 taaagcatttaggaacggttttaggggcaattatcggattattacctcaagtcctttcaaaactattcaaaatcagcagtaac  
 agaataatttaggcattgtgaattggtttatataccgtagatgccatgataactataaacttactggactactgcaattataagtt  
 actatgtaatacaactgcgatgagatggggaaaaacaattgaatagggggataatagataagttttatagtggtactcc  
 aggaagtggaaaagcttaatatagctagatacatatggattaagttcgcacatgctaacaacaaataaataacttgttaat  
 atgacagttaatagagagtatcttattacatcaaaactgaagcaacttgttaataaaatagattgaaattaaacttaaac  
 ctattaactaagttaaaagactatggcaaaatctattctataagactcgtacagctgaacacaaaatttctagaagattat  
 gctatgaaattcacatggtgggcattgaaggacaatcaaaaataataatagatgaggcacaactgatttgggtcccaaac  
 ggtgatgaaaaataaaaagcaggtagaccctaattatcgtgaacgctggatagagtttatgacactccataga

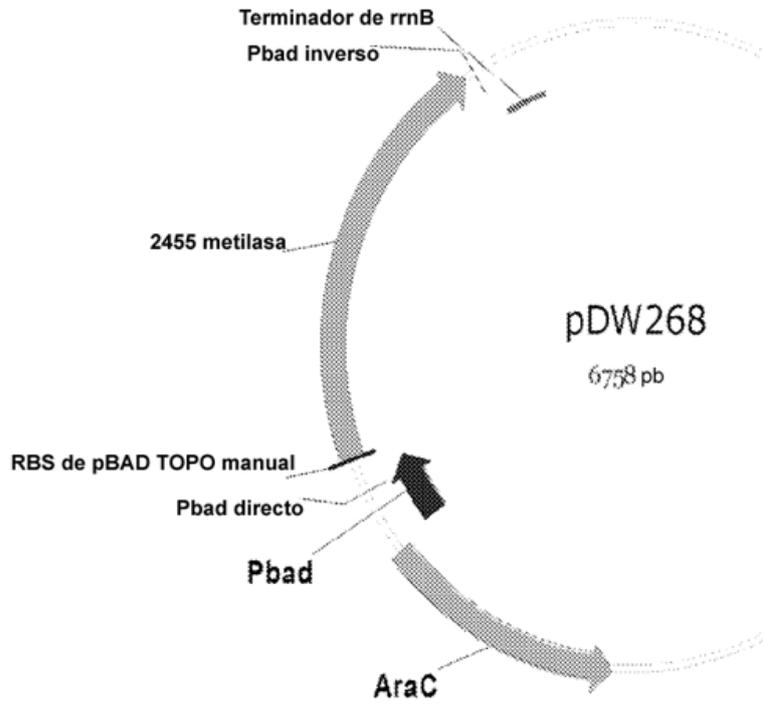
**FIG. 22B**

cacttaggtttgacatgataattataagtcaattgataggtgatagatgcacaaatacgtgtctattgaatacaatcatatt  
 catcggaaagtcaataacttttgataggtattggctaaacctattcaaaataaaagtattgcagaagtgcaatattggtat  
 ggagttagagcaaggattggagtaatttctcgctattactccatggacttcaaaacactataggaaaattataacgcac  
 ataaaagggtctcagatttaaagggaaagaaaaaagtagcgtagcgtggactttttctccctttaaatacaagaaatata  
 atgttcgtaaaaaaatgaatcctgatgtcatggatcacgtggcagcagicaatatttagatctaaaaattgaataatcca  
 acaaataggaggtgtgtaaaataatgttcgtgattataggttaagttaagtctgaggtaaatctatgaaatgcgatta  
 agggccggccagtgggcaagttgaaaaattcacaaaaatgtggtataaatctttgttcattagagcgataaactgaattt  
 gagagggaaacttagatggattttgaaaaattgataaaaatagttggaacagaaaagagatatttgaccactactttgcaa  
 gtgtacctgtacctacagcatgaccgttaaagtgatcacacaaataaaggaaaagggaaatgaaactatactctgc  
 aatgctttattatattgcaatgattgtaaacccgcatcagagttaggacggcaatcaatcaagatggggaattggggat  
 atgatgagatgataccaagctatacaatattcacaatgatactgaaacattttccagcctttggactgagtgtaagtctgac  
 ttaaactatttttagcagattatgaaagtatagcgaacggatggaacaatcatagaatggaaggaaagccaaatgc  
 tccggaaaacatttttaatgtatctatgataccgtggcaacctcgatggcttaatctgaattgcagaaaggatagattatt  
 tgattcctattttactatgggaaatattataaagaagataacaaaattatacttctttggcaattcaagttcatcacgcagt  
 atgtgacggatttcacatttgccgttttgtaaacgaattgcaggaattgataaatagttaacttcaggtttctgtaactaaaa  
 acaagtatttaagcaaaaacatcgtagaaatacgggtgtttttgttacctaagttaaacctcttttgataatctatgacca  
 aaatccctaacgtgagtttctgctcactgagcgtcagaccccgtagaaaagatcaaaggatctcttgagatcctttttct  
 ggcgtaatctgctgctgcaaacaaaaaaaccaccgctaccagcgggtggtttgttgccggatcaagagctaccaactc  
 ttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttagtgtagccgtagttaggccaccacttc  
 aagaactctgtagcaccgctacatacctcgctcgtctaatcctgttaccagtggtgctgcccagtgccgataagtcgtgc  
 ttaccgggtggactcaagacgatagttaccggataaggcgcagcggctcgggctgaacggggggtcgtgcacacagc  
 ccagctggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccagcttcccgaa  
 gggagaaaggcggacaggtatccggtaagcggcagggctcgaacaggagagcgcacgaggggagcttccagggg  
 gaaacgcctggtatctttatagctctgctgggttcgccacctctgactgagcgtcgattttgtgatgctcgtcaggggggc  
 ggagcctatggaaaaacgccagcaacgcggccttttacgggtcctggccttttctggccttttctcacatgttcttctgc  
 gtatccctgattctgtggataaccgtattaccgctttgagttagctgataaccgctcggcgcagccgaacgaccgagcg  
 cagcagtgtagtgagcaggaagcgggaagagcggccaatacgcagggccccctgctcgggggtcattatagcgatttt  
 ttcggtatattccatctttttcgcacgatatacaggattttgccaaagggtcgtgtagacttccctgggtgatccaacggcgtc  
 agccgggcaggataggtaagttaggcccaccgagcgggtgttcttcttactgtcccttattgcacctggcgggtg  
 ctcaacgggaatctgctcgtcggaggtggccggctaccgcccggtaacagatgagggcaagcggatggctgatga  
 aaccaagccaaccaggaaggcagcccacatcaagggtgactgcctccagacgaacgaagagcgtattgagga  
 aaaggcggcggcggccggcatgagcctgtggcctacctgtggcctggccagggctacaaaatcacgggcgtc  
 gtggactatgagcacgtccgagcgtggcccgatcaatggcgacctgggcccgtggcggcctgctgaaactctgg  
 ctaccgacgacccgagcagggcgggttcgggtgatgccagatcctcgccctgctggcgaagatcgaagagaagca  
 ggacgagcttggcaaggatgatggcggtggtccgcccagggcagagccatgacttttttagccgctaaaacggcc  
 ggggggtgctgctgattgccaagcagctccccatgctcctcaagaagagcagcttccgggagctgggtgaag

**FIG. 22C**

tacatcaccgacgagcaaggcaagaccgatcgggccccctgcaggataaaaaaattgtagataaaatttataaaatagttt  
 atctacaatttttatcaggaacagctatgaccgcgccgccagctatagcagctactcttggattattatcaaaatgcttaa  
 taaaatagattacaaaagtgtctatacatgatagatatatattaatgatataaggggggtgatagattgtttacaaggaacc  
 agaaactaaaaataagctttagttcttagaatgacagaaacgcaaaagaagatacttgagattatggctaagagagagg  
 ttatcacaatcagaattaattatgatattatggagaatgaattcaagaagcctgtattagaataaagcagcaagattaaactt  
 gccgccttgatagcggagcaacggttttatccaagcggtaaacaatattctaacagcgggtgtttaaaattatcaactagaa  
 gtgtattaatggctgcggaaagaaatattaaccagctactacacaatcgcaccttaaagtaaggttttaatgttaatttgg  
 cacggaactgctcttcttgatataattacaacaagtcggctaaaattgaaatttaacgttatccgaaaggggggcaaaattt  
 ggatgagaagatacttaaagatgaagggttctaaaaatcattacaatcggttcataataataatcagtataataagttgattg  
 taggttattacaatcaatacatagaagattctagacctgtaaagaagaaaaagactatttggattatactagatttactatgaa  
 gattatttgttgaaaaattagaacataaaagagataagttagctaattgtaataagaaatgggaagttgaagtttatgaaaaa  
 cttaaagtaaaagattatgtgtctactttattatgtaatgataagtttttagtaattgtaagaaagtaagcaagctcaaggatg  
 gcgaaaaatatgcccttgctgaacagataaagataaattatatacaaatggtttaactacaccaaatattgtagatcatacag  
 ggaagaattgaaaaagagattaaaaagcaatttaagcattaacttattaacagaatatttaaaaggtaaaaaacaag  
 taaagggttagattttgatattggatacttaggtgcaataaggtcgttggaggtaacttatagcggtgactattatcatccgcattt  
 gcatttgatattagattggataatcaaaatgaattataacagataaaaaaaatataaataactattcttatgattattataaaaa  
 aagaccaactagattatttcagattttgaaatattgttacagaaatcttggtatctttatataatggggaaagattgactaagga  
 aaatatagataaaactggaaaaagggtatagttgcatgatggataaggcaaaagaagatgatttttagaagttttaatacat  
 ggtgaagaatgatccggcagaggagaatgtaaaggtagtaacaaaatgacttataaaaattttagagtattagaatagc  
 attgcatagataagacagatacaagggtatggagtttttataatattaagatatattaatggctgaagaaglaaatgaaatgt  
 atgaatggataagagagatttaatacaaaaatgaaggagaagctcctgcataatcgtgtgagaagatacagaagcttctaga  
 tgatactgagtatactcttatcaaggaaaaaaatattacgtatthaagaaaaatatactctgaataataacattatagcataa  
 agagggcttaattgctctcttttttaattcttttaagcttcaattgggtgatgtttaaagattacagtaaaatcgccgaaagcca  
 cggttcaatcgtgggatgaaaggcgtttcttttaattcttctgttcagttcagtttaaaactgatactataaatatattgggacaag  
 attatagaagaacacaaacaacagtatcttaataaactatcattttgtttctgtccaaggtacagacgtaaagttctagttgga  
 gaagttgaaataaaaattaaacagcttctcaatgagattgtaaagacattgaaatagaaatttggcaatagaatgtgataaa  
 gaccactgccatcttttgcattgcacttctcatttaagtcagcagacataatggcaaaagtgaaaggagtgacttctcgat  
 tattaaggcaggaattaaacatctgcgacatttgccaagctttggacaagaagctatttgtatctaccgcaggaaatgtatc  
 aagtgaactataaaacgatatg

FIG. 23



**FIG. 24A**

gcaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgca  
ggaccacttctgcgctcggcccttccggctggctggttattgctgataaatctggagccggtgagcgtgggtctcgcggtat  
cattgcagcactggggccagatggtgaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggat  
gaacgaaatagacagatcgcctgagataggtgcctcactgattaagcattggttaactgtcagaccaagttactcatatata  
ctttagattgattacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacact  
gccagcgccttagcgcgccctccttctgcttcttcccttcttctcgcacgltcgcggcttccccgtcaagctctaaatc  
gggggctccccttaggggtccgatttagtgctttacggcacctcgaccccaaaaaacttgattgggtgatgggtcacgtagt  
gggcatcgcctgatagacgggtttcgcctttgacgttggagtcacgttcttaatagtggtactctgttccaaactgaa  
caactcaaccctatctcgggctattctttgattataagggatttggcgtattcggcctattggttaaaaaatgagctgatt  
aacaanaaattaacgcgaatttaacaaaatattaacgtttacaatttaaaggatctaggtgaagatccttttgataatctc  
atgacaaaatccctaactgagtttctgctcactgagcgtcagacccgtagaaaagatcaaggatcttcttgagat  
cctttttctgcgcgtaactctgctgctgcaacaaaaaaaccaccgctaccagcgggtggtttgttgcggatcaagagct  
accaactctttccgaaggtaactggctcagcagagcgcagataccaatactgtccttctagttagccgtagttaggc  
caccactcaagaactctgtagcaccgctacatacctcgcctgtaactcctgttaccagtcaggcatttgagaagcaca  
cggtcacactgcttccggtagcaataaaccggtaaacagcaatagacataagcggctatttaacgaccctgcctga  
accgacgaccgggtcgaatttgccttgaatttgccttcatccgcttattatcactattcaggcgtagcaccaggcgtta  
agggcaccaataactgccttaaaaaaattacgccccgcctgccactcatcgcagtagtctgtaattcattaagcattctg  
ccgacatggaagccatcacagacggcatgatgaacctgaatcgcagcggcatcagcacctgtcgccttgcgtataat  
atttgccatggtgaaaacgggggcgaagaagttgtccatattggccacgtttaaataaaaactggtgaaactcaccag  
ggattggctgagacgaaaaacatattctcaataaacccttagggaaataggccagggtttaccgtaacacgccacatc  
ttgcaatatatgtgtagaaactgccggaaatcgtcgtggttactccagagcgtgaaaacgtttagttgtcatgg  
aaaacgggtgaacaagggtgaactatcccatacaccagctcaccgcttctcattgccatacggaaattccggatgagc  
attcatcaggcgggcaagaatggaataaaggccggataaaaactgtgcttattttcttaccggtcttaaaaaggccgtaa  
tatccagctgaacggctcgtttaggtacattgagcaactgactgaaatgcctcaaaatgttcttaccgatgccattgggat  
atatcaacgggtggtataaccagtgatttttctccattttagcttcttagctcctgaaaatctcgataactcaaaaaatagcc  
cggtagtgatcttattcattatggtgaaagttggaacctctacgtgccgatcaacgtctcattttcgccaaaagttggccca  
gggcttcccggatcaacagggacaccaggatttatttctgcgaagtatctccgtcacaggatttattcggcgcaaa  
gtcgtcgggtgatgctgccaactactgatttagtgatgatgggttttgagggtcctcagtggtctgttctatcagctgtc  
cctcctgttcagctactgacgggggtgctgtaacggcaaaagcaccgcccggacatcagcgtagcggagtgatactg  
gcttactatgttggcactgatgaggggtcagtgagtgctcatgtggcaggagaaaaagggtgcaccgggtgcgtcag  
cagaatatgtgatacaggatatactccgcttctcgtcactgactcgtacgctcggctcgtcactgcggcgagcggaa  
atggcttacgaacggggcggagatttctggaagatgccaggaagatacttaacaggggaagtgagagggcggcggc  
aaagccgttttccataggctccgccccctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaaccgg  
acaggactataaagataaccaggcgttccccctggcggctccctcgtgcgtctcctgttctcgttccgttccgggtgtc  
attccgctgtatggccgcttgtctcattccacgctgacactcagttccgggttaggcagttcgtcctcaagctggactgtat  
gcacgaacccccggtcagtcaccgctgcgccttaccggtaactatcgtcttgagccaacccggaaagacatgca  
aaagcacc

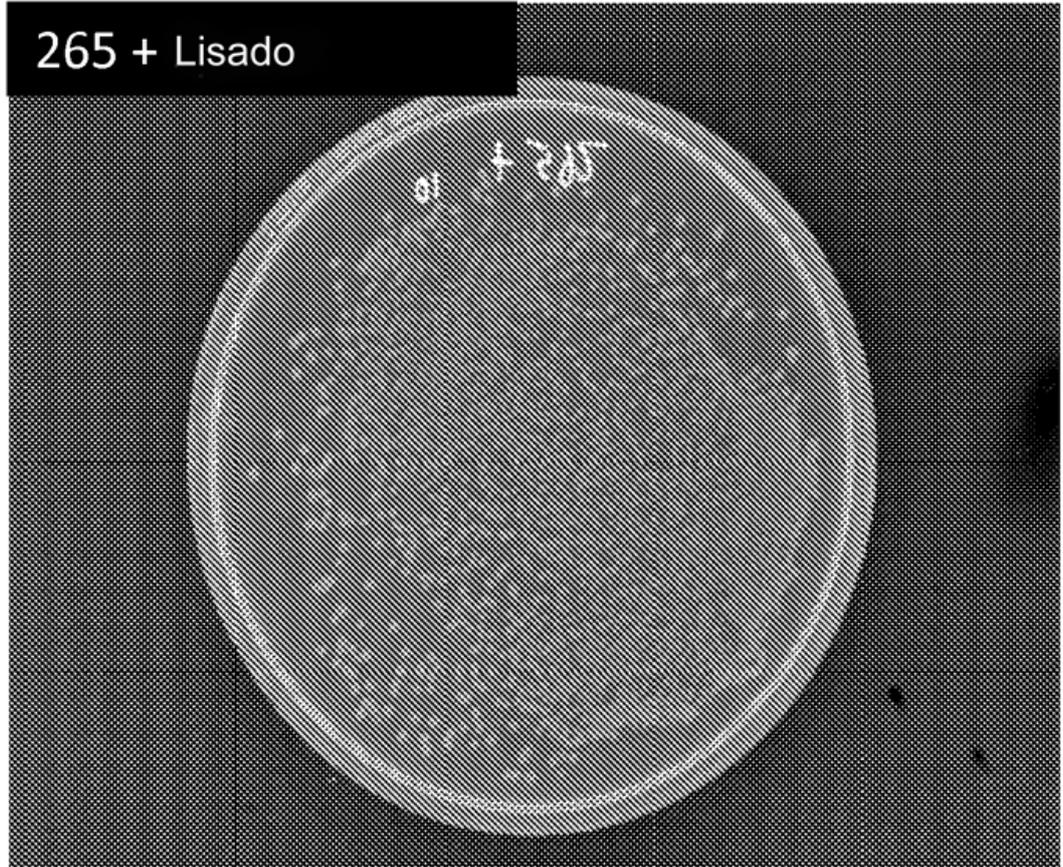
**FIG. 24B**

actggcagcagccactggtaattgattagaggagtagtctgaagtcacgcccggtaaggctaaactgaaaggaca  
 agttttggtgactgcgctcctccaagccagttacctcggtcacaagagttggtagctcagagaacctcgaaaaaccgccc  
 tgcaaggcggtttttcgtttcagagcaagagattacgcgcagacaaaacgatctcaagaagatcatcttattaatcaga  
 taaaatatttgctcatgagcccgaagtggcgagcccgatcttccccatcgggtgatgctggcgatagggcgccagcaacc  
 gcacctgiggcgccggtagtgcggccacgatgcgtccggcgtagaggatctgctcatgtttgacagcttatcatcgatgc  
 ataatgtcctgtcaaatggacgaagcagggattctgcaaacctatgctactccgtcaagccgtcaattgtctgattcggt  
 accaattatgacaactgacggctacatcattcactttttctcacaaccggcacggaactcgtcggggtggccccgggtgc  
 atttttaaatacccgcgagaaatagagttgatcgtcaaaaccaacattgacgaccgacgggtggcgatagggcatccgggtg  
 gtgctcaaaagcagcttcgctggctgatacgttggctcctgcgcccagcttaagacgctaaccctaaactgctggcggaa  
 aagatgtgacagacgcgacggcgacaagcaaacatgctgtgacgctggcgatatacaaatgctgctgcccagggtg  
 atcgtgatgtactgacaagcctcgcgtaccggattatccatcgggtggatggagcgcactcgtaatcgttccatgcgccg  
 cagtaacaattgctcaagcagattatcgccagcagctccgaatagcgccttcccctgcccggggttaatgattgccc  
 aacaggctcgtgaaatgcccgtggtgcttcatccgggcgaaagaaccccgtattggcaaatattgacggccagttaa  
 gccattcatgccagtagggcgcgaggacgaaagtaaaccactgggtataccattcgcgagcctccggatgacgaccgt  
 agtgatgaatctcctggcggaacagcaaaaatcacccggctggcaaaacaattctcgtccctgattttaccacc  
 cctgaccgcaatggtgagattgagaatataaccttaccagcggctcgtgataaaaaatcgagataaccggtg  
 gcctcaatcggcgtaaacccgccaccagatgggcataaacgagatcccgccagcaggggatcatttgcgctttag  
 ccatactttcactcccgcattcagagaagaaccaattgtccatattgcatcagacattgccgtcactgctctttact  
 ggctcttcgctaaccaaaccggaaccggcttataaaagcattctgtaacaaagcgggaccaaaagccatgacaaa  
 aacgcgtaacaaaagtgctataatcacggcagaaaagtccacattgattattgacggcgtcacactttgctatgccat  
 agcattttatccataagattagcggatcctacctgacgctttttatcgcaactctactgtttctccataccggtttttgggcta  
 gcgaattcgagctcggtaaccggggagggaataataaatggccgtactccgcaatattgatgagcaactgaccgaggaa  
 ttaagaaactgccgatcgaactattgggactttgagggtagggacacgaaagaactgacgcacggcctgcacaactatc  
 cggcgggtgatggttatccgatctaccgtaacattatcgacatcgtgaagcgtcacgggtgaggtcgaaaccttctggacc  
 gttatgggtagcggtagcggcctgggtggaaggcaagctggcgggttcaacaaagtgtacggtagcggatctgaatcctc  
 tggcagtgctgctgagcaaggtaagaccaccgcttgaagaggatagcgtggatattcaggacaagctgctgcgcg  
 agaatttgagcaggcgttcgtgtccagcaaacagctgctggataacattgacaattacattgaggagaaggcctgga  
 cgtcagcgccaaagacggctggggctctgatgcgcatgctatttgcgagatctggatacctacaacagcggctgga  
 aatcccagacttaagaatatgggtattgggtcaaaccgcggtattctggagctgcaactgattaaggatattctgc  
 agatcgagaatgaggactccgtaacttcttctggctgcttctgaaactgcccgtactgagcaaacaccggtaatgg  
 gactcaagctgtccgtatcaagaaagaaaaagtggcagattcaatccggacgtaagatcgagtttacaataatctg  
 gatcgtaacatcgaaaagattaaagactttgacaaacggtgtaacaacgattgcaagttagcgttgcittgaaagatacc  
 cgattctggactcgggtccggacaatagcatcgaatcgtgattaccagcccaccgtacggcgatagcaaaactacgggt  
 ggcgtacggtaatttagccgtccgtcttgggtgggtggatctggaattgatggacatcgaagagctgaatcaagttgac  
 aacaatctgctgggtggaagaagggtggacaaagacttcgagtgtaactgagctcccgtaccttgagaaggcgatta  
 aagaaatcaagaaaaggacctggaccgcacgtgacgtttatagctctacgaggattggataaggctatggagtc  
 cattacgaaaaagatgc

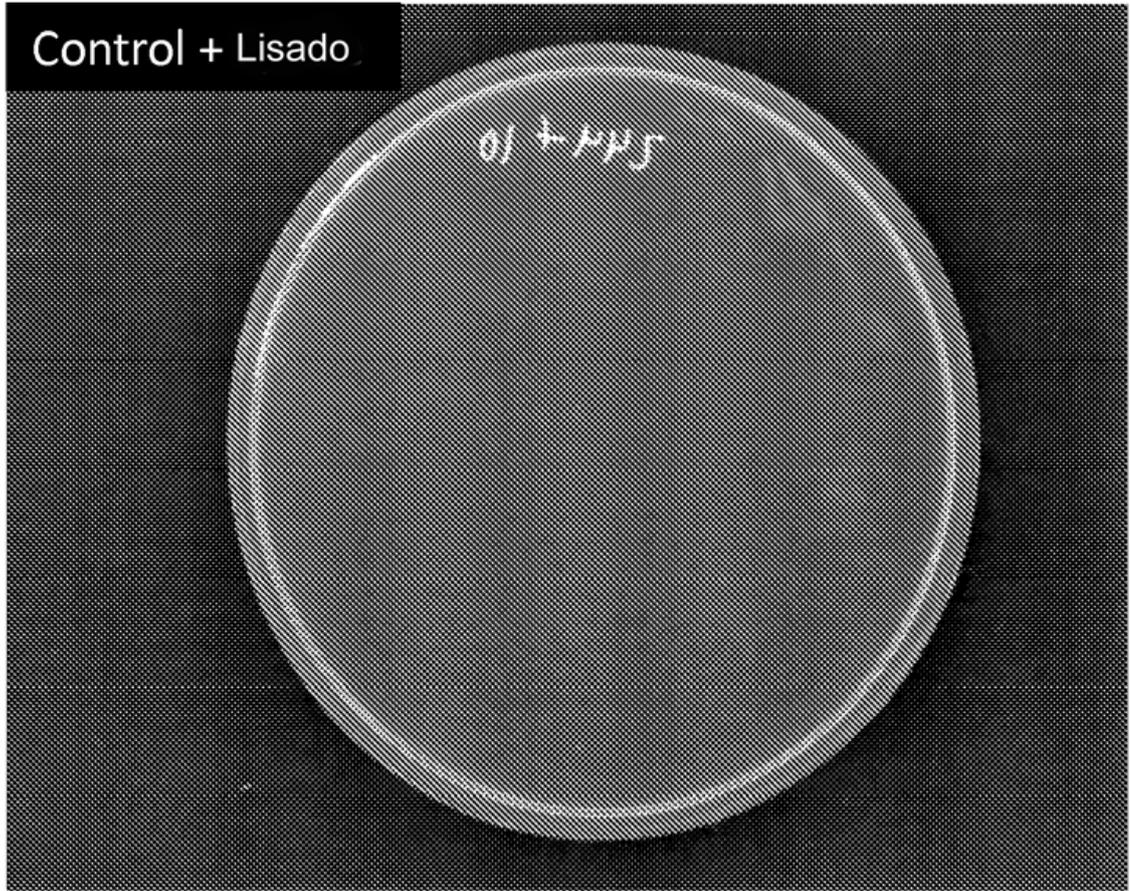
**FIG. 24C**

gtcataacagctaccagttctggggtgtcggtaaccgtaccgttaaagaagtcaaactgctgaccaacgaaat  
cattagcgaactgggcgagaaatatggtttggtgaggttacgatatcccgcgtaacatcccgaataaggatcat  
gccgagccgtaattccccgaccaatgaaaccggcaagacggtcagcaccatgacgaacgagcacatcgtc  
gtgctgcgcaaagatcgttgaggctgtttggcggatgagagaagatttcagcctgatacagattaaatcaga  
acgcagaagcggctgataaaacagaatttgctggcggcagtagcgcgggtgggtcccacctgaccccatgc  
cgaactcagaagtgaaacgccgtagcggatggtagtgtgggtctccccatgcgagagtagggaactgc  
caggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgtttatctgttgttgcggtaacgct  
ctcctgagtaggacaaatccgcccgggagcggatttgaaactgacgagcaacggcccggagggtggcggg  
caggacgcccgcataaactgccaggcatcaaattaagcagaaggccatcctgacggatggccttttgcgtt  
tctacaaactctttgtttatcttaataacattcaaatatgtatccgctcatgagacaataaccctgataaatgctt  
caataatattgaaaaaggaagagatgagtattcaacattccgtgtcgccttattccctttttgcggcattttgcc  
ttcctgttttgctcaccagaaacgctggtgaaagtaaagatgctgaagatcagttgggtgca

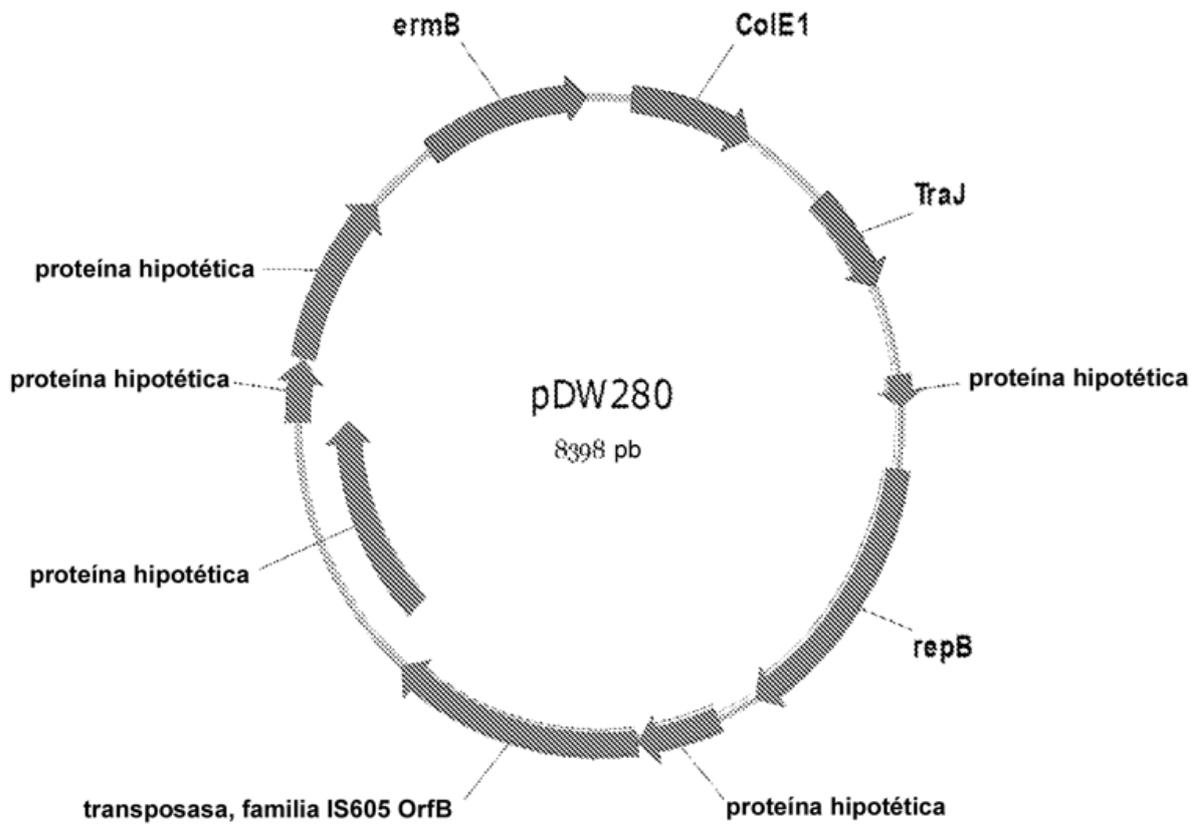
**FIG. 25**



**FIG. 26**



**FIG. 27**





**FIG. 28B**

tttagagtattagaatatgcattgcatagtataagacagatacaagggtatggagttttataatattaaagatatattaatgg  
 ctgaagaagtaaatgaaatgtatgaatggataagagagatlttaaatcaaaaatgaaggagaagctcctgcataatcggtt  
 gagaagatacagaagcttctagatgatactgagtatactcttataatcaaggaaaaaatatttacgtatttaagaaaaatat  
 actctgaataataacattatagcataaagagggcctaattgctctctttttaaattcttttaagcttcattgggtgatgtttaata  
 gattacagtaaaatcgccgaaagcccacggttcaatcggggatgaaaggcgttcttttaaatcttctgtgacagttcagttt  
 aaaactgatactataaataatgggacaagattatagaagaacacaacaacagtatcttaataaactatcatttggtttct  
 gtccaaggtacagacgtaaaagtctagttggagaagttgaaataaaatlttaaacagcttctcaatgagattgtaaagaca  
 ttgaaatagaaatlttgcaatagaatgtgataaagaccactgccatcttttgcaatgcacttctcatttaagtccagcag  
 acataatggcaaaagtgaaaggagtgacttctcgattattaaggcaggaattaaacatctgcgacattgccaagctttg  
 gacaagaagctatttctatctaccgcaggaaatgtatcaagtgaactataaaacgatatgttgaagaacaaaaaca  
 aggggtgaaacaatgcagataacagtaaaatlttaataatltttgacaaaagaacaagtacaactaataagaatctatat  
 caaaagaatataatccatactgttaatagccttgttcatctacgctccaatcagaagaagagtaaagctatcatctaaaga  
 tgttttgcaaatatgccaagtgacgtaaaaaatcaatctattagagatgccaaaagtatctgtactaagtaacaagaagct  
 atcaaggctaattccaactgcctactgataaacaanaagtaaatcaatgtagctacccttaaaaaacctgtctgtatagg  
 aataatcaaaatltcactaaagacggtatttcttagtttcccgttattatagatgggaaatcgcagcgtattcaaaactaga  
 actatcatgacagactatcagctaaaacaactagaaggctcattgggagcattgctgataactaagaaaagcaataaat  
 atatcgtcaataaggttgtaaaaagtatctatagttaaagggtgatgttgaatgggtgtgacttaggcctaaaagttc  
 ctgctgtagctgaaccgattcaggaaaaacgttttttggaaacggtaggcaaaaataaatacgtcaaacgtaaatataa  
 agcgaacgtaaaaaacttggaaaagccaagaagcctaaagcttaaaaagctgatgataaagaacaacgttggaa  
 tgacagaccaagaccacaaagtaagtagagaaataaataatlttgacgtaaaataataatgtttctgatattcggctgaaa  
 aattaacgaatatcagaaacacggcaagaacaagccgtaaaaacgaaaaaatctacatacatggctcattctatcgtc  
 tagctcaattcatagagtataaggcactattgaaggggataaagggtgaaatgttgatcctaaatacacttctcaaatatgc  
 cctgaatgtaagaactaaataaagcaagagatagaaaataaataatgctcctgtggttttaaacaacatagggatagag  
 taggtgtataaataaataatgtcacctgtagtagatggtaaaagtctactagcctagggtagctatgtactgtcttagga  
 ggggtaatggcataccctaagcttgaggctactccgatagcagaaatgtactcggtttaactcaagaatcccact  
 gcttagctgtgggagtgcaaatgaagcatgatggctcattatctgtaactagtgaaaggaagattgtattatgctggtagtca  
 aaaaattagtttaatagtggtatacctttaaatacaggagatggagttgttgggaatgaaatcaagatttaattcaact  
 ctgatgttattccgatgttactttaaaggatgaaatgcaaatcaaatatccaataaataatlttgaatatgatggaaaaga  
 accgattagcaatccgttttgggattatgaaaacttacatacaggtagtagaagattgatataggtgcaatccagattat  
 cagctctagtagggaaaacatatgaagatgttattagtgaaaatccaagtcaacaaaatcctatggtgcctccgatacca  
 ttctctgattcatggttggcaaatggaaagatatagttaacgatagtggaacatggcaaggggaaggcatagatggaag  
 tactggaactgcaatagatagctcctcattagatattcttggaaacgtggcaaggcaaatggcttggacagcagacggct  
 aattagtttctgatggttcttttcagggtctgacggaacaacatggcaaggaacatatacgcatacaggaatagggttcag  
 aatctgtactaaatccaccactaaccggatttaacaggaataacaggttgggtatcatctataagttcatggttaactag  
 ttgtttgcgtttcaactgatttagtttgaatlttagaccggtgaaaaatctacctatagcaacaaaatltccttctgtttgccatt  
 tgattaaaaaatagcattgaatcattgcaatctcctgtcgtgtccagttttacgactacttggaaattaccctttatcaagg  
 agatatagagattaatlttagcagctatggaacgattgacacaaataacacgttggggaacgtaattgtatttaacttgggtt  
 aatactgttacaaggaagggttatcatgatatggcaagcactagcatcttttaataatctacttataaagcattaggaacg  
 gtttaggggcaattatcggattattacctcaagctcttcaaacatttcaaacattcagcagtaacagaatatttaggatgtt  
 gaattggt

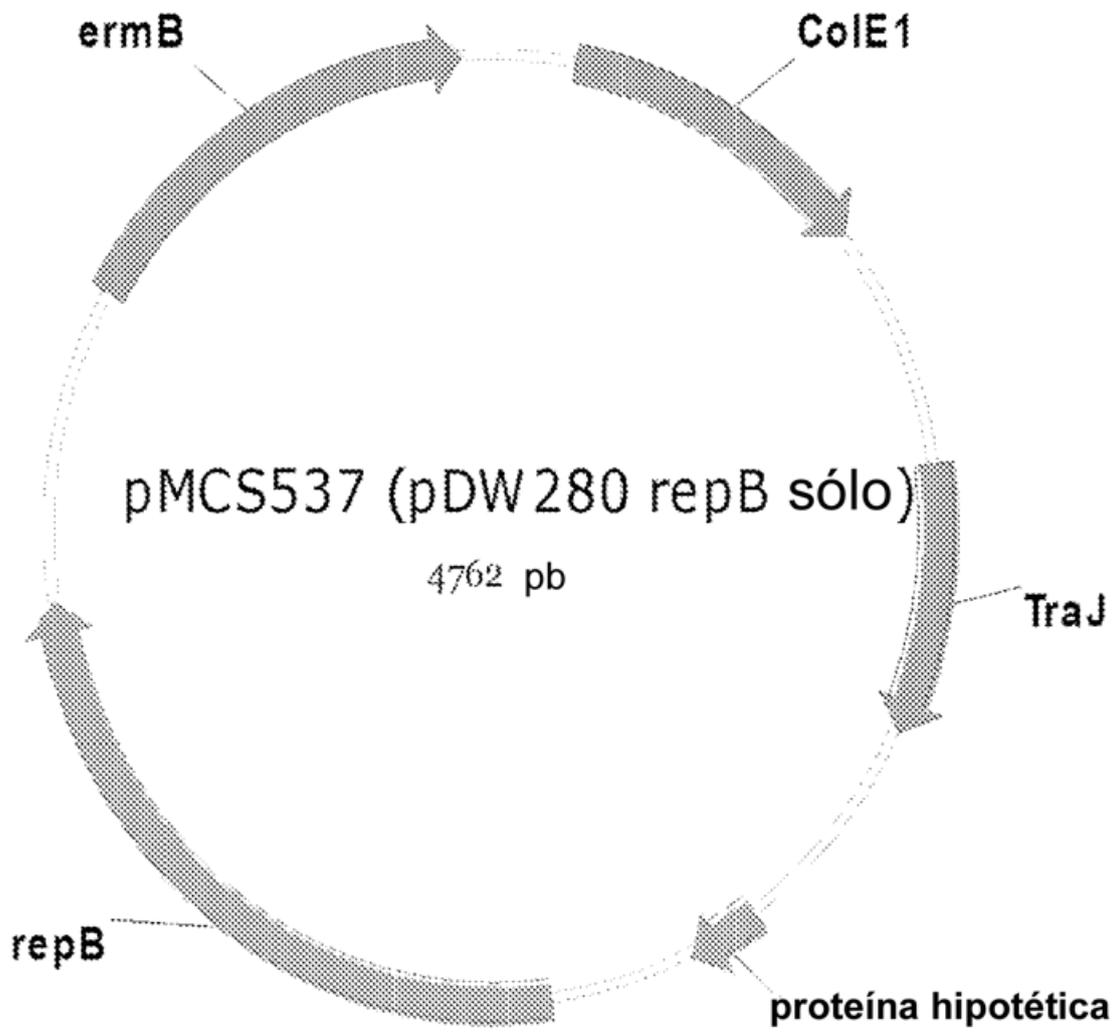
**FIG. 28C**

ttatatccgtagatgccatgataactatattaacttactggactactgcaattataagttactatgtaatatcaactgcgatgag  
 atggggaaaaacaattgaatagggggataatatgataagttttatagtggtactccaggaagtggaaaaagtcttaatat  
 agctagatacatatggattaaagttcgacatgctaaacaaaatataataacttgtaatatgacagtaatagagagtatctta  
 ttacatcaaaactgaagcaacttgtaataaaaattagattgaaattaaacttaaacttataactaagttaaaagactat  
 ggcaaaatctattciataagactcgatcagctgaacacaaaatttctagaagattatgctatgaaattcacatgggggca  
 ttgaaggacaatcaaaaataataatagatgaggcacaactgatttgggtccccaacggatgaaaaataaaaagcag  
 gtagaccctaattatcggaacgctggatagagttatgacactccatagacacttaggtttgacatgataattataagtca  
 attgatagggtgatagatgcacaaatcgttcttattgaatacaatcatattcatcggaagcaataactttgtataggtt  
 attggctaaacctattcaaaaataaaagtatttgcagaagtgcaatattggatggagttagagcaaggattggagtaatttc  
 ttcgctattactccatggacttcaaaacactatagggaaaattataacgcacataaaaagggttcagattaaagggaaag  
 aaaaaagtagcgtagcgtggactttttctcccttaaatcaagaaatataatgttcgtaaaaaaatgaatcctgatgcat  
 ggatcacgtggcagcagtcataatttagatctaaaaattgaataatccaacaaataggagggtgtgtaaaataaatgtt  
 cgtgattataggtaatgtaagtctgaggtaaatctatgaaatgcgattaagggccggccgaagcaaaactaagagtg  
 gttgatagtgacagatcttaaaattttgtataataggaattgaagttaaattagatgctaaaaatttgaattaagaaggagtg  
 attacatgaacaaaaatataaaatattctcaaaacttttaacgagtgaaaaagttactcaaccaaataaaaaacaattg  
 aatttaaagaacacgataaccgtttacgaaattggaacaggtaaagggcatttaacgacgaaactggctaaaaataagt  
 aaacaggtaacgtctattgaattagacagtcattctcaacttatcgtcagaaaaataaaactgaatactcgtgtcacttt  
 aattcaccaagatattctacagttcaattccctaacaacagaggtataaaattgttgggagtattccttaccatttaagcac  
 acaaattataaaaaagtggttttgaaagccatgctctgacatctatctgattgttgaagaaggattctacaagcgtacctt  
 ggatattcaccgaacactaggggttgccttgcacacicaagtcctcgattcagcaattgcttaagctgccagcggaatgcttc  
 atcctaaaccaaagtaaacagtgcttaataaaaacttaccgccaataccacagatgtccagataaatattggaagctat  
 atacgtactttgttcaaaatgggtcaatcgagaataatcgtcaactgttactaaaaatcagtttcatcaagcaatgaaacac  
 gccaaagtaacaatttaagtaccgttacttatgagcaagattgtctattttaaatagttatctattttaacgggaggaat  
 aattctatgagtcgcttttgtaaatttgaaagttacacggtactaaagggatgtgtt

**FIG. 29**



**FIG. 30**



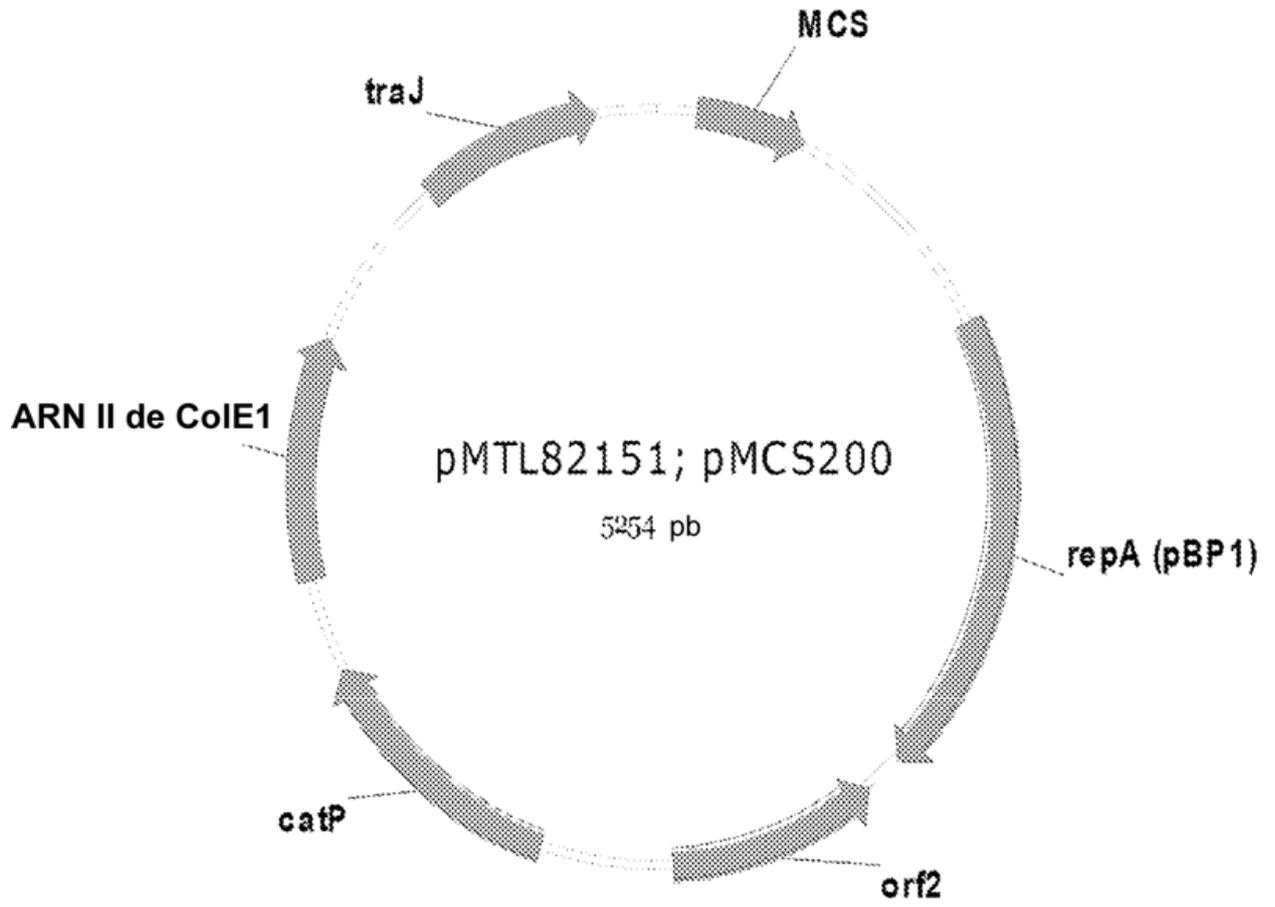
**FIG. 31A**

aaactccttttgataatctcatgacccaaaatccctaacgtgagtttccactgagcgtcagaccccgtagaaaagatca  
 aaggatcttctgagatcctttttctgcgctaatctgctgcttgcacacaaaaaaaccaccgctaccagcgggtgtttgttc  
 cggatcaagagctaccaactcctttccgaaggaactggctcagcagagcgcagataccaaatactgttcttagttagc  
 cgtagttaggccaccactcaagaactctgtagcaccgctacatacctcgtctgctaactcgttaccagtggtgctgcca  
 gtggcgataagtcgtgtcttaccgggttgactcaagacgatggtaccgggataaggcgcagcggcgggctgaacgggg  
 ggtcgtgcacacagcccagctggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcg  
 ccacgctcccgaagggagaaaggcggacaggtatccggtaagcggcagggctggaacaggagagcgcacgagggga  
 gctccagggggaacgcctggtatctttagtctgtcgggttcgccacctctgactgagcgtcgatgtttgtgatgctcgtca  
 ggggggaggcctatggaaaaacgccagcaacgcggccttttaccggtcctggccttttctggccttttctcacatgttctt  
 tctctggttaccctgattctgtggataaccgtattaccgctttagtgagctgataccgctcggcgcagccgaacgaccga  
 gcgcagcagtcagtgagcaggaagcgggaagagcgcaccaatacgcagggccccctgctcggggcattatagcgatt  
 tttcggatatacctatcctttctgcacgataatacaggattttgccaaagggctcgtgtagacttcttgggtatccaacggcgtca  
 gccgggacagataggatgaagtagggccaccgcgagcgggtgttcttctcactgtcccttattcgcacctggcgggtgctca  
 acgggaatcctgctctgcgaggctggccggctaccgcccggcgaacagatgagggaacgggatggctgatgaaacca  
 agccaaccaggaagggcagcccacctaacaaggtgtactgcttccagacgaacgaagagcgtgaggaaaaggcg  
 gcggggccggcagtgagcctgctggcctacctgctggcctcggccagggctacaaaaacacggggcgtcgtggactatga  
 gcacgtccgcgagctggcccgcataatggcgacctgggcccgcctgggcccctgctgaaactctggctcaccgacgac  
 ccgcgacggcgcggttcgggtgatgccacgatcctcggcctgctggcgaagatcgaagagaagcaggacgagctggca  
 aggtcatgatggcggtggctcggccgagggcagagccatgacttttttagccgctaaaacggccgggggggtgcgcgtgatt  
 gccaaagcagctccccatgcgtccatcaagaagagcgcactcgcggagctggtgaagtacatcaccgacgagcaaggc  
 aagaccgatcgggccccctgcaggataaaaaaattgtagataaaatltataaaatagttttatctacaatlttttatcaggaac  
 agctatgaccgcccggccagctatagcagctacttgggtattattatacaaaatgcttaataaaatagatttacaaggtgct  
 atacatgatagatataatgataataggggggtatagattgtttacaaggaaccagaaactaaaaataagcttttagt  
 tcttagaatgacagaaacgcaaaagaagatacttgagattatggctaatgagagaggttatcacaatcagaattaattatga  
 tattatggagaatgaattcaagaagcctgtattagaaataaagcagcaagattaaactgccgccttggatagcggagcaa  
 cggttttccaagcggtaaacatattctaaacagcgggtgtttaaattatcaactagaagtgattaatggctcgggaaaga  
 aatattaaccagtaactatcacaattcgcaccttaaaagtaaggltttatgtttaaatttggcaggaactgatataatcaaaa  
 caagtcggctaaaattgaaatlttaacgltatcctgaaaggggggcaaaatltggatgagaagataactaaagatgtaaggg  
 ttcaaaaaatcatttacaatcgggtcataataataatcagataataagttgattgtaggtattacaatcaatcatagaagattct  
 agacctgaaagaagaaaaagactatlttgattatactagatttactatgaagattatltgtgaaaaatagaacataaaaag  
 agataagttagctaattgtaataagaaatgggaagttgaagttatgaaaaacttaaagtaaaagattatgtgctactttattat  
 gtaatgataagttttagtaattgtaagaaagtaagcaagcttcaaggatggcgaaaaatagccttctgctgaacagtata  
 aagataaattatatacaaatgglttaactacaccaaattgttagatcatacaggggaagaattgaaaaaagagattaaaaa  
 gcaatlttaagcattaaacttatttaacagaataatlttaaaaggtaaaaaacaagtaaaagggttttagattttgatattggatacttag  
 gtgcaataaggtcgttggaggttaactatagcgggtgactattatcatccgcatttgattgatattagattggataatcaaaatg  
 aattataacagataaaaaaaataaataactattcttagattataaaaaaagaccaactagattatlttcagattttgaaa  
 tattgttacagaaacttgggtatctttatataatggggaaagattgactaaggaaaaatagataaaactggaaaaagggtatag  
 ttgatgatggataaggcaaaagaagatgatttttagaagtttttaatacatggtgaagaatgatccggcagaggagaatg  
 taaaaggtagtaacaaatgacttataaaaaatlttagaglattagaatatgcattgcatagataagacagatacaagggtatg  
 gagtttttataatataaagatataatltggctgaagaagtaaatgaaatgtatgaatggataagagagattttatcaaaaaat  
 gaaggagaagctcctgcatacgtgtgagaagatacagaagcttctagatgatactgagataactcttatatcaaggaaaaa  
 aatatttacgtatttaagaaaaatatactctgaataataacattatagcataaagagggcttaattgctctcttttaatttcttaaa  
 gcttatttgggtgatgtttaaagattacagtaaatcgcctgaaagcccacggttc

**FIG. 31B**

aatcgtgggatgaaaggcgttctttaatcttctgttgcagttcagtttaactgatactataaatattagcgttggactttttctccctt  
aatcaagaaatataatgttcgtaaaaaatgaatcctgatgtcatggatcacgtggcagcagtcfaatatttagatctaaaaattga  
ataatatccaaacaaataggagggtgtgtaaaataaatgttcgtgattataggtaaatgttaagtctgagggtcaatctatgaaatgag  
attaaggccggccgaagcaactaagagtggtgatagtgagtcagtcctaaaatttgataataggaattgaagttaaattagat  
gctaaaaattgtaattaagaaggagtgattacatgaacaaaaatataaaatattctcaaaccttttaacgagtgaaaaagtactc  
aaccaataataaaaacaattgaattaaaagaaaccgataccgtttacgaaattggaacaggtaaaagggtcattaacgacgaa  
actggctaaaataagtaaacaggtaacgtctattgaattagacagtcactattcaactatcgtcagaaaaattaactgaatact  
cgtgtcactttaattaccaagatattctacagttcaattccctaacaacagaggataaaaattgtgggagttccttaccattta  
gcacacaaattataaaaaagtggttttgaaagccatgctgacatctatcgtattgtgaagaaggattctacaagcgtaccttg  
gatattcaccgaacactaggggtgctctgcacactcaagtcgattcagcaattgcttaagctgccagcgggaatgcttcatcctaa  
acaaaagtaaacagtgcttaataaaaacttaccgccataccacagatggttcagataaatattggaagctatatacgtactttgtt  
caaatgggtcaatcgagaatatcgtaactgtttactaaaaatcagttcatcaagcaatgaaacacgccaaagtaacaattta  
agtaccgttacttatgagcaagtattgtctattttaatagttatctattttaacgggaggaaataattctatgagtcgctttgtaaattg  
gaaagttacacgttactaaaggaatgtgtt

**FIG. 32**



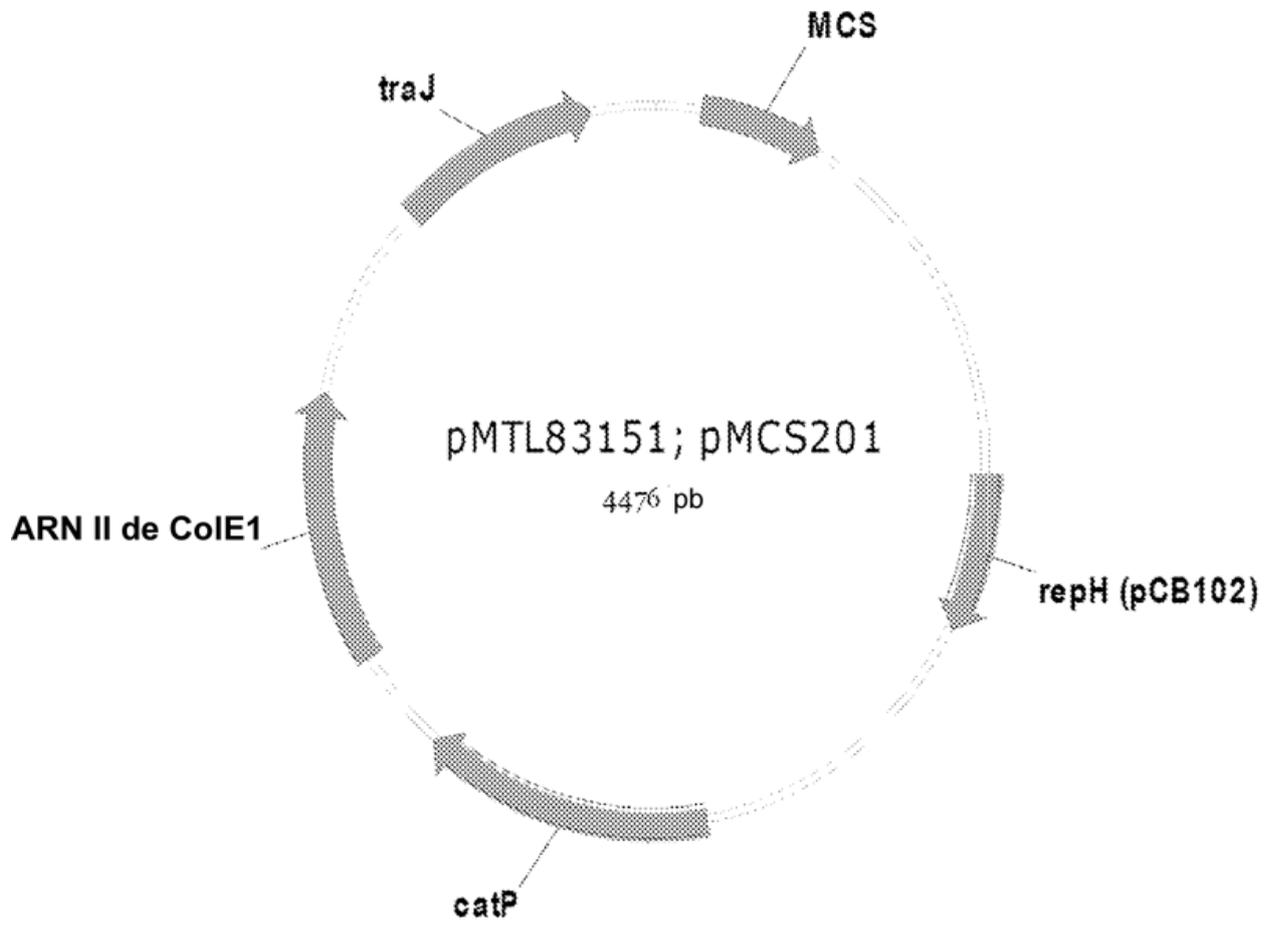
**FIG. 33A**

cctgcaggataaaaaaattgtagataaattttataaaaatagttttatctacaattttttatcaggaacagctatgaccgcgccgct  
 gtatccatgatgacctgattacgaattcgagctcggatcccggggatcctctagagtcgacgacgctccatggagatctcga  
 ggctgcagacatgcaagcttggcactggccgctggtttacaacgctcgtgactgggaaaaccctggcgttaccacacttaatcgc  
 ctgacgacatcccccttccgagctggcgtaatagcgaagaggcccgaccgatcgccctcccaacagttgagcagcctg  
 aatggcgaatggcgctagcataaaaaataagaagcctgcatttgcaggcttctttttatggcgccgcttctgaatccttagctaa  
 tggttcaacaggtaactatgacgaagatgacccctggataagctctgtaatggattcaaggcattaatgaagacgtgatataa  
 aatgtgctaataaaaaagaaaatgctgtaaaagagcctaaaatgagttcaaatggtttgaaatgattggtagtttaattatata  
 tttttctattggctatctcgatacctatagaatctctgttactttgttttgaaatataaaaaggggcttttagcccttttttaaaactc  
 cggaggagtttctcattctgatactatacgaactatttcggattgacttattgcaattaagctagtaaaatcaatggtaaaaaa  
 caaaaaactgcattttctacctagtaattataatttaagtgctgagtttaaaagtataattaccaggaaaggagcaagttttta  
 taaggaaaaattttcttttaaaacttatttctgtatagactaattataatcaaaaaatgaaaaataaacaagaggttaaaactg  
 cttagagaaatgtactgataaaaaaagaaaaatcctagattacgctacacatagcacccttaactactaagaaaaatltgaa  
 aggactccactgtggagattattgtttatgttgatgagcagacttagaacatttaaaattacataaaggtaattttgctggaata  
 gattttgccaatgtgtagttggcgactgctgtgaaggatagtttagaaatctattcttatggagcatttaagaaaaagaagaaat  
 aaagagttatattttaaactctacaactccaaatgtaaaaagtatgatcttaattattctataaacaatataataaatcttttaaaa  
 ataatggagcgttaaggaagttaaggatataactaaaggtatataagaaaatagaagtaactaccaaaaggaaaaataca  
 taacaaaggattatggaataaaaaaagattatcaaaaaaaggactgaaatgggtgatttagaacctaattttgatactt  
 ataactctcatttcatgtagttatgtagtaataaaagtattttacagataaaaaatattatataaatcgagaaagatggttgaat  
 atggaagtttctactaaggatgattctataactcaagttagttagaaaagcaaaaaatgattataaagaggttacgaactt  
 gcgaaatattcagtaagacactgatttaataatcgaggccagattttgaaatttttataaagcattaaaaggcaagcaggta  
 ttgtttttagtgatttttaaaagatgcacacaaatgtacaagcaaggaaaacttgatgtttataaaaagaaagatgaaatlaaat  
 atgtctatagtttattataaattgggtcaaaaaacaatagaaaaaactagaataagggaacttacggaagatgaaaaagaag  
 aatlaaatcaagatttaagatgaaatagaaatagattaaagtgaactatactttatataatgattaaaaaaataaaaaaca  
 acagcctattagggtgtgttttttttttatttataattttttaatttttagtttttagtcttttttaaaataagtttcagccctttttcaatattttta  
 aagaaggagtattgcatgaaltgcttttttcaacagacttaggaaatatttaacagatctcttgcgcccgtgattttggaacttc  
 ataactactaattataattatttttttttaattgtaacagttgcaaaagaagctgaacctgtcttcaactagttatcatcttcaat  
 ataattcttgacctatagataaataatattttattatattttactttttctgaaatctatttttataatcataaaaagtttaccacca  
 aagaaggttactccttctggccaacataattttttactatattatctaaataattttgggaactgggtgtgtaatttgattaatcgaaca  
 accagttacttaaaaggaattataactataaaaaatataaggattatcttttaaaattcatttggccctcttttttaaaattatgttac  
 cataaaaaggacataacgggaatagtagaataatttttaalgtagacaaaatttacataaataaagaaaggaaggtttgtttta  
 aattttatagcaaatcaaaaattagggggataaaaaatlagaaaaaaagggtttcgatgtttttatgtttaaactttaaagtttg  
 tggttttttacaaatcgccggccagtgggcaagtgaaaaatcacaaaatgtggtataatctttgttcattagagcgataa  
 acttgaattgagaggggaacttagatggtattgaaaaaattgataaaaaatagttggaacagaaaagagattttgaccactactt  
 gcaagtgacctgtacctacagcatgaccgttaaagtgatcacacaaaataaggaaaagggaatgaaactatactctgc  
 aatgctttattatattgcaatgattgtaaaccgcatcagagtttaggacggcaatcaatcaagatggtgaaatggggatataatgat  
 gagatgataccaagctatacaatattcacaatgatactgaaacattttccagcctttggactgagtgtaagctgactttaaactatt  
 itagcagattatgaaagtgatcgaacggatggaacaatcatagaatggaaggaaagccaaatgctccggaaaacatttt  
 taatgtatctatgataccgtggcaacctcgtatggcctttaaactgaatttcagaaaggatagattatttgattctattttactatggg  
 gaaatattataaagaagataacaaaattacttctttggcaattcaagttcaccgcagatgtgacggatttcaatttggcgtt  
 ttgtaaacgaattgcaggaattgataaatagtttaactcaggtttgtctgtaactaaaaacaagtaatttaagcaaaaacatcgtaga  
 aatacgggtttttgttaccctaagtttaaacctttttgataatctcatgacaaaaatcccttaacgtgagttttcgttccactgagcgt  
 cagaccccgtagaaaagatcaaaggatctcttgagatcttttttctgcgctaatctgctgctgcaaacaaaaaaaccaccg  
 ctaccagcgggtgtttgtttgcccgatcaagagctaccaactcttttccgaaggttaactggctcagcagagcgcagataccaaa  
 tactgttctctagtgtagccgtagttaggccaccactcaagaactctgtagcaccgctacataacctcgtctgctaactctgttac  
 cagttggctgctccagtgccgataagctgctgtctaccgggttgactcaagacgatagttaccggataaggcgcagcggctcgg  
 gctgaacgggggggtcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatg  
 agaaagcggcagcttccgaaggaggaaaggcggacaggtatccgg

**FIG. 33B**

taagcggcagggctcggaacaggagagcgcacgagggagctccaggggaaacgcctggtatctttatagtcctgtcgggttcg  
ccacctctgacttgagcgtcgatcttctgtgatgctcgtcagggggcggagcctatggaaaaacgccagcaacgcggccttttacgg  
ttcctggccttttgctggcctttgctcacatgcttcttctgcttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgata  
ccgctcgcgcagccgaacgaccgagcgcagcagcagtcagtgagcaggaagcgggaagagcgccaatacgcagggccccc  
tgcttcggggcattatagcgatcttctcgggtatccatccttttcgcacgatatacaggatcttgcctccttctcactgtcccttattcga  
gtgatccaacggcgtcagccgggcaggataggtgaagtaggcccacccgcgagcgggtgtccttctcactgtcccttattcga  
cctggcgggtgctcaacgggaatcctgctcgtcagggctggccggctaccgccggcgtaacagatgagggcaagcggatggctga  
tgaaaccaagccaaccaggaagggcagcccacctatcaaggtgtactgcctccagacgaacgaagagcgattgaggaaaag  
gcggcggcggccggcatgagcctgtcggcctacctgtggccgtcggccagggctacaaaatcacgggcgtcgtggactatgag  
cacgtccgcgagctggcccgcacatgagcgcacctggggcctggggcggcctgctgaaactctggctaccgacgacccgcgc  
acggcgcgggtcgggtgatgccacgatcctcgcctgtggcgaagatcgaagagaagcaggacgagcttggaaggtcatgat  
ggcgtgggtccgcccagggcagagccatgacttttttagccgctaaaacggccgggggtgctcgtgattgccaagcacgtcccc  
atgctcctcatcaagaagagcgcactcgcggagctgggaagtacatcaccgacgagcaaggcaagaccgatcggggccc

**FIG. 34**



**FIG. 35A**

```

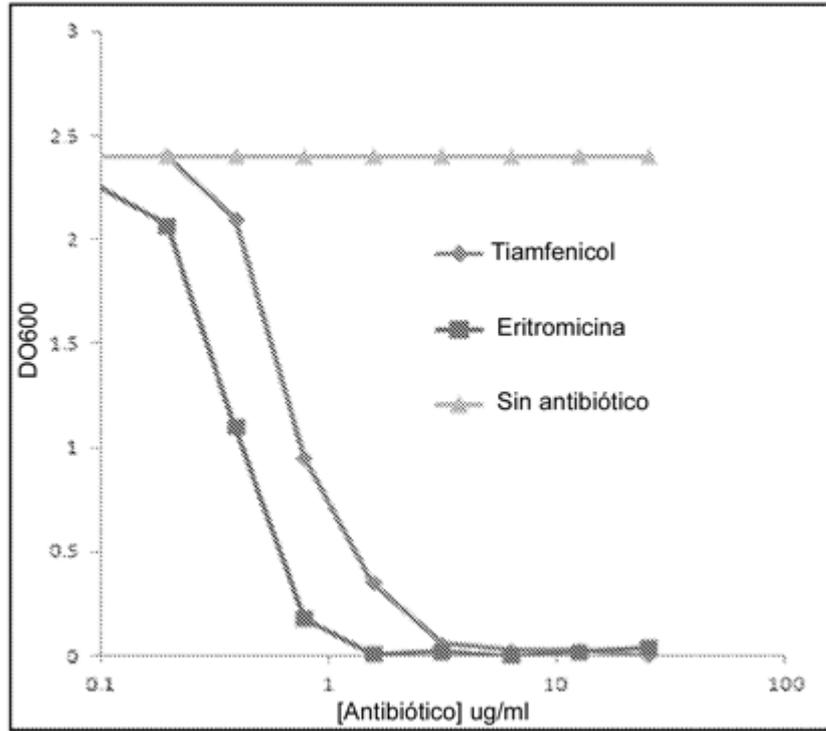
1  ggataaaaaaattgtagataaattttataaaatagttttatctacaatttttttatcagg 60
61  aaacagctatgaccgcggcogctgtatccatatgaccatgattacgaattcgagctcgg 120
121  acccggggatcctctagagtcgacgtcacgcgtccatggagatctcgaggcctgcagaca 180
181  tgcaagcttggcactggcogtcgttttacaacgtcgtgactgggaaaaccctggcggtac 240
241  ccaacttaatcgccctgcagcacatccccctttcgccagctggcgtaatagcgaagaggc 300
301  ccgcaccgatcgcccttcccacagttgocgagcctgaatggcgaatggcgctagcataa 360
361  aaataagaagcctgcatttgcaggettcttatttttatggcgcccgccattatttttt 420
421  gaacaattgacaattcatttcttattttttattaagtgatagtcaaaaggcataacagtg 480
481  ctgaatagaaaagaatttacagaaaagaaaattatagaatttagtatgattaattatact 540
541  catttatgaatgtttaattgaatacaaaaaaaaaaatacttgttatgtattcaattacggg 600
601  taaaatatagacaagttgaaaaatttaataaaaaaataagtcctcagctcttatatatta 660
661  agctaccaacttagtatataagccaaaacttaaatgtgctaccaacacatcaagccggtta 720
721  gagaactctatctatagcaatatitcaaatgtaccgacatacaagagaaaacattaactat 780
781  atatattcaatttatgagattatcttaacagatataaatgtaaattgcaataagtaagat 840
841  ttagaagtttatagcctttgtgtattggaagcagtacgcaaaggcttttttatttgataa 900
901  aaattagaagtatattttttttcataattaatttatgaaaatgaaaggggggtgagcaa 960
961  agtgacagaggaaagcagtatcttatcaaataacaaggatttagcaatatcattattgac 1020
1021  tttagcagtaaacattatgacttttatagtgtttagctagctaagtagtacgaaagggggag 1080
1081  ctttaaaaagctccttggatacatagaattcataaattaatttatgaaaagaagggcgt 1140
1141  atatgaaaacttgtaaaaattgcaaagagtttattaaagatactgaaatatgcaaaatac 1200
1201  attcgttgatgattcatgataaaaacagtagcaacctattgcagtaaatacaatgagtaa 1260
1261  gatgtttacataaagggaaagtccaatgtattaattggtcaaagatgaaccgatatggat 1320
1321  ggtgtgccataaaaatgagatgttttacagaggaagaacagaaaaaagaacgtacatgca 1380
1381  ttaaatattatgcaaggagctttaaaaaagctcatgtaagaagagtaaaaagaaaaaat 1440
1441  aatttatttattaatttaattatgagagtgccgacacagtatgcactaaaaaatatct 1500
1501  gtgggtgtagtgagccgatacaaaaggatagtcactcgcattttcataatacatcttatgt 1560
1561  tatgattatgtgtcgggtgggacttcacgacgaaaaccacaataaaaaaagagttcgggg 1620

```

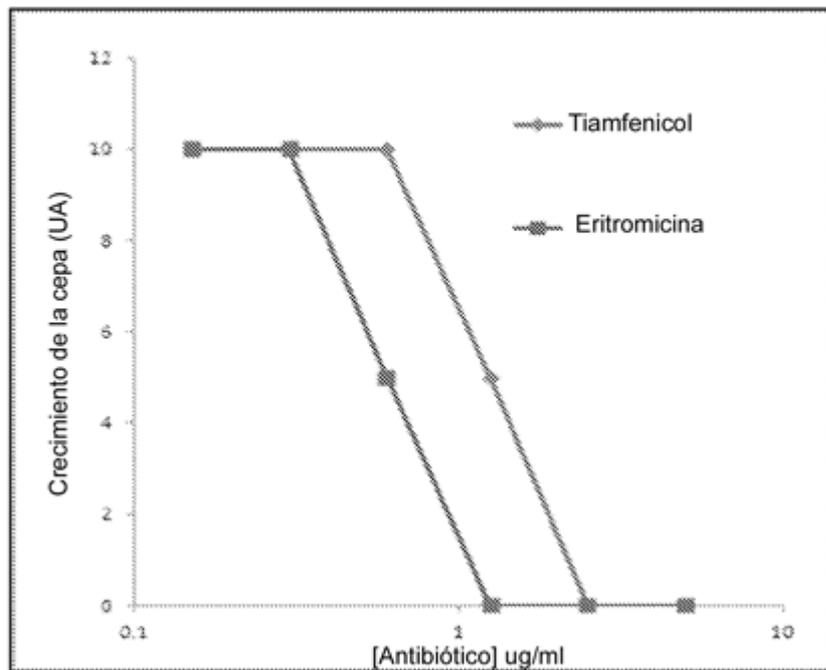
**FIG. 35B**

1621 taggggtaagcatagttgaggcaactaaacaatcaagctaggatagcagtagcagaccg 1680  
 1681 taaggctcgttgtttagggtgtgtgtaatacatacgctattaagatgtaaaaaatcgggata 1740  
 1741 ccaatgaagggaaaaggtataaattttggatgtagtttggtttggttcatctatgggcaaact 1800  
 1801 acgtccaaagccggttccaaatctgctaaaaagtatatcctttctaaaatcaaagtcagg 1860  
 1861 tatgaaatcataaataaagtttaattttgaagttattatgatattatgtttttctattaa 1920  
 1921 aataaattaagtatatagaatagtttaataatagtatatacttaatgtgataaagtgtctg 1980  
 1981 acagtgtcacagaaaaggatgattgttatggattataagcggccggcagtgaggcaagtgtg 2040  
 2041 aaaaattcacaaaaatgtggtataatctttgttcattagagcgataaaacttgaatttg 2100  
 2101 agagggaacttagatgggtatttgaaaaatgataaaaaatagttggaacagaaaagagta 2160  
 2161 ttttgaccactactttgcaagtgtaacctgtacctacagcatgaccgtaaaagtggatata 2220  
 2221 cacacaaataaaggaaaagggaaatgaaactatatacctgcaatgctttatttatattgcaat 2280  
 2281 gattgtaaaccgcccattcagagtttaggacggcaatcaatcaagatgggtgaattggggat 2340  
 2341 atatgatgagatgataccaagctatacaatatttcacaatgatactgaaacattttccag 2400  
 2401 cctttggactgagtgtaagctgactttaaatcatttttagcagattatgaaagtgtatc 2460  
 2461 gcaacgggatggaaacaatcatagaatggaaggaaaagccaaatgctccggaaaacatttt 2520  
 2521 taatgtatctatgataccgtgggtcaaccttcgatggctttaatctgaatttgcaaaaagg 2580  
 2581 atatgattatttgattcctatttttactatggggaaatattataaagaagataaacaanaat 2640  
 2641 tatacttcccttggcaattcaagttcatcacgcagtagtgacggatttcacatttgccg 2700  
 2701 ttttgtaaacgaattgcaggaattgataaaatagtttaacttcagggttggctgtgtaactaa 2760  
 2761 aacaagtatttaagcaaaaacatcgtagaaatacgggtgtttttgttaccctaagtttaa 2820  
 2821 actccttttgataatctcatgaccaaaatcccttaacgtgagtttccgttccactgagc 2880  
 2881 gtcagaccccgtagaaaagatcaaaggatcctccttgagatccttttttctgcccgtata 2940  
 2941 ctgctccttgcaaacaaaaaacaccocgctaccagcgggtgggttggttgcccgaataaga 3000  
 3001 gctaccaactccttttccgaaggtaactggcttcagcagagcgcagataccaaatactgt 3060  
 3061 ccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacata 3120  
 3121 cctcgtctgtctaactcctgttaccagtgctgctgcccagtgccgataaagtctgtcttac 3180  
 3181 cgggttggactcaagacgatagttacoggataagggcgcagcggctcgggctgaacgggggg 3240  
 3241 ttctgtcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcg 3300  
 3301 tgagcattgagaaagcggccacgcttccogaaggggagaaaggcggacaggtatccggtaag 3360  
 3361 cggcagggctcggaaacaggagagcgcacgagggagcttccaggggaaacgcctggtatct 3420  
 3421 ttatagctctgtcgggttctgccaactctgacttgagcgtcgatttttggatgctcgtc 3480  
 3481 aggggggaggagcctatggaaaaacgcagcaacgcggcctttttacggttcctggcctt 3540  
 3541 ttgctggccttttctcaccatgtcttttctgcttataccctgattctgtggataaaccg 3600  
 3601 tattaccgcctttgagtgagctgataccgctcggcgcagccgaacgaccgagcgcagcga 3660  
 3661 gtcagtgagcggaggaagcgggaagagcggcccaatagcagggcccctgcttcggggctcat 3720  
 3721 tatagcgatttttccggtatataccatcctttttcgcacgatatacaggattttgccaag 3780  
 3781 ggttcgtgtagactttccttgggtgtatccaacggcgtcagccgggcaggataggtgaagt 3840  
 3841 aggcccaccccgagcgggtgttcttcttactgtccttattcgcacctggcgggtgct 3900  
 3901 caacgggaaatcctgctctgagaggctggccggctaccgcccggcgtaacagatgagggcaa 3960  
 3961 gcggatggctgatgaaaccaagccaaccaggaagggcagcccacctatcaaggtgactg 4020  
 4021 ccttccagacgaacgaagagcgtattgaggaaaaaggcggcggcggcggcggcggcggcggc 4080  
 4081 ggcctacctgctggccgtcggccagggtacaaaaatcacgggctcgtggactatgagca 4140  
 4141 cgtccgcgagctggcccgcacatcaatggcgacctgggcccgcctggggggcctgctgaaact 4200  
 4201 ctggctcaccgacgacccgcgacggcggggtcgggtgatgccacgatcctcgcctgct 4260  
 4261 ggcgaagatcgaaagagaagcaggacgagcttggcaaggctcatgatggcgtgggtccgccc 4320  
 4321 gagggcagagccatgacttttttagccgctaaaacggccgggggggtgcgctgattgcca 4380  
 4381 agcacgtcccatgcgctccatcaagaagagcgaacttcgcccggagctgggtgaagtacatca 4440  
 4441 ccgacgagcaaggcaagaccgatcggggcccctgca 4476

**FIG. 36A**



**FIG. 36B**



**FIG. 37**

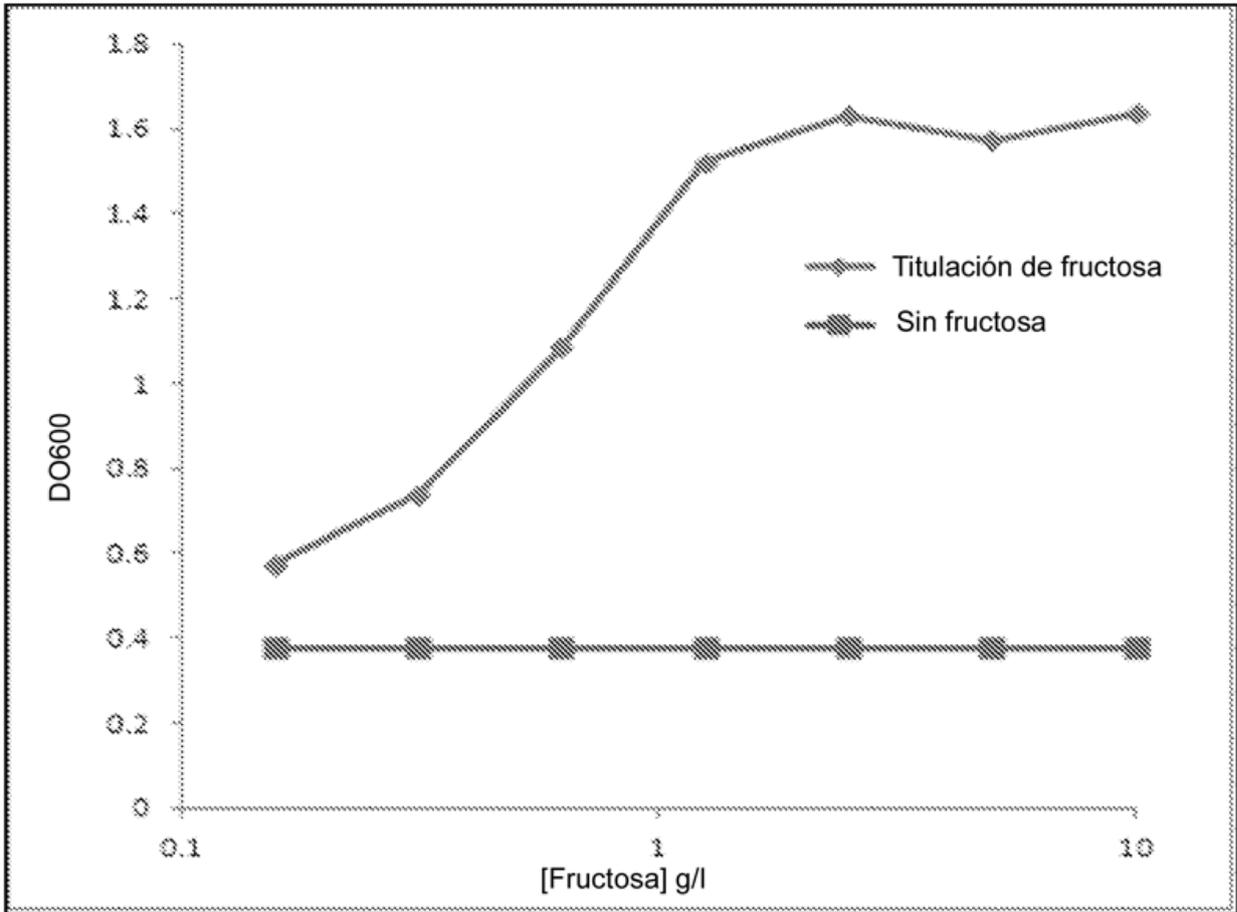


FIG. 38A

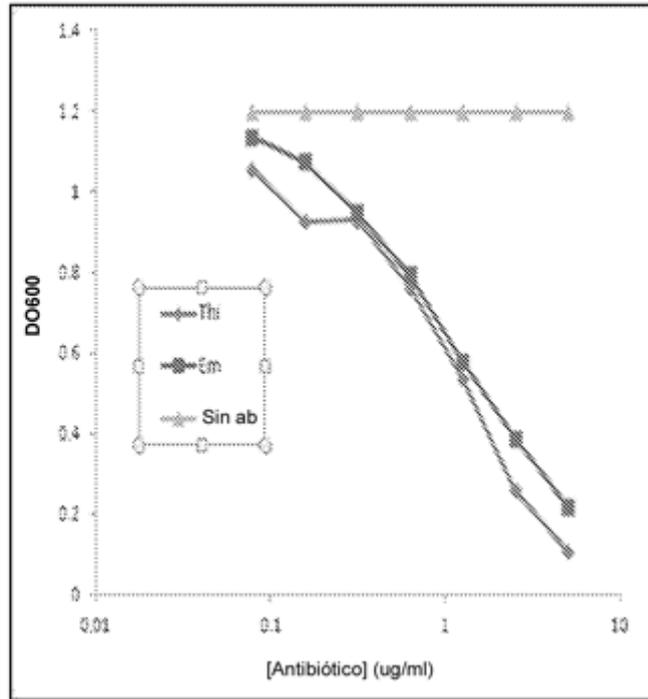
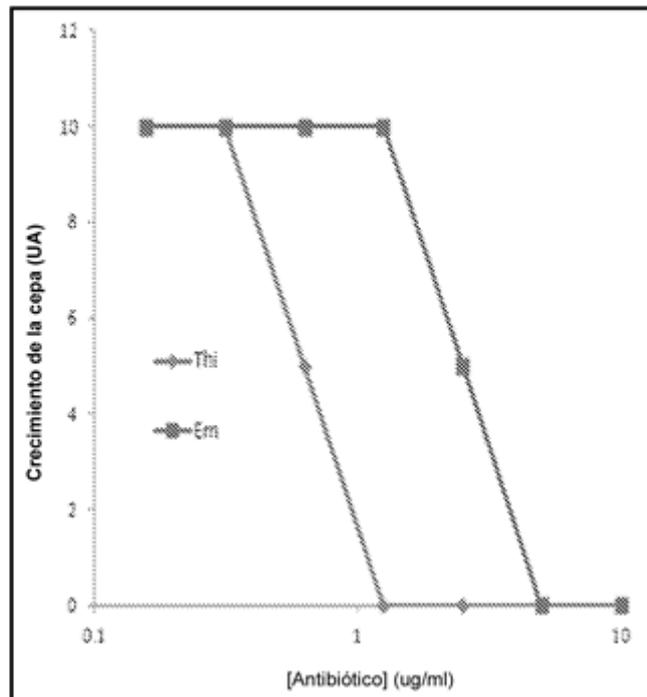


FIG. 38B



**FIG. 39**

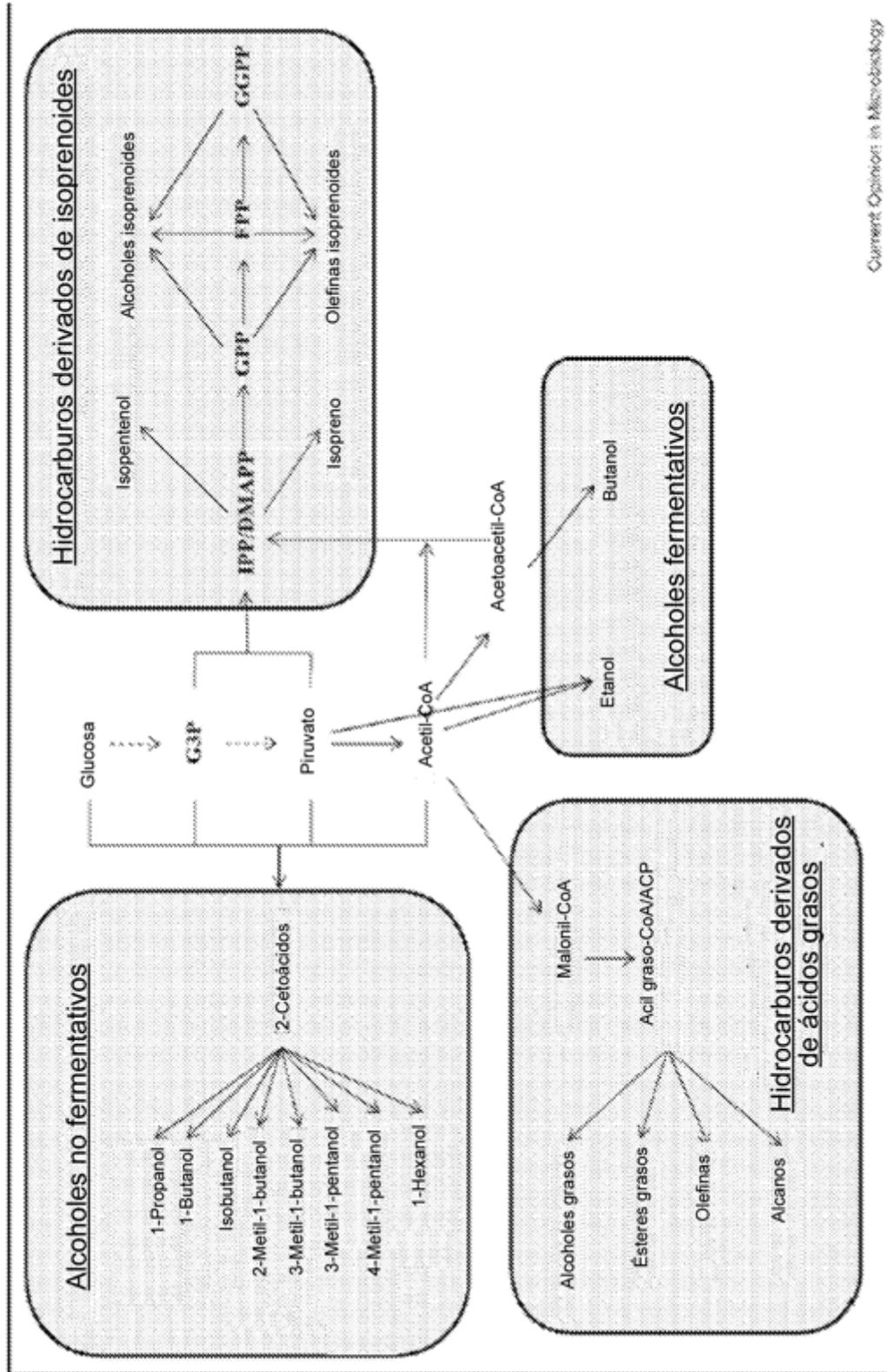


FIG. 40

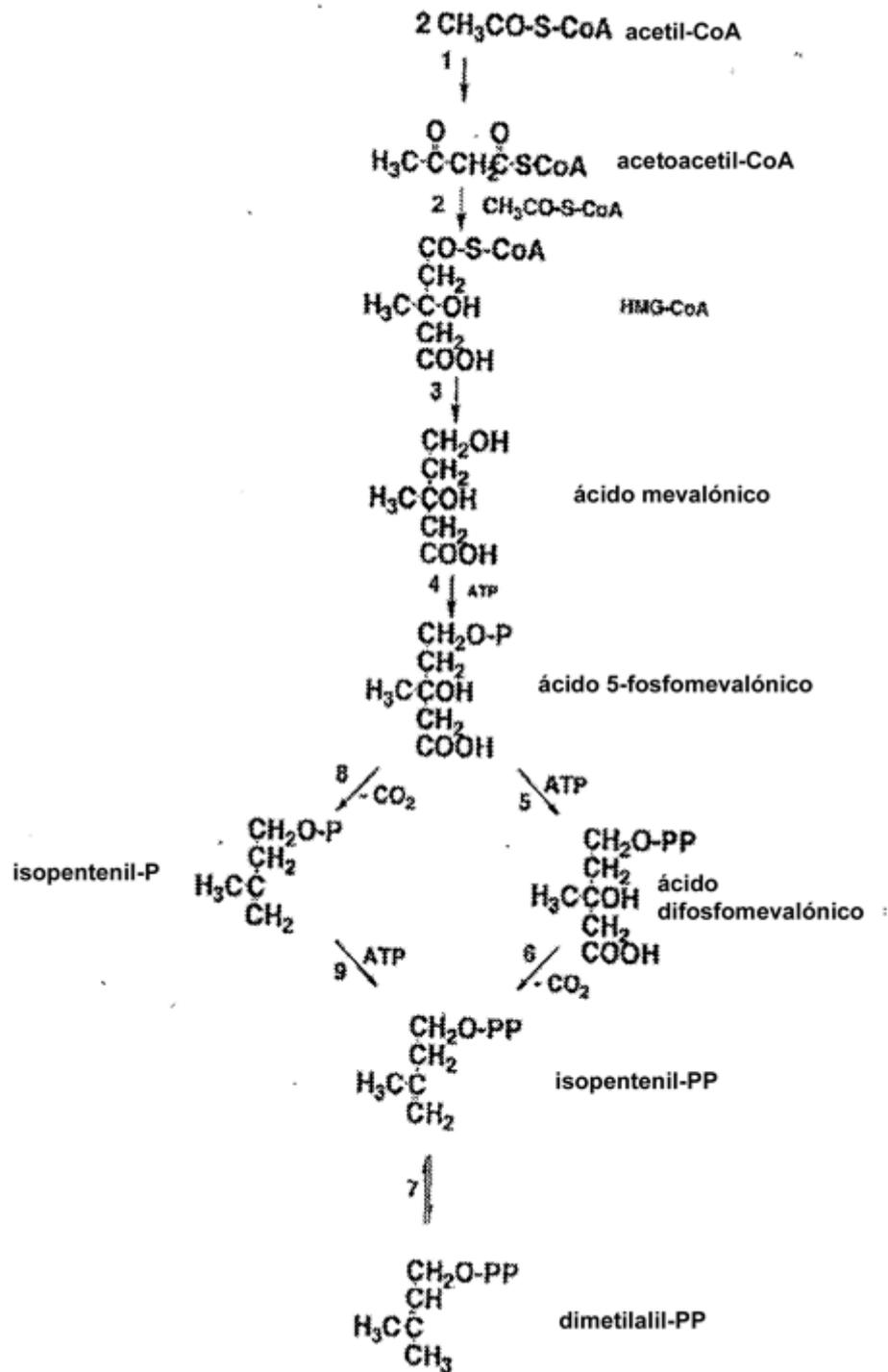




FIG. 42

Expresión de la ruta de MVA superior

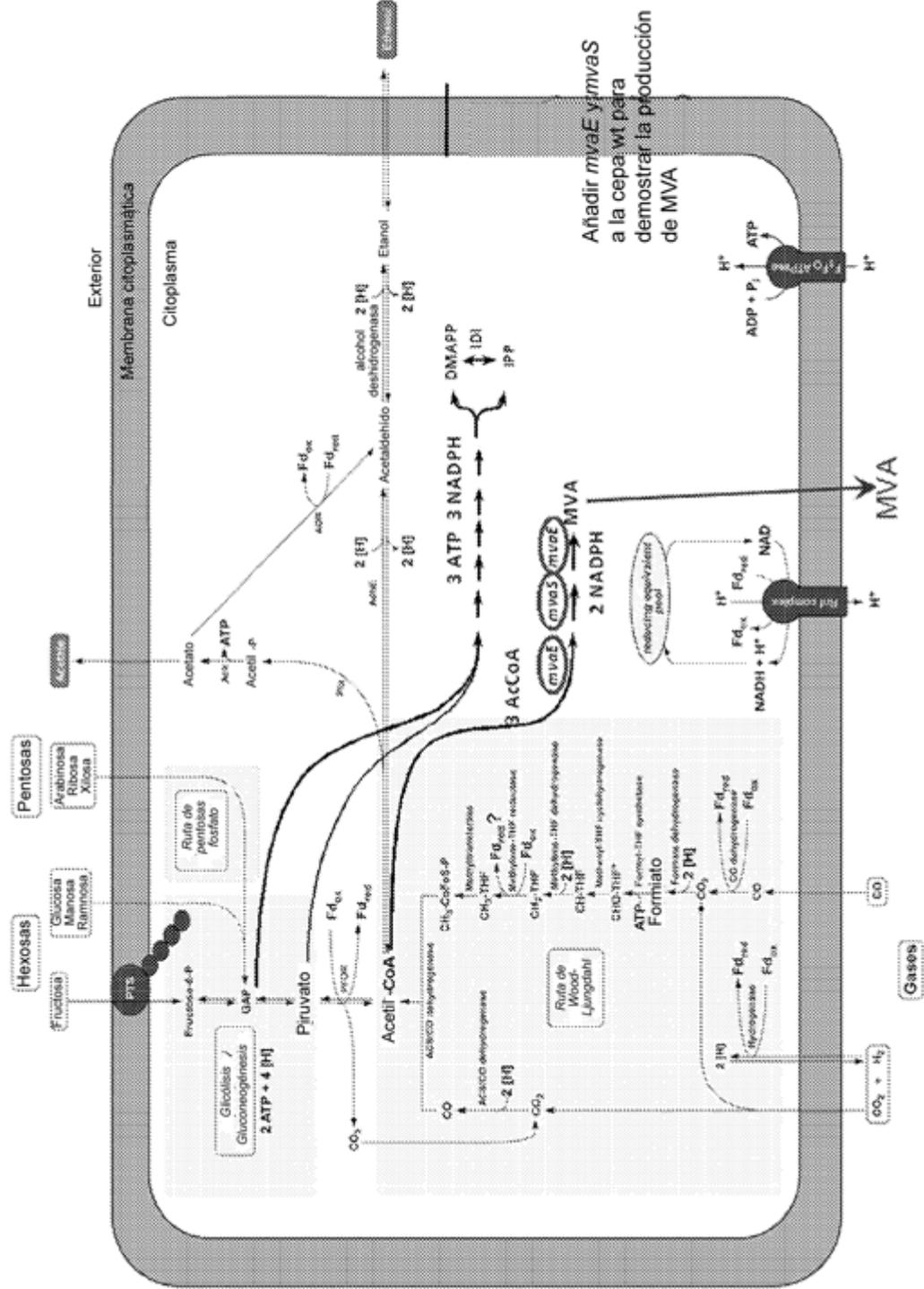
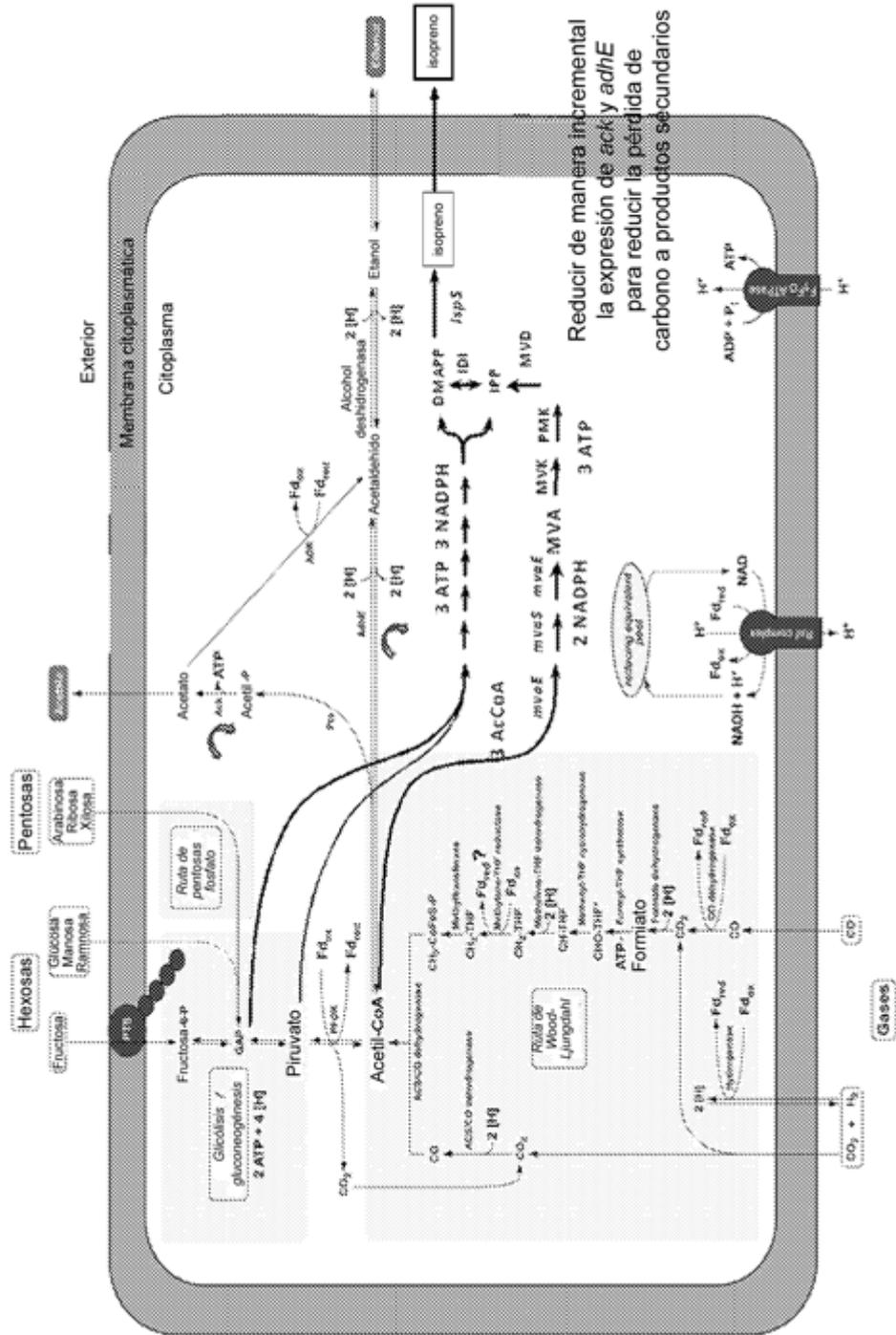






FIG. 45

Redirección del flujo de carbono lejos del acetato



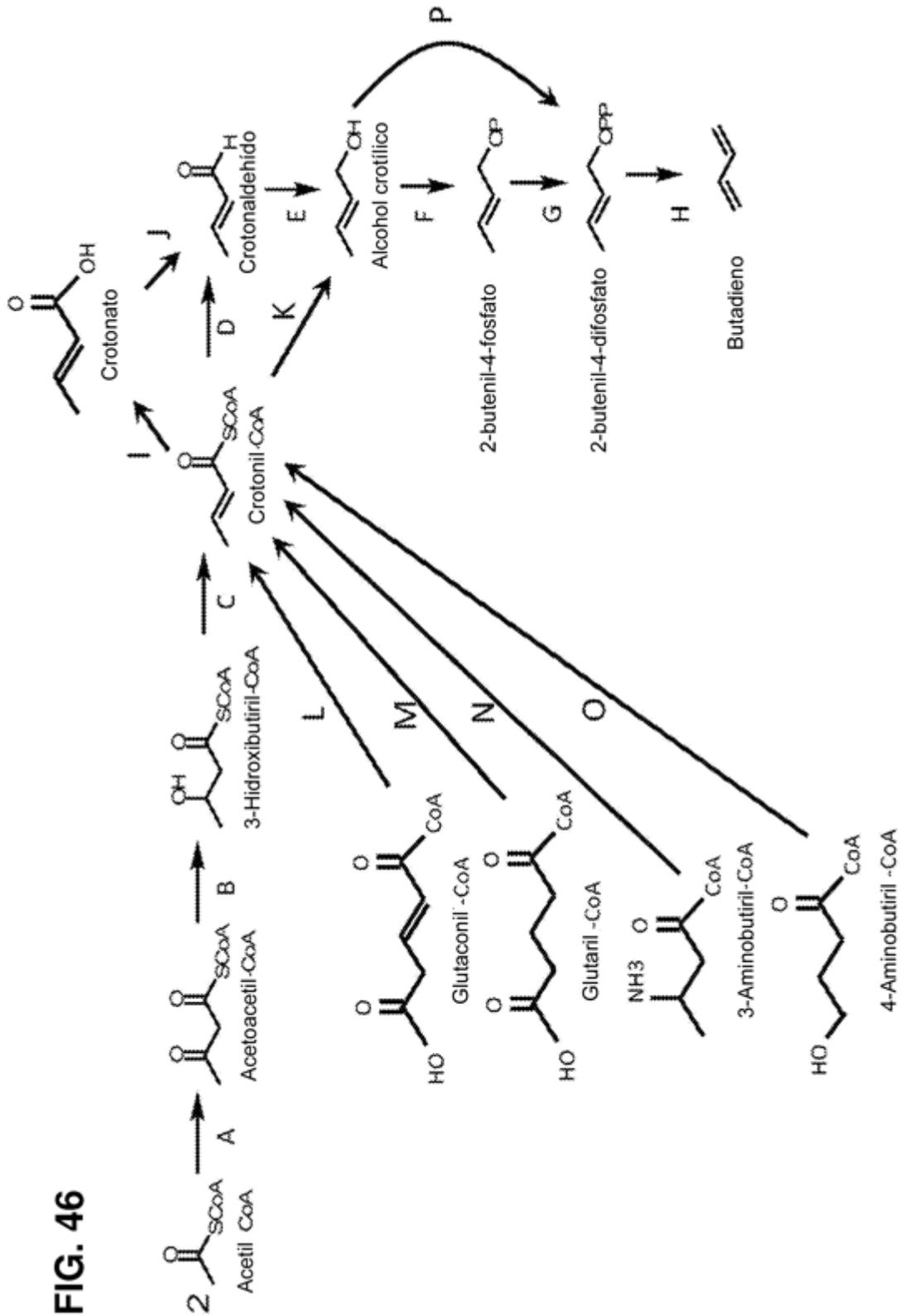


FIG. 46

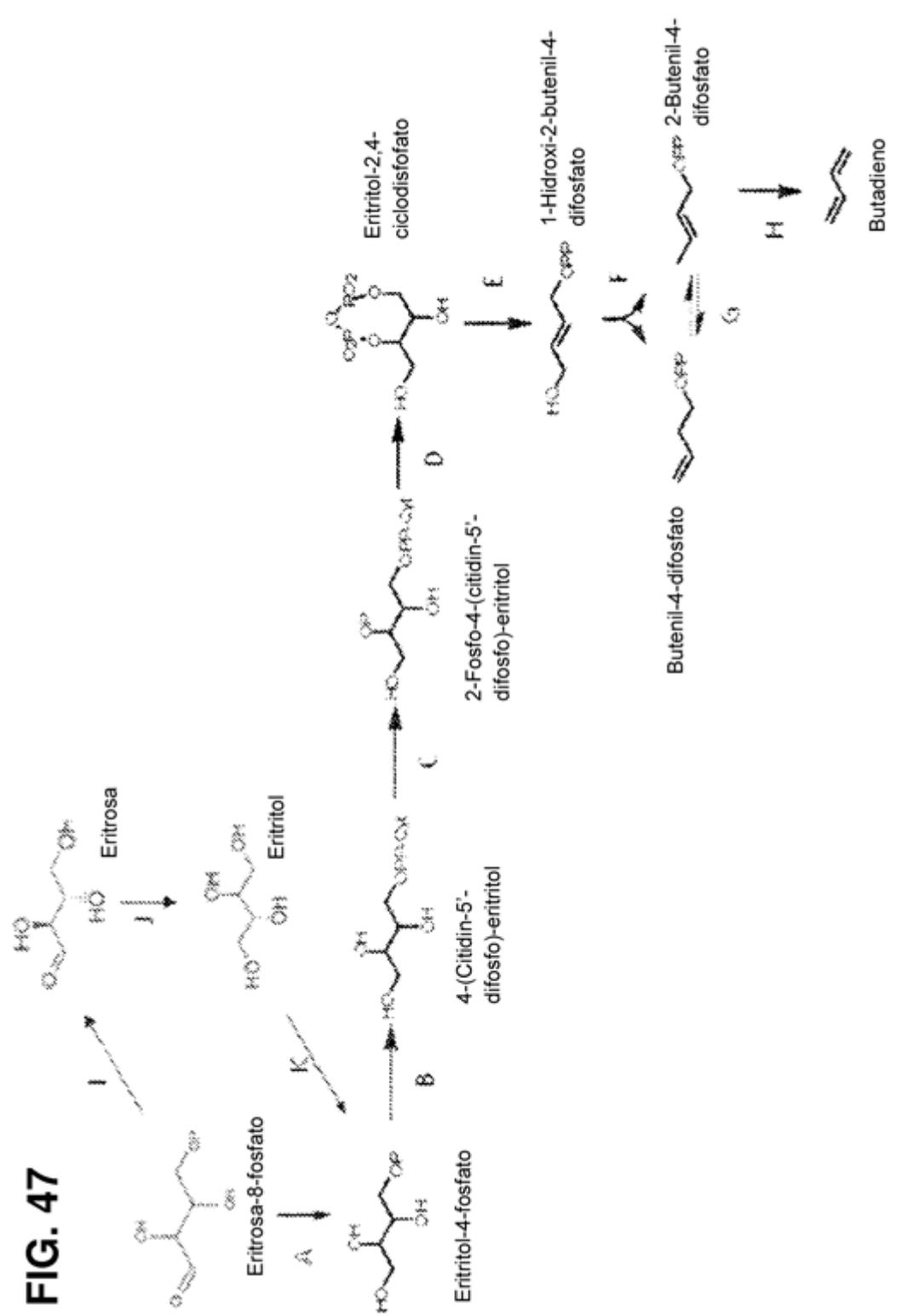




FIG. 49

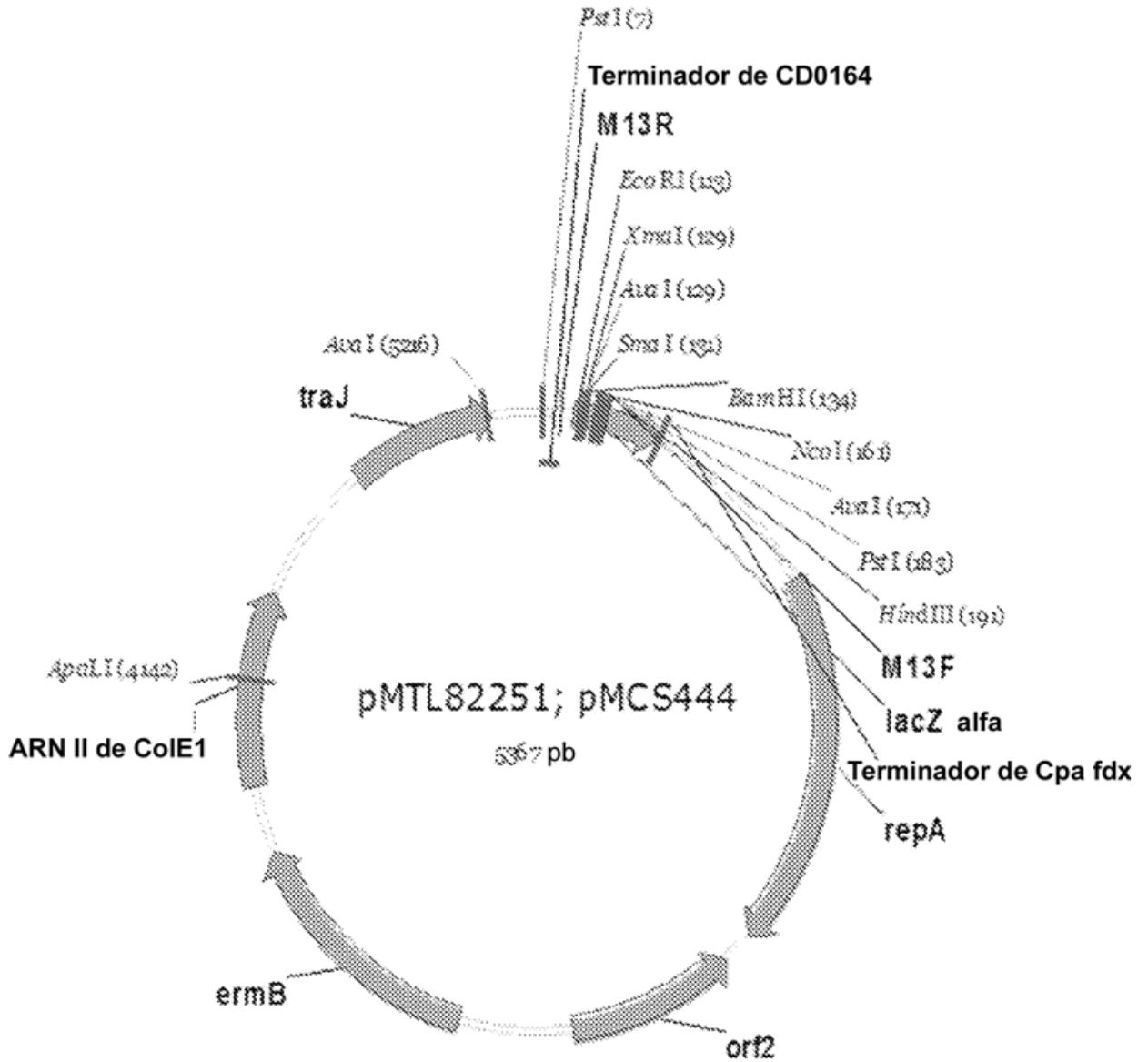
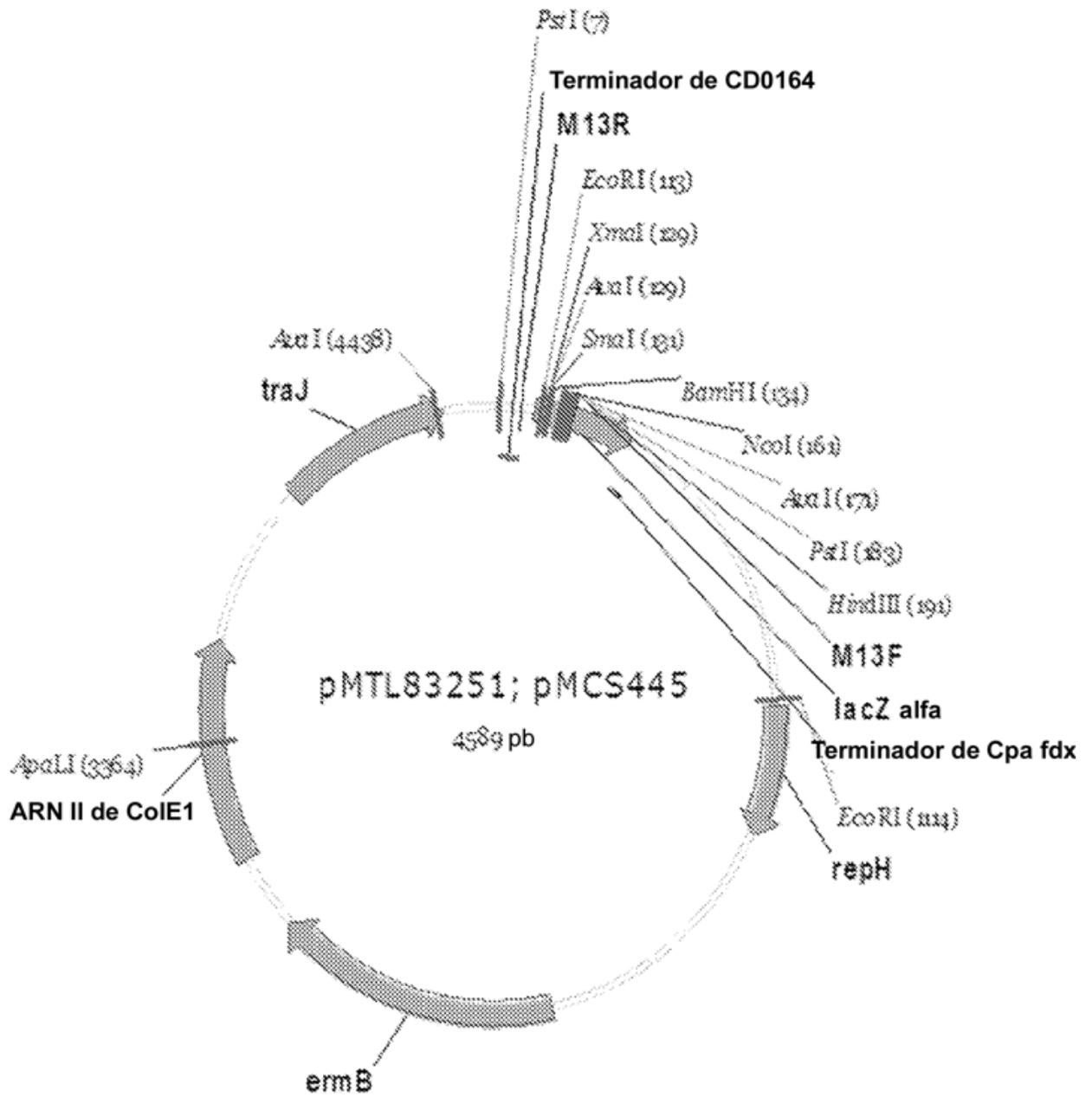




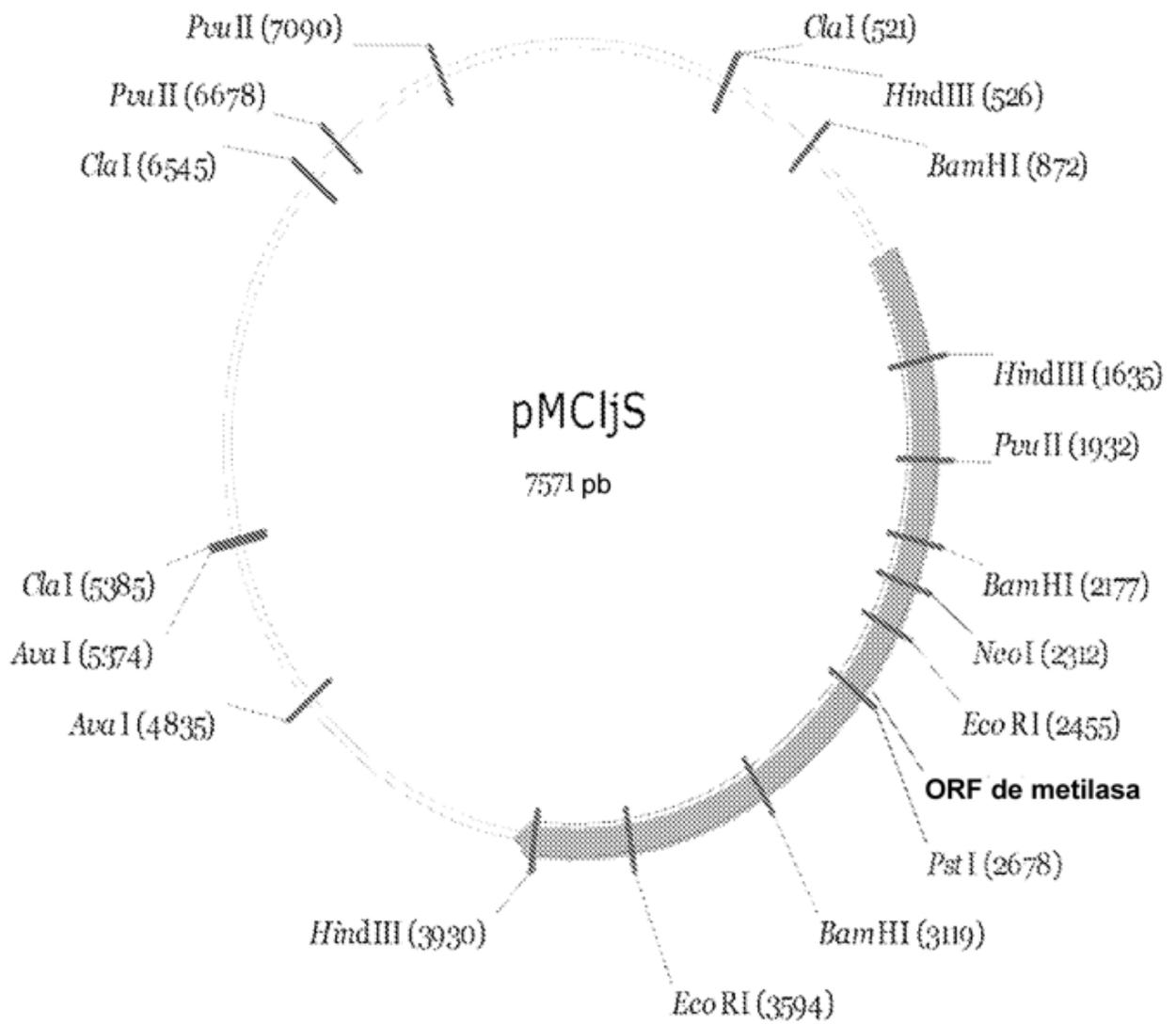
FIG. 51



**FIG. 52**

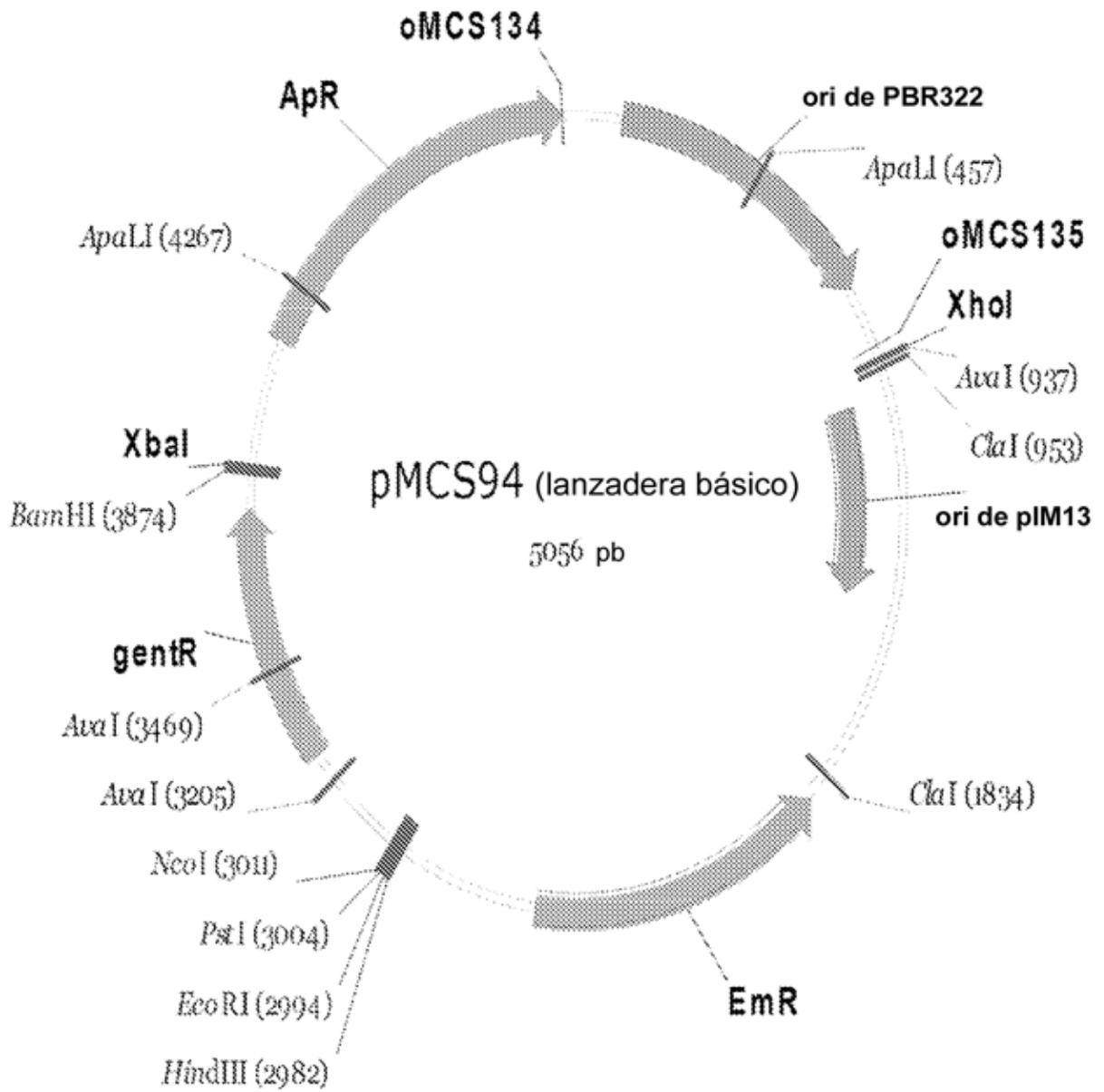
cctgcaggataaaaaattgtagataaaattataaaaatagtttatctacaattttttatcaggaaacagctatgaccgagccgctgtatccatag  
 accatgattacgaattcgagctcggatcccggggatcccttagagtcgacgtcacgctccatggagatctcgaggcctgcagacatgcaagcttggc  
 actggccgtcgttttacaaagctcgtgactgggaaaaccctggcgttaaccaacttaacgcttgcagcacatcccctttccgagctggtgataatgc  
 gaagaggcccgcaccgactgcccttcccaacagttgcgagcctgaatggcgaatggcgttagcataaaaaaagaagcctgattgaggcttctt  
 attttatggcgcgcgcgattatttttgaaacaattgacaattcatttcttttttattaagttagatgcaaaaggcataacagtgctgataagaag  
 aaatttacagaaaaaagaattatagaatttagatgattaattatactcattatgaatgittaattgaaatacaaaaaaaaactgttatgtatccaat  
 tacgggttaaaatatagacaagtgaaaaatttaataaaaaaataagfctcagctcttatattaagctaccaacttagtatataagccaaaactta  
 aatgtgctaccaacacatcaagcgttagagaaactctatctatagcaattttcaaatgtaaccgacatacaagaagaaacattaactatatattcaat  
 ttatgagattacttaacagataaaatgtaaattgcaataagtaagatttagaagttatagcctttgtgtattggaagcagtagcgaagccttttta  
 ttgataaaaattagaagatatttttttcaataattatgaaaatgaaagggggtgagcaaggtgacagaggaagcagtagtatttatcaaat  
 aacaaggtattagcaataatcattatgacttttagcagtaaacattatgacttttatagtgctttagcctaagtagtaagaaagggggagctttaaag  
 ctcttggaaatacagaaattcataaattaattatgaaaagaaggcgtatagaaaacttgtaaaaattgcaaaagatttaataagatactgaaat  
 atgcaaaaatacattcgttagattcatgataaaacagtagcaacctattgcagtaatacaatagtagcaagatgttacataaagggaagctcaat  
 gtattaattgttcaaaagtagaacgatatggatgggtgctcataaaaatgagagtggtttacagaggaagaaacagaaaaaagaacgtacatgcattaa  
 atattatgcaaggagctttaaaaagctcatgtaaaagagtagtaaaagaaaaataattatttataaattaattatgagagtgccgacacagtag  
 gcaataaaaaataatctgtggtgtagtagccgatacaaaaggatagctactcgcatttcaataacatcttagttatgattatggtcggtgggac  
 ttcacgacgaaaaaccacaataaaaaagagttcggggtagggttaagcagtagttaggcaactaacaatcaagctaggatagcagtagcagac  
 cgttaagctgtgttttaggtggttgaatacatacgcctattaagatgtaaaaaatcggataccaatgaaagggaagataattttggatgattgtg  
 tttgttcatctatgggcaactacgtccaagccgtttccaaatctgctaaaagtatatcctttcaaaatcaagtcagtagaaatcataaataaa  
 gtttaatttgaagttattatgatatattgttttctattaaataaataaagtataagaatgttataatagtagtatacttaagtgataaggtctgga  
 caggtgcacagaaaggatgattgttatggattataagcggccggccgaagcaacttaagagtggtgtagatgtagcagtagtctaaatgtgataat  
 aggaattgaaagtaaatagtagtctaaatattgtaattaagaaaggagtgattatgaaacaaaataataaatatctcaaaacttttaacgagtg  
 aaaaagtagtcaaccataataaaacaattgaatttaaaagaaaccgataccgtttacgaaattggaaacaggtaaagggcatttaacgacgaaac  
 tggctaaaataagtaaacaggttaacgtctattgaaattagacagtagtctatcaactatcgtcagaaaaataaaactgaatactcgtgctacttaat  
 tcaacaagataattctacagttcaattccctaacaacagaggtataaaatgttgggagtagtcttaaccatttaagcacaacaattattaaaaagtg  
 gttttgaaaagccatgctgtagcatctatctgattgtgaaagaggatctcaaacgtagccttggatattcaaccgaaactagggttctctgcaac  
 tcaagctcagattcagcaatgcttaagctgcccagcgaatgcttcaactcaaaacaaagtaaacagtagtcttaataaaacttaaccgccataccac  
 agatgttccagataaattggaaagctatatacgtactttgtttcaaaaatgggtcaatcgagaatactgcaactgttactaataaaatcagtttcaaa  
 gcaatgaaacacgcaaaagtaaaacaattaaagtagtaccgttactatgagcaagtagtcttttaatagttatctatttttaacgggaggaaataat  
 tctatgagtcgtttgttaaatttggaaagttacagcttactaaagggaatgtgttaaacctctttttagtaaatctatgaccaaaatcccttaacgtga  
 gttttcgttccactgagcgtcagaacccgtagaaagatcaaaaggatcttcttgagatcctttttctgctgtaactcgtcgtcttgcacaacaaaaaa  
 ccaaccgctacagcgggttgggttccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagatcaaaatactgttc  
 ttctagtagcctgtaggcaaccactcaagaactctgtagcaccgctacataactcgtctgtaactcgttaccagtggtgctgctgcaagtgcc  
 gataagctgctcttacgggttggactcaagacgataagttaccggataagcgcagcggctcggctgaacgggggggtcgtgcacacagccagcctt  
 ggagcgaacgacctacaccgaaactgagatacctacagcgtgagctatgagaagcgcacgcttccgaaagggaagaaaggcggacaggtatccgg  
 taagcggcagggctggaacaggagagcgcacgagggagctccaggggggaaacgcctgggtatctttatagtcctgctgggttccgcaacctctgactt  
 gagcgtcgattttttagtctgctcagggggggcggagcctatggaaaaacgcccagcaacgccccttttaacggttcctggtcttttctggccttttgc  
 tcaatgttcttctcgttatccctgattctgtagataaccgtattacgcttttagtgagctgataaccgctcgcggcagccgaaacgacagcagcga  
 ggcagtagttagcgaaggaaagggaagagcggcaataacgagggggccctgcttgggggtcattatagcgaatttttccggtataatccatcttttccg  
 cagatatacaggaattttgcaaaagggttcgtgtagacttctcttgggtatccaacggcgtcagccgggaggaataggtgaaagtaggcccacccgag  
 agcgggtgttctcttctactgctcttattcgcaccctggcgggtctcaacgggaatcctgctcgcagggctggcgggtaacgcccggcgaacagatg  
 aaaaaaaatgaaaccagccaaccaggaaggcagcccaactcaagggtgtagcttccagacgaaacgaaagcagattgag

**FIG. 53**





**FIG. 55**





**FIG. 57**

